Cognitive Processing in Second Language Acquisition

Edited by Martin Pütz
and Laura Sicola

CONVERGING EVIDENCE IN LANGUAGE AND COMMUNICATION RESEARCH 13

JOHN BENJAMINS PUBLISHING COMPANY

Cognitive Processing in Second Language Acquisition

Converging Evidence in Language and Communication Research (CELCR)

Over the past decades, linguists have taken a broader view of language and are borrowing methods and findings from other disciplines such as cognition and computer sciences, neurology, biology, sociology, psychology, and anthropology. This development has enriched our knowledge of language and communication, but at the same time it has made it difficult for researchers in a particular field of language studies to be aware of how their findings might relate to those in other (sub-)disciplines.

CELCR seeks to address this problem by taking a cross-disciplinary approach to the study of language and communication. The books in the series focus on a specific linguistic topic and offer studies pertaining to this topic from different disciplinary angles, thus taking converging evidence in language and communication research as its basic methodology.

Volume 13

Cognitive Processing in Second Language Acquisition.
Inside the learner's mind
Edited by Martin Pütz and Laura Sicola

Cognitive Processing in Second Language Acquisition

Inside the learner's mind

Edited by

Martin Pütz
University of Koblenz-Landau

Laura Sicola
University of Pennsylvania

John Benjamins Publishing Company
Amsterdam / Philadelphia

Library of Congress Cataloging-in-Publication Data

Cognitive processing in second language acquisition : inside the learner's mind / edited by Martin Pütz, Laura Sicola.

p. cm. (Converging Evidence in Language and Communication Research, ISSN 1566-7774 ; v. 13)

Includes bibliographical references and index.

1. Second language acquisition. 2. Cognitive learning. I. Pütz, Martin, 1955- II. Sicola, Laura.

P118.2.C635 2010

418.001'9--dc22 2010000584

ISBN 978 90 272 3902 0 (Hb ; alk. paper)

ISBN 978 90 272 8832 5 (Eb)

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands

John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Acknowledgements

This collection of 18 papers has its origin in the 33rd International LAUD Symposium, held in Landau, Germany, at the University of Koblenz-Landau, in March 2008.

We owe a debt of gratitude to a large number of people who have contributed to this volume. First of all, many thanks are due to the participants in the symposium for the stimulating discussions in a very pleasant atmosphere, and to the contributors of the present volume, who have responded with alacrity and professionalism to all the requests that have been made of them. Furthermore, the editors of the volume would like to thank all persons who kindly agreed to act as referees and linguistic experts: Salvatore Attardo, Kees de Bot, Anna Cieślicka, David Deterding, René Dirven, Helen Fraser, Stephan Gries, Beate Hampe, Michael Harrington, Juliane House, Jan H. Hulstijn, Ulrike Jessner, Janice Johnson, Ronald Leow, Seth Lindstromberg, Jeannette Littlemore, Shawn Loewen, Geoffrey Nathan, Shannon Sauro, Farzad Sharifian, Dan Slobin, Paul Toth and Marjolijn Vespoor.

Thanks to the generous support by the DFG (German Research Council), the University of Koblenz-Landau, the Ministry of Education and Culture of the Rhineland-Palatinate, the Gillet Foundation (Edenkoben) and the Friends and Supporters of the University of Koblenz-Landau (Landau Campus), many internationally well-known scholars were able to participate in the symposium. We would also like to thank the organizing staff of the symposium, in particular Isabel Zehfuß, Laura Bornemann, Ulrike Laux, and Jens Ruckdäschel for their enthusiasm and kind assistance.

Thanks are also due to Carmen Fischer for proofreading major parts of the manuscript.

Finally, we owe a debt of gratitude to Hanneke Bruintjes (Acquisition Editor, John Benjamins, Amsterdam), Martine van Marsbergen (Production Department, John Benjamins, Amsterdam) and Marjolijn Verspoor (Editor of the CELCR series).

Martin Pütz — University of Koblenz-Landau, Landau Campus
Laura Sicola — University of Pennsylvania, Philadelphia

INTRODUCTION

Cognitive processing in second language acquisition

Martin Pütz and Laura Sicola

The present volume grew out of the 33rd International LAUD Symposium, held in March 2008 at the University of Koblenz-Landau in Landau, Germany. This edited volume represents state of the field research linking cognition and second language acquisition. The aim of the volume is to offer new and insightful research on cognitive issues from the learner's perspective as can be observed in second language acquisition contexts. Its unique perspective reflects the experience of the learner when engaged in noticing, input/output processing, retrieval, and even attrition of target forms. Contributions are both theoretical and practical; they include a range of target forms in the areas of lexis, morpho-syntax, phonology and pragmatics; and represent more than a dozen L1, L2 and L3 combinations from Europe, Asia, and the Americas. Spoken and written languages are investigated in face-to-face, paper-based, and computer-mediated contexts.

The psycholinguistic and cognitive processes underlying the learning of a foreign or second language have always been a subject of major interest to both (i) Second Language Acquisition (SLA) researchers and (ii) practitioners involved in language pedagogy, i.e. second and foreign language teaching. Most SLA theories have thus far assumed a top-down operation from Universal Grammar to L1 and L2 grammars. Just as for first language acquisition, it has historically been assumed that, given sufficient input, a second language system develops in the mind automatically. The L2 input is said to trigger the language acquisition device, which is seen to operate quite autonomously, that is, independently from and without any interaction with other cognitive abilities or faculties such as bodily experiences, image schemas, perception, attention, categorization, emotion, or still other mental faculties.

In contrast, most of the chapters in this volume take a radically opposing view in the sense that they focus more on the essential role of bottom-up processing in SLA. It is argued that language acquisition and learning can, like all human learning processes, only be understood and explained if they are seen as bottom-up, exemplar-based and usage-oriented processes. These can and have to be accommodated in a much broader framework of how people interact with the world around them, store and acquire knowledge in some symbolic form or other, and thus establish a link between cognitive development and language acquisition.

The book explores the even more complex process of L2 learning and acquisition from three different inroads:

i. Cognitive theoretical foundations of language and learning,
ii. Mental processes and acquisition procedures followed by language learners, and
iii. Cognitive language pedagogy: Classroom studies with applications for teaching.

Section 1

We begin our investigation into the mind of the learner from a theoretical perspective in Section 1, "Cognitive theoretical foundations of language and learning". Whether dealing with a second, foreign, or additional language, *per se,* as are represented in the contributions of this first section, the opportunity to identify and even visualize models that can delineate the many possibilities for processing everything from lexis and morpho-syntax to pragmatics can only help provide a variety of contexts within which to ponder the findings offered in the more empirical studies in subsequent sections.

To start with, **Mark Fifer Seilhamer** introduces a general discussion of relevant theories and models and elaborates the ideas of concept-stretching and model merging which provide an attempt to account for L2 processing and acquisition of grammatical constructions. The purpose of the paper as stated by the author is to "stretch" and enrich the Evidential Access Model and the Local Coherence Model inherent in Jurafsky's theory of grammatical construction access and disambiguation, by merging it with two other approaches (Abdel's Dual Idiom Representation Model and Chang's model of grammatical construction learning in L1 child language acquisition) so as to create a new integrative model capable of better explaining L2 processing and L2 learning.

Subsequently, **Nick Ellis** sets out to explain that learners' understanding of language, and of how it works, is based upon their experience of language and that learners have to estimate the system from a sample of usage. He specifically shows how acquisition is affected by the frequency and frequency distribution of exemplars within construction, by their prototypicality, and, using a variety of psychological and corpus linguistic association metrics, by their contingency of form-function mapping. The greater the token frequency of an exemplar, the more it contributes to defining the category and the more it is considered the prototype. Frequency, distinctiveness and prototypicality have become associated in language and thus in learning. Ellis shows that psycholinguistics, cognition, second language acquisition, and corpus and cognitive linguistics are all necessary for our understanding of how learners construct a second language.

As cognition plays a role in all areas of language development, **Maria Jodłowiec** investigates the role of pragmalinguistic theories in SLA studies, i.e. an account of the cognitive processes that are at play during verbal communication, also from a cognitive perspective. Specifically, the paper seeks to explore interfaces between a pragmalinguistic model, namely Sperber and Wilson's relevance theory, and processes underlying the development of competence in L2. The author demonstrates how the model affords significant insight into psycholinguistic and sociolinguistic processes underlying the development of competence in a second/foreign language and explains the way it also illuminates important practical aspects directly related to target language instruction.

From there, **Bilal Kırkıcı** tests the applicability of the Dual-Mechanism Model (for L1 processing) to L2 processing, particularly with regard to the development of regular and irregular English past tense morphology. L1 Turkish learners of English at both advanced and beginner levels, as well as L1 English speakers all performed similar production tasks including real and nonce English verbs. Patterns of how native and non-native speakers process regular and irregular verbs lend support to the applicability of the Model in the L2 context from an early stage of development.

Venturing into the realm of the mental lexicon, **Ksenya Filatova** applies a cognitive linguistics framework to third language (L3) acquisition on a multilinguistic continuum of macrocategories. She presents both cognitive and neurophysiological evidence relating to the interplay of these categories and the changing roles of the L1, the L2 and other languages in the acquisition of additional foreign languages. Her case study of the experience of L1 Russian university students learning Spanish demonstrates the relevance of her model of synonymic categories across languages.

Section 2

Filatova's chapter provides a doorway into Section Two, i.e. "Mental processes and acquisition procedures followed by language learners". In this section, theoretical models from Section One will give way to the data of empirical studies, as we explore more concrete evidence of what occurs in the mind of actual learners under a variety of conditions.

To begin this section, **Rebekah Rast** looks at the behaviour of learners at a stage that is particularly difficult to capture for research purposes: that of true first exposure to an L2. She frames her study of L1 French students' first exposure to Polish in the context of establishing the overlap and distinction between the notions of input and intake in a historic sense, and establishing how the current operationalizations thereof help to explain the selectivity of the intake process. The sentence repetition test and sentence translation test broadly illuminate the influence of phonological, morphological, syntactic and other factors on what input is more likely to be converted to intake in the earliest stage of development.

Narrowing the focus back to the mental lexicon, **Kees de Bot** and **Wander Lowie** convincingly argue that lexical representations are not stable entities on which operations can be carried out, but meaning bearing structures that constantly change with use. Data from a simple word naming task in English and Dutch with advanced bilinguals show that even for a selected set of stimuli the stability of representations as measured by correlations between individuals and between sessions within an individual is low. This argues for a dynamic and episodic rather than invariant perspective on the multilingual lexicon which may have substantial consequences for research on this topic.

Following this theme, in their article on the blended bilingual lexicon, **Wander Lowie, Marjolijn Verspoor** and **Bregtje Seton** adopt a usage-based approach which means that words are conceptualized as dynamic constructs, based on both physical and linguistic experience, and therefore have individual and cultural associations. The authors have tested this assumption by means of a number of paper and pencil word association tests

and two related response time experiments on Dutch advanced learners of English and English advanced learners of Dutch. Both experiments clearly show that even though the advanced learners move towards more native-like L2-concepts, a clear vestige of the L1 concepts remains.

Broadening the focus from the lexical unit, in her contribution, **Anna Cieślicka** addresses the question of how Polish second language learners of English process formulaic language – more specifically, idioms. To this purpose she seeks to verify the so called "superlemma theory of idiom production", which is said to be a hybrid model, assuming both a unitary and a compositional nature of formulaic expressions. Her result supports the view that literal analysis of an idiomatic string is obligatory in L2 idiom production, especially in light of the fact that the design of the experiment might allow participants to ignore the idiom and focus on the naming task instead.

Subsequently, **Aline Godfroid, Alex Housen** and **Frank Boers** examine the role of 'noticing' in incidental L2 vocabulary acquisition and present evidence from an eye-tracking study. The authors define 'noticing' as a cognitive process in which the amount of attention paid to novel or not fully acquired language forms in the input exceeds a critical threshold, thus turning these forms into intake and, as such, into candidates for further processing and long-term storage (uptake). The aims of their experiment in which four groups of Dutch-speaking students of English participated were to gauge (a) the occurrence of noticing events during silent reading, (b) the role of the context in which novel words are encountered, and (c) the degree to which learners might differ in their inclination for noticing.

Shifting the analysis a step further, **Juana Marín Arrese** presents a case study comparing the impersonalization strategies used by L1 Spanish learners of English when translating newspaper articles from L1 to L2, and from L2 to L1, when defocusing agency of an event. Data revealed a variety of construal patterns employed by the participants, reflecting differences in discourse-pragmatic features of the respective languages. Emergent patterns provide insight into how learners perceive the construal of events in English and Spanish.

In contrast to the above studies, **Merel Keijzer** takes a unique angle in her analysis of cognitive processes in L2 acquisition by juxtaposing it with L1 attrition. She compares two language systems in flux, characterized by advanced L1 acquisition in monolingual Dutch-speaking adolescents and the L1 of first-generation Dutch emigrants in Anglophone Canada. Based on the findings of an empirical study (controlled language tasks, spoken data), the outcomes revealed that the interlanguage of attriters (i.e. the emigrants) showed clear correspondences to the output of the adolescents, as opposed to the Dutch control subjects'. So-called mirror symmetries were identified where those features that are acquired late in Dutch-speaking adolescents also proved vulnerable to attrition. In line with findings from cognitive linguistics, it was found that the usage-based concepts of exemplar-based language learning, entrenchment and analogy could account for many of the study's results.

Section 3

Having established a variety of frameworks and evidence representing many of the processes in which learners engage on the journey of L2 acquisition, the question must be posed regarding how these ideas can and do relate to the experience of learners in a more formal classroom setting. As such, Section Three, "Cognitive language pedagogy: Classroom studies with applications for teaching", will explore this issue.

In his paper, **Peter Robinson** argues that a weaker form of the "Cognition Hypothesis" (Cromer) has much to offer in explaining the course and extent of cross-linguistic influence on second language acquisition, i.e. to calling for proposals for task-based language teaching materials and syllabus design in the L2 classroom. The fundamental pedagogic claim of the Cognition Hypothesis, Robinson argues, is that pedagogic tasks should be sequenced for learners in an order of increasing cognitive complexity. Furthermore, this promotes L2 development and improvements in the ability to perform target tasks in the L2. To illustrate his methodology, he describes the results of a study involving three interactive tasks performed by Japanese L1 speakers of English. The results of the study are largely consistent with the claims of the Cognition Hypothesis.

From a cognitive linguistics perspective, **Luna Filipović** and **Ivana Vidaković** address the issues that arise when second language acquisition is studied within the context of a semantic typology. With reference to Talmy's typology, the authors maintain that languages offer different frameworks to speakers in the process of lexicalizing experience. On the basis of how motion events are lexicalized in English and Serbian, these two languages are said to belong to the same typological group, unlike e.g. Spanish, which exhibits a different lexicalization pattern. Seeing this typology as a cline rather than a dichotomy, the authors assess the importance of intra-typological differences between English and Serbian for second language acquisition. Their central focus is to estimate the extent to which this typologically driven study could help predict and explain learners' errors and progress in competence.

This leads to the cognitive issue of metaphor, or figurative language, which is taken up by **Jeannette Littlemore** who fills a gap of research by investigating the extent to which language learners are able to transfer their metaphor interpretation and production skills from their mother tongue (L1) to the target language (L2), i.e. the way metaphoric competence develops in the bilingual lexicon. To this purpose, she analyzed four dimensions of metaphoric competence in upper-intermediate French-speaking university students of English studying at a university in Belgium: (i) the tendency to find meaning in metaphor, (ii) the speed in finding meaning in metaphor, (iii) the ability to identify multiple interpretations for a given metaphor, and (iv) novel metaphor production. Littlemore found that scores on all four tests in the L1 correlated with scores on the equivalent test in the L2. In other words, if students displayed a high level of metaphoric competence in the L1, they also displayed a high level of metaphoric competence in the L2.

From there, in their exploration of an area that challenges even the most advanced learners, **Katarzyna Bromberek-Dyzman** and **Anna Ewert** set out to investigate how Polish monolinguals and bilinguals comprehend and interpret figurative meanings with particular reference to conversational implicatures (Grice). From a theoretical perspective they argue that the two languages of an L2 user are interconnected in multicompetence and inextricably connected with general cognitive processing. Specifically, they try to

empirically answer the question of how L2 figurative competence interacts with L1 figurative competence. The results showed that advanced L2 learners may achieve a similar level of figurative proficiency in both languages since monolinguals and bilinguals interpreted the unsaid, intended meaning in L1 in a similar manner.

Taking a more explicit task-based approach to cognitive processing in L2 instruction and learning, **Laura Sicola** explores the potential of having learners draw each other's attention to targeted phonological forms through dyadic tasks. Although two-way information gap tasks have more commonly been used to draw learners' attention to lexical or morpho-syntactic targets, Sicola demonstrates how conditions established by such pedagogic tasks can provide ample opportunity for learners to engage cognitive processes of noticing and awareness by triggering the exchange of corrective feedback and increasing the need for phonologically modified output, having an effect on productive and perceptual accuracy.

Nowadays, a compendium such as this would be remiss not to acknowledge the ever-growing popularity of learning in a 'virtual classroom' space. As such, **Susana M. Sotillo** investigates the cognitive process of noticing as well as other psycholinguistic factors such as awareness and attention in computer-mediated interaction, specifically, synchronous computer-mediated communication (SCMC) compared to face-to-face contexts. She identifies how language learners are able to notice linguistic gaps in their knowledge as they shift attention from meaning to form in the different interaction settings. Sotillo's study is based on data from chat logs and transcribed tape recorded exchanges via Yahoo Instant Messenger between tutors and ESL learners as they collaborated on five learning tasks. It examines noticing as tutor- and learner-initiated language-related episodes, the type and quality of corrective feedback, learner uptake and successful uptake.

Conclusion and outlook

The range and depth of exploration presented in these pages hold powerful implications for future research and pedagogical application. While there may indeed be some top-down processing involved, the ideas and evidence provided here lend undeniable weight to the argument that there are essential bottom-up cognitive processes that learners undergo upon acquiring second and additional languages. This volume has taken some great steps towards identifying how these processes work, and demonstrated ways in which they can be exploited in more formal language learning contexts by teachers and students alike, but we have just begun to scratch the surface.

The challenge now is to push the metaphorical envelope. To what extent can the theoretical models be honed and tested? What might be the result of replicating the more empirical studies with learners of different L2s, from different L1 backgrounds, or of different ages or proficiency levels, for example? And how can practitioners and motivated learners apply this knowledge in such a way as to maximize their success in different learning contexts? We hope that the information presented in these pages inspires further enquiry into what learners attend to and notice, what input converts to intake and uptake, why, and how, and to what extent our knowledge of what occurs in the learners' minds can be translated into tools for success that can be placed into their own hands.

SECTION I

Cognitive theoretical foundations of language and learning

CHAPTER 1

Concept stretching and model merging

An attempt to better account for L2 processing and acquisition of grammatical constructions

Mark Fifer Seilhamer
National University of Singapore

1. Introduction

The idea of concept stretching, the expansion of a given conceptualization in order to account for a wider range of phenomena, is likely to occur in any field of inquiry. According to Paivio (1986), even mathematics, arguably the most formal and rule-governed of all disciplines, has seen its share of concept-stretching, with mathematicians in the early 20th century constantly extending the definition of *polyhedra* (three-dimensional forms) in order to account for increasingly complex real and imagined forms. In areas of study involving psychological processes, concepts and theories are inevitably subject to much revision and expansion as more and more becomes known about the inner workings of the human mind.

The concept subject to expansion in this paper is one of language processing – specifically, Jurafsky's (1993, 1996) model of grammatical construction access and disambiguation. The theory that Jurafsky proposes actually consists of two distinct models – the Evidential Access Model and the Local Coherence Model, to account respectively for construction access (the retrieval of candidate structures from the mental store) and construction disambiguation (choosing the most appropriate structure from among those accessed). For ease of reference, however, these two models will be collectively referred to in this paper as the Evidential Access & Local Coherence Model (henceforth, the EALC Model). This probabilistic on-line model utilizes Construction Grammar (Kay & Fillmore, 1999; Goldberg, 1995; Fillmore, Kay, & O'Connor, 1988; Lakoff, 1987) and operates on the same basic assumption as this syntactic theory; namely that all aspects of linguistic knowledge (e.g., morphological knowledge, lexical knowledge, & syntactic knowledge) are represented uniformly by constructions of varying sizes. By allowing such constructions to be accessed in a parallel and uniform manner, as well as acknowledging the powerful influence of frequency effects on language comprehension, Jurafsky's model has the potential to account for the interpretation of a huge array of linguistic input. Because it assumes the existence of a final-state adult grammar, however, the EALC Model completely

fails to address the question of how L2 learners, who are not yet armed with complete arsenals of grammatical constructions, manage to acquire the constructions and process the linguistic input they encounter.

In order to provide the EALC Model with the explanatory power necessary to account for L2 processing and learning, this paper will seek to stretch Jurafsky's original concept by fusing it with insights and procedures from two other recent models: Abel's (2003) Dual Idiom Representation Model and Chang's (2001) model of L1 grammatical construction learning. The first half of the paper will provide a more detailed explanation of the EALC Model, as well as summaries of the models by Abel and Chang. The second half will explore areas in which the three models can potentially be merged to accommodate the aforementioned issues that were not considered in Jurafsky's EALC model.

2. The three models

2.1 Jurafsky's EALC model

By assuming lexical and syntactic forms to be uniformly represented by constructions, Jurafsky (1993, 1996) is able to apply insights from prior psycholinguistic models proposing parallel processing for word and idiom recognition to a general model of cognitive processing. Marslen-Wilson's (1987) Cohort Model of auditory word recognition, for example, states that, in retrieving individual words from the mental lexicon, a huge array of candidate lexical items are initially activated upon encountering the beginning of a word. As additional phonological and contextual information is perceived, however, candidate items are dropped from consideration until one is chosen as the best possible candidate. The point at which the single most-likely candidate is selected is referred to as the *access point*.

The idea of parallel processing for idioms was first proposed by Swinney & Cutler (1979), whose Lexical Representation Hypothesis claims each idiomatic expression is represented in the mental lexicon as one lexical item, as if it were a long word, and conceptualizes the comprehension process as a race between the literal and idiomatic meaning, with both initially being activated and beginning the race in parallel. This race, according to the Lexical Representation Hypothesis, is normally won by the idiomatic meaning since it is accessed as one chunk and, therefore, has a processing advantage over the item by item access of individual components required to compute the literal meaning.

In a series of cross-modal priming studies, Cacciari & Tabossi (1988) found that the point at which a given idiomatic expression's idiomatic meaning is activated can vary dramatically depending on the individual component words that make up the expression. In the case of conventional idiomatic expressions that begin with common delexicalized verbs such as *give*, *take*, and *make* (e.g., *get your goat*; *take someone to the cleaners*), subjects' activation of the idiomatic meanings did not occur until after the offset of the expressions. The model Cacciari & Tabossi developed based on these findings, later dubbed the Configuration Model, claims that an idiomatic expression's literal meaning is activated initially and continues to be accessed throughout processing. The idiomatic meaning is accessed

only when evidence for it is sufficient. Cacciari & Tabossi refer to this point of sufficient evidence as the *idiomatic key*, which is conceptually analogous to the Cohort Model's *access point* for word recognition. According to the Configuration Model, the *idiomatic key* can occur at any place in an idiom's word string or at its offset, causing processing time to vary from idiom to idiom.

Envisioning long-term linguistic memory as a storage facility for not only lexical and idiom constructions, but also grammatical constructions which themselves contain semantic meaning, Jurafsky (1993), stretching the concept of the mental lexicon, coins the term *constructicon* to refer to this cognitive warehouse of linguistic information. In his EALC Model, candidate grammatical constructions are retrieved from the *constructicon* when evidence for their consideration (both bottom-up and top-down phonological, semantic, and syntactic evidence) reaches an acceptable threshold. Jurafsky's criterion for weighing this evidence is purely conditional probability. In other words, the comprehender is conceptualized as subconsciously posing the question, "Given the evidence provided to me by the input thus far, is the probability of construction x being the correct one sufficient to warrant its consideration?"

Once deemed sufficiently probable, a given construction, along with others that have reached the acceptable probability threshold, are shuttled into a conceptual space that Jurafsky (1993) terms the *access buffer*. The *access buffer* serves as a sort of holding cell for constructions to rest before being integrated with the input at the *integration point*, yielding interpretations which are then subject to the EALC Model's disambiguation processes. At the disambiguation phase of the model, candidate interpretations are dropped from consideration or, in Jurafsky's terms, "pruned," as additional evidence renders them improbable candidates. In the same manner in which unlikely lexical candidates are ruled out by contextual and phonological evidence in the Cohort Model, the EALC Model's disambiguation sub-model of Local Coherence prunes all interpretations except for the one that seems most coherent in relation to the available evidence. Jurafsky (1993) refers to this point, at which one candidate interpretation emerges from the pool of eligible interpretations, as the *selection point*, which corresponds to the Cohort Model's *access point* and the Configuration Model's *idiomatic key*.

To illustrate the EALC Model with an example, let us consider the on-line comprehension of the exclamatory utterance "What a fool Harry is!" within the context of someone in the comprehender's immediate environment doing something exceedingly stupid. Upon hearing *what*, the first word of the utterance, the comprehender (assuming this person is an adult native speaker of English) retrieves, from his or her *constructicon*, the most probable candidate WHAT constructions. Based on sheer statistical probability, the most likely candidate would be the WH-NON-SUBJECT-QUESTION construction (e.g., *What is Harry doing?*; *What should we do now?*; *What drugs did Harry take?*). The single input word *what* also provides evidence not only for the WH-EXCLAMATION construction (e.g., *What a fool Harry is!*; *What nice people they are!*) but also a vast array of other WHAT constructions, including those directly related to the WH-NON-SUBJECT-QUESTION construction through inheritance links, such as the WHAT'S X DOING Y construction (e.g., *What's Harry doing swimming in the shark tank?*). All of these WHAT constructions that the comprehender judges to be sufficiently probable candidates are then integrated

with subsequent input evidence as it becomes available and are promptly pruned if they do not cohere with the expectations provided by this evidence. While an NP following *what* in the input does provide some evidence for the WH-NON-SUBJECT QUESTION construction, the determiner *a*, in the input presently under consideration, renders this an improbable choice. Given the input evidence of *what*, followed by the NP unit *a fool*, all candidate interpretations are likely pruned, except for the one provided by the WH-EXCLAMATION construction, which emerges as the selected interpretation.

By conceptualizing the comprehension process in this on-line parallel manner based on conditional probabilities, the EALC Model is able to explain the results of a large body of psycholinguistic research, not only in the areas of lexical and idiom access, but also those involving parsing preferences in ambiguous and garden-path sentences (e.g., Milne, 1982; Ford et al., 1982; Taraban & McClelland, 1988). According to the EALC Model, the reason comprehenders are led astray in interpreting garden-path sentences is simply that, in such sentences, the correct interpretations are improbable ones. Since, in the disambiguation phase of processing, unlikely candidate interpretations are quickly pruned, comprehenders reject the correct interpretation in favor of an incorrect one that is, in fact, more probable, given the available evidence. The following example, provided by Jurafsky (1993, 1996), illustrates this phenomenon well:

(1) The complex houses married and single students and their families.

Since the lexical item *complex* occurs far more frequently as an adjective than it does a noun, and since it is also much more statistically probable for *houses* to be used as a noun than a verb, the choice that stands the strongest chance of being correct is to interpret *the complex houses* as an NP, doing away with the correct, but far less probable interpretation. Basically, the EALC Model characterizes the difficulties caused by such sentences as the side effects of a probabilistic processing system that works entirely too well.

2.2 Abel's Dual Idiom Representation model

Although, in formulating the EALC Model, Jurafsky drew much insight from prior studies and models of idiom processing, all of the work in this area, at that time, dealt only with L1 processing. The manner in which non-native speakers of a language extract idiomatic interpretations from linguistic input has generally been ignored. In more recent years, however, several researchers (Liontas, 2002; Abel, 2003; Cieślicka, 2006, this volume) have sought to remedy this neglect and have taken preliminary steps to better understand the nature of L2 idiom processing. Of these, it is Abel's (2003) work that I will focus on in this paper. With insights gleaned from her own experiments with L2 learners, Abel has formulated the Model of Dual Idiom Representation (henceforth referred to here as the DIR Model), a theoretical psycholinguistic model which promises to accommodate both L1 and L2 processing of phrasal idioms.

Such idioms, examples of which are the English expressions *speak your mind* and *kick the bucket*, have undergone extensive reanalysis in recent years. Once thought to make up one homogenous class of lexical items that was, by definition, non-decomposable, or not

at all understandable based on the meanings of the individual components, these idiomatic expressions are now widely believed to reside instead at varying points along a decomposability continuum. Some expressions, such as *speak your mind*, are generally thought to be highly decomposable, while others, like *kick the bucket*, reside decidedly at the non-decomposable end of the continuum. Many other expressions, such as *spill the beans*, can be seen as lying somewhere in the middle of the decomposability continuum.[1]

Abel's (2003) investigation, a replication of a prior L1 study by Titone & Connine (1994), required participants, native speakers of German who had studied English an average of eight and a half years, to provide decomposability judgments for English idiomatic expressions. When she compared her results with those of Titone & Connine's L1 subjects, Abel found that the two groups showed opposite tendencies. While the native speakers of English tended to judge idiomatic expressions to be non-decomposable, the non-native speakers favored the decomposable rating. Upon analyzing data from a biographical questionnaire that accompanied the experiment, Abel also found that the non-native subjects who reported reading English texts daily judged significantly more idioms to be non-decomposable than their fellow non-native speakers who read English texts less often, suggesting that frequency effects have a strong influence on lexical representations.

These findings informed Abel's development of the DIR Model. Following Paivio's (1986) Dual Coding Approach to mental representations, the main claim of the DIR Model is that idiomatic expressions can be represented in the mental lexicon in two ways: (1) by an idiom entry, or (2) by conceptual representations. The latter is separate from linguistic knowledge and, therefore, functions at a cognitive level, representing general world knowledge and situations common to human experience. The DIR Model states that if an idiomatic expression is decomposable, this conceptual representation can be tapped into by activating the lexical entries of the expression's individual constituents, rendering the existence of an idiom entry for the entire expression unnecessary. If an idiom entry does exist, it simply provides additional information and facilitates processing. If an idiomatic expression is completely non-decomposable, however, activation of the expression's individual constituent lexical entries will not tap into the appropriate conceptual representation. For completely non-decomposable idioms, an idiom entry is, therefore, absolutely essential for comprehension to occur.

Abel (2003) hypothesizes that the formation of an idiom entry is largely influenced by frequency effects and, because non-native speakers have encountered far fewer L2 idioms than native speakers, L2 idiom comprehension (for expressions that are, to some degree, decomposable anyway) is reliant on the bottom-up activation of constituent entries to access conceptual representations. The finding that the non-native speakers in Abel's study who read English texts daily judged considerably more expressions to be non-decomposable than those subjects who read English texts less often supports this hypothesis.

The fact that frequency effects are granted a substantial role in determining the manner in which idioms are processed is but one of several aspects that differentiates the DIR Model from the models of idiom comprehension that came before it. Another obvious difference is that, while other models took the existence of an idiom entry for granted and were, therefore, only applicable to the L1 lexicon, the DIR Model allows for the gradual

formation of idiom entries and, hence, is applicable to both L1 and L2 idiom processing. Lastly, previous processing models ignored the conceptual aspects of idiom comprehension and instead focused exclusively on lexical representations. By integrating conceptual and lexical representations, the DIR Model acknowledges the fact that non-linguistic factors themselves, while indeed linked to lexical representations, play an important processing role. These three aspects of the DIR Model allow it to have far greater explanatory power than the earlier theories that helped inform the development of the EALC Model, such as the Lexical Representation Hypothesis and the Configuration Model.

2.3 Chang's model of construction learning

The term *model merging* in the title of this paper refers not only to the issue at hand of merging three different theoretical models, but also to one of the processes by which, according to Chang's (2001) model, grammatical constructions are learned. This model, intended to explain the manner in which children acquiring their native languages progress from the single-word stage to possession of a fully productive grammar, proposes that language learning occurs as a result of three processes: (1) reorganization of known constructions (alternately referred to by Chang as "models") through merging, on the basis of similarity, or composition, on the basis of frequently observed co-occurrence, (2) hypothesizing new constructions when known constructions prove insufficient to account for input data, and (3) reinforcement of constructions.

In this model, the constructions that are learned depend crucially on what constructions are already known. Following Tomasello's (1992) Verb Island Hypothesis, it claims that acquisition begins with lexically specific constructions. A learner, for example, might already know the lexically specific THROW-TRANSITIVE construction (e.g., *You throw the ball.*), where *throw* requires the argument roles of a thrower (*you*) and a throwee (*the ball*). If this learner encounters the utterance, *You kick the ball* in the context of family members kicking a ball around the yard, he or she will perform a mental search of all known constructions in the *constructicon* (borrowing Jurafsky's terminology) and attempt to find the construction that best accounts for the forms in the utterance and the meanings in the situation. If, from among the retrieved known constructions, the THROW-TRANSITIVE construction seems to be the best match, the learner will weigh the odds of its correctness based on how well it explains the input utterance in context. If these odds are deemed acceptable, the THROW-TRANSITIVE construction will be reorganized, through the process of merging with the newly encountered KICK-TRANSITIVE construction, to create a more general construction that includes both transitive verbs. The basis for this merging process is similarity, since *throw* and *kick* have identical arguments.

The other reorganization process proposed by Chang's model is *composition*, which occurs as a result of observing two known constructions co-occurring frequently in input data. If a learner already knows the two word HUMAN-KICK (e.g., *You kick*) and KICK-OBJECT (e.g., *kick the ball*) constructions, the frequently observed co-occurrence of these two constructions in utterances such as *She kicked the door* will eventually prompt the learner to form (or compose) the larger construction containing both the human actor and object theme. This composition process and the merging

operation both occur only if the learner determines that there is sufficient evidence for such construction reorganization. In other words, the question posed by the learner is, "Given the forms that I'm hearing, paired with the situations that I'm witnessing, what is the probability that this more general construction I'm considering will account for all similar input in the future?"

If the learner, upon evaluating the input utterance and its context in relation to the known construction, determines that the odds are not in favor of the known construction accounting for all the input, a new construction must by hypothesized. This is done by first determining which form-meaning relations in the input utterance and situation do have correspondences in the known construction. If, for example, the learner encounters the input utterance, *Don't kick the cat* in the context of a parent behaving disapprovingly, the form-meaning relations of the known KICK-TRANSITIVE construction remain relevant and are retained to be integrated into a new construction. The learner's task is then to determine which form relation of the utterance and meaning relation of the situation are not included in the KICK-TRANSITIVE construction. These unmapped relations, in this case the form *don't* and the parent's disapproving behavior, are paired and a new DON'T KICK-TRANSITIVE potential construction is created and subsequently analyzed to determine if its inclusion into the grammar (*constructicon* or stockpile of known constructions) is feasible.

Chang characterizes this analysis of a new hypothesized construction not only as an evaluation of how well it matches up with the form-meaning relations of the utterance and situation, but also in terms of overall cost to the grammar. She explains that, in the process of construction acquisition, there are two competing forces at play. One is the desire to have an economical grammar, consisting of just a small number of very general constructions. The opposing force, however, is the desire to have a very specific grammar, requiring a large number of highly specific constructions, in order to better account for the range of input data encountered. The learner is, therefore, conceptualized as continuously balancing these opposing pressures by probabilistically weighing the benefit of specificity to be gained by the addition of a given new construction against the cost of a less economical grammar. Jurafsky (1996) comments as well on these two opposing forces, noting that each is emphasized differently in the disciplines of linguistics and psychology. While linguists have tended to strive for economy in their theoretical models, the field of psychology has generally regarded the human mind as a vast storage facility capable of holding large amounts of information. By detailing a means by which a balance can be achieved between economy and specificity, Chang's model attempts to bridge the concerns of traditional linguists with Construction Grammar's requirement that the mind's storage capacity be sufficient to store a large, but finite, number of constructions.

If, in Chang's model, the learner's analysis determines that a newly hypothesized construction does not meet minimum requirements for improvement of the grammar, the new construction is not simply discarded. Instead, since it could potentially be employed at some future time, given appropriate input data, this new construction is retained in a pool of potential stand-by candidate constructions. If an input utterance includes any lexical units that are unknown to the learner, the utterance, along with its context, is, in a similar manner, retained and set aside for possible future use.

In the reinforcement process of Chang's model, frequency effects play the familiar role of strengthening known constructions. Those constructions that are encountered frequently, successfully comprehended, and successfully produced, will become quite durable and resistant to forgetting. Those that are seldom heard, seldom used, or used unsuccessfully will become weak and eventually fade from the learner's grammar.

3. When models collide

Now that the three models under consideration have each been summarized, I shall begin to explore the ways in which Abel's DIR Model and Chang's model of construction learning can each augment Jurafsky's EALC Model to help remedy its shortcomings in the areas of L2 processing and learning of grammatical constructions. I will begin by examining the role of bottom-up semantic evidence in the EALC Model and contributions that the DIR Model can provide in this area. I shall then discuss the problem of L2 learning, with Abel's model and Chang's model each supplying insights relevant to this issue, before proceeding with an overview of how a combined model would function. I will then end with a brief discussion of the role of frequency effects in L2 construction learning.

3.1 Bottom-up semantic evidence

In his discussion of the various types of evidence considered by the comprehender in the EALC Model's access of grammatical constructions, Jurafsky (1993) acknowledges that bottom-up semantic evidence can be employed in determining which constructions to access and cites idiom decomposability research by Gibbs, Nayak, & Cutting (1989), which found semantically decomposable idiomatic expressions (e.g., *pop the question, spill the beans*) to be accessed faster by native English speakers than non-decomposable expressions (e.g., *kick the bucket*). Jurafsky explains that the semantics of the word *kick*, for example, do not include any evidence for the KICK THE BUCKET construction, since there is no relation, metaphorical or otherwise, between the action of kicking and the idiom's figurative meaning of dying. The semantics of the word *spill*, however, trigger the conceptual representation of a spilling action, providing metaphorical evidence for the SPILL THE BEANS construction to be retrieved from the *constructicon*. Although this argument is logical, it assumes the existence of the SPILL THE BEANS construction in the *constructicon* and is, therefore, not applicable to the L2 learner who does not yet possess such a *constructicon* entry.

Anyone who has ever attempted to comprehend L2 input is well aware of the fact that the process, at least at the beginning to intermediate levels of proficiency, relies largely on the bottom-up analysis of an utterance's individual component words. This analysis has traditionally been regarded as a syntactic one, with a comprehender's primary concern being the task of discerning grammatical relations between constituents. A growing body of anecdotal and empirical evidence (e.g., Bernhardt, 1987; Bley-Vroman, 1991; Pienemann, 1998) suggests, however, that insufficient grammatical knowledge can be compensated

for by a reliance on non-linguistic conceptual information. According to the DIR Model, this conceptual information, which, in adult L2 learners, comprises a developed system of mental representations, schematizing general world knowledge, as well as specific situations common to human experience, can be triggered by the individual words of decomposable phrasal idioms, allowing the figurative meanings of such expressions to be at least partially understood even if the construction for a particular idiom is not present in the learner's *constructicon*.

Although Abel (2003) makes no claim that her DIR Model might be applicable to any domain of language beyond phrasal idioms, the fact that decomposing is, in principle, simply inferencing behavior that employs both conceptual and lexical information, suggests that it could be utilized by L2 learners to aid in the comprehension of a wide variety of linguistic input. Even if we restrict its applicability exclusively to idiomatic language, that alone accounts for a substantial proportion of the input a learner encounters. Indeed, the phenomenon of idiomaticity in language use is widespread. It is not just phrasal expressions of the sort Abel studied that can be viewed as residing at various points on a decomposability continuum, but also compounds (e.g., *butterfly*, *headline*) and phrasal verbs (e.g., *shut up*, *break up*), as well as sentence-length grammatical constructions, such as the WHAT'S X DOING Y construction (e.g., *What's this fly doing in my soup?*). Idiomaticity, or some degree of non-compositionality, is, in fact, central to the very notion of grammatical constructions. According to Goldberg's (1996: 68) definition, a construction is "a pairing of form with meaning/use such that some aspect of the form or some aspect of the meaning/use is not strictly predictable from the component parts or from other constructions already established to exist in the language." In the Construction Grammar framework, the idiomaticity of a given structure, or the aspects that are "not strictly predictable from the component parts" are captured in the semantics of the construction itself. Viewing constructions in this manner, the relevance of the DIR Model can easily be extended far beyond the decomposability of expressions like *speak your mind*.

In the same way that *speak your mind* and *spill the beans* can be decomposed and semantic meaning inferred from mental representations, such processes can also be employed in extracting meaning from sentence-length grammatical constructions. Many such constructions have been, in recent years, subject to extensive analysis. Kay & Fillmore (1999), for example, examine the WHAT'S X DOING Y construction at length. The LET ALONE construction (e.g., *I can't run one block, let alone a mile*) receives a similar treatment by Fillmore, Kay, & O'Connor (1988), and Goldberg (1995) explores the semantics contained in the DITRANSITIVE construction (e.g., *I faxed you the application*), the CAUSED-MOTION construction (e.g., *The crowd laughed him off the stage*), the RESULTATIVE construction (e.g., *My father slapped me silly*), and the WAY construction (e.g., *I pushed my way through the crowd*). All of these constructions can be conceptualized as lying at various points along the same decomposability continuum as phrasal idioms and compounds. Just as the DIR Model claims is the case with phrasal idioms, I am proposing here that grammatical constructions, if they are decomposable to some degree, can also be understood with the aid of mental representations activated by individual words, even if there is no *constructicon* entry for the given constructions. If such entries do exist, they

work hand in hand with the conceptual representations (which are, of course, still triggered) to facilitate comprehension.

Utilizing conceptual world knowledge as a means of deciphering sentence-level linguistic input is not a concept that is at all new or original. Bley-Vroman (1991) describes how an L2 learner might, when encountering unknown grammatical constructions, rely on conceptual representations of given situations to assign theta roles to sentence constituents. When encountering a ditransitive sentence such as *The doctor gave me penicillin*, a learner knowing only the individual lexical items and not the DITRANSITIVE construction, for example, could achieve comprehension based simply on a mental representation of the giving/transfer event – knowledge that this sort of situation requires a giver (the theta role assigned to *doctor*), a receiver (*me*), and a theme/thing given (*penicillin*). It is a process such as this one that could have been employed by the L2 participants in a study by Bernhardt (1987). These participants, all non-native learners of German, were found, in an eye-tracking study of German text, to spend a significant amount of time attending to content words, while practically ignoring grammatical function words. In contrast, subjects that were more proficient with German, both native German speakers and advanced-level non-natives, were found to fixate considerably longer on grammatical function words. The readers with lower German proficiency, if they were, in fact, constructing meaning from the mental representations activated by content words, would have had little use for these function words.

Applying such processes of the DIR Model directly to Jurafsky's EALC Model does not greatly change its basic operations. The comprehension process can still proceed in an on-line parallel fashion with the learner applying evidence as it becomes available. The key difference, however, is that, for the learner whose *constructicon* does not include an appropriate grammatical construction, the lexical constructions of component words are accessed from the *constructicon*, activating a semantic network of mental representations that, after leaving the *access buffer*, yield interpretations. These interpretations are then edited, or pruned, with each successive piece of lexical input in the same way they would be if they had resulted from grammatical constructions. The EALC Model's *selection point*, where the comprehender chooses the interpretation that is deemed most likely, given the available evidence, remains fully in effect regardless of whether it is grammatical constructions or just lexical constructions triggering mental representations that are originally accessed.

By having two distinct modes of representation – one linguistic constructions, and the other non-linguistic conceptual knowledge – some of the EALC Model's uniformity is compromised. Despite this lack of uniformity in representation type at the access phase, the combined model utilizes situational evidence available to the comprehender in the context of an utterance in much the same way that Jurafsky's original model does. In the Construction Grammar framework, semantic and pragmatic information is packaged in the constructions themselves. Since such information is encoded in the various constructions accessed for consideration, Jurafsky (1993, 1996) deals with situational evidence available to the comprehender only in the disambiguation phase, when interpretations that do not cohere with situational evidence are pruned. In this final selection phase, an L1 speaker, upon encountering the utterance *What's this fly doing in my soup?*, would, for

instance, utilize situational evidence to prune the literal interpretation and select the interpretation that resulted from the WHAT'S X DOING Y construction. For the L2 learner that is accessing a web of non-linguistic representations based on the semantics of individual component words, contextual clues can also be conceptualized as impacting processes at the disambiguation phase. For this L2 learner hearing *What's this fly doing in my soup?*, individual words would activate mental representations that include the knowledge that a fly is not ideally a soup ingredient. The surprised, or perhaps disgusted tone of the speaker's voice then serves to prune red herring interpretations, such as any that might have resulted from the word *doing*. The pruning mechanism works the same way, regardless of what system the interpretations result from. Although a dual representation system of access detracts somewhat from the uniform elegance of Jurafsky's model, when L2 learners are considered, such a model can arguably never be so neat and elegant. A degree of messiness seems unavoidable.

3.2 The problem of learning

Contextual factors are granted a substantial role in Chang's (2001) model of grammatical construction learning. Conceptualizing the construction learning process as basically a series of educated guesses concerning the likelihood of perceived forms matching all perceived evidence in a given situation, this model can perhaps best be characterized as playing the linguistic odds. In order for a learner to pair a construction with situational evidence, making such judgments possible, however, all components of the construction must first be noticed.

According to Schmidt (1990, 1994), noticing, or the conscious perception of linguistic forms, is a necessary prerequisite for input to be utilized in any sort of hypothesis formation. As Godfroid, Housen and Boers (this volume) show in each of their reviews of the literature on noticing, Schmidt's claim has been validated by quite a few SLA and cognitive psychology studies. Chang's model of construction learning allows for noticing by restricting the processes of merging, composing, and hypothesizing of new constructions to perceived input in which all lexical forms are known, and by proposing that if unknown forms are encountered in an input construction undergoing analysis, these unknown forms, along with a notation of the situation in which they are encountered, are entered into a storage pool of unexplained data, to be available for future use. Although these unknown forms are, at this point, still unexplained, the process of identifying them as unknown, associating them with a given situation, and placing them in the pool of data for future use does itself constitute noticing.

For an L2 learner at the beginning to intermediate levels of proficiency, the number of noticed constructions, both lexical and grammatical, that are available for mental representation or hypothesis formation, is quite limited. The role of the DIR Model processes of mental representation in the learning process is significant at this beginning to intermediate level when lexical constructions for some component words of larger grammatical constructions have been learned, and can, therefore, activate mental representations, while lexical constructions for other words in the grammatical constructions encountered have not yet been learned. It is through the repeated process of extracting individual words

from linguistic input and employing the resulting conceptual representations to gain at least partial comprehension that learners become more and more adept at assigning form to meaning. As this is done, the known lexical constructions utilized in triggering the conceptual representations are also strengthened, resulting in increased automaticity for these lexical constructions. This, in turn, frees up more space in working memory for previously unknown lexical items, or an entire phrase or sentence level grammatical construction, to become noticed. As their L2 conceptual representations acquire added dimensions and become closer to those of L1 speakers of the given language, the learners' pairing of form to meaning is, of course, also facilitated.

For L2 learners, however, it is not just non-linguistic conceptual representations that mediate the mapping of construction forms with meaning. As Filatova (this volume) points out, a learner's first or previously learned language cannot help but play a significant role in this process. Rather than assume, as the L1 learner does in Chang's model, that there is a lexically specific THROW-TRANSITIVE construction, an L2 learner will likely be inclined to transfer semantic categories and verb arguments from the L1 or previously learned language, assuming, or at least hypothesizing, that the verbs that share grammatical constructions in one language also do so in the other. The L2 learners, therefore, would generally begin the construction learning process with more general verb constructions than the lexically specific ones Chang proposes L1 learners start with. Instead of a lexically specific THROW-TRANSITIVE construction, L2 learners might already know the more general FORCE-TRANSITIVE construction. Because a great deal of human experience is universal, semantic categories across languages and cultures tend to be remarkably similar. The learner generalizing categories of grammatical constructions will, therefore, find such hypotheses to be correct much of the time and these already known general constructions can proceed to aid in the formation of larger more complex constructions. Since semantic categories and the arguments of individual verbs can and do vary from language to language, however, input evidence will often prove such L1 construction transfer hypotheses to be incorrect. When this occurs, the L2 learner, upon noticing the conflicting input evidence, would hypothesize a new construction to account for the input utterance in question and proceed with the probabilistic analysis phase of Chang's model.

There would, of course, be additional factors affecting the ease with which L2 learners could carry out the operations of Chang's model compared to the child L1 learners for which it was intended. Chang (2001) calls attention to the fact that child-directed speech tends to deal with concrete objects and observable actions, facilitating the task of mapping form-meaning relations. Adult-directed speech, in contrast, is far more likely to include metaphor and forms that cannot easily be mapped to situational contexts, complicating the construction learning task considerably. The processes, however, of making generalizations, based on similarity and frequency, and forming hypotheses in order to understand information that remains unexplained are both learning operations that all adult language learners consciously employ. In utilizing these processes, perhaps in a fashion similar to the one Chang's model proposes, the fact that they have a more difficult time and are nowhere near as successful as L1 learners should not come as a surprise to anyone.

3.3 Overview of the combined model

My conceptualization of a combined model would retain the basic form of the EALC model, with Jurafsky's four key conceptual spaces – the *constructicon*, the *access buffer*, the *integration point*, and the *selection point* – all continuing to operate as they did in Jurafsky's original model. In the merged model, however, these spaces are stretched to accommodate additional operations from Abel's and Chang's models at both the access and the disambiguation phases.

In the access phase of a combined model, an L2 learner, upon encountering an input utterance, would search his or her *constructicon* for known grammatical constructions that best match the input evidence (both utterance and situation). If no such constructions yet exist in the learner's *constructicon*, the learner would instead access the lexical constructions of the utterance's component words (also housed in the *constructicon*), which would activate conceptual representations. Candidate known constructions or conceptual representations (whichever the case may be) would then be shuttled to the *access buffer*, which could accommodate either.

In the disambiguation phase of the combined model, known candidate constructions and mental representations alike would both leave the *access buffer* and proceed to the *integration point*, where all evidence is taken into account, interpretations for both known candidate grammatical constructions and conceptual representations are considered, and the learner starts making the necessary educated guesses. Unlikely interpretations, be they interpretations yielded from grammatical constructions or conceptual representations, would then fall out of consideration. A known construction may be deemed relevant, but lack some crucial element noticed in the input utterance. If this is the case, the learner, at the *integration point*, would hypothesize a new construction that incorporates this element into the known construction.

In the combined model, the subsequent *selection point* is the conceptual space that must be stretched most substantially to accommodate all the learning operations of Chang's model. As in Jurafsky's original EALC model, the selection point is where final judgments are made, but this final judgment for L2 construction learners frequently involves more than making a simple selection. For the L2 learner at this selection point, the final judgments made could be merging judgments, combining two similar candidate constructions, or composing judgments, forming one large construction from two smaller ones on the basis of frequent co-occurrence. Any hypothesized construction, newly formed at the *integration point*, would also face judgment here as to whether it meets minimal criteria for acceptance into the *constructicon*. When any of these judgments, or a more straightforward final selection of one interpretation based on a known construction or conceptual representations, are made at the *selection point*, an interpretation is decided upon that the learner, while not necessarily certain, at least believes to be likely.

After final judgment is made at the *selection point*, newly merged, newly composed, and newly hypothesized constructions that are deemed acceptable for entry into the *constructicon* are shuttled there directly via a learning loop. Those constructions that do not meet the minimal criteria for acceptance go instead to yet another conceptual space that serves as a storage pool for both potential stand-by candidate constructions and any unexplained data

from the utterance or situation. The *constructicon* itself is reserved for lexical and grammatical constructions that the learner is relatively confident about, but since stand-by constructions and unexplained data could prove useful in interpreting future utterances, this storage pool maintains a link to the *constructicon* for potential future access.

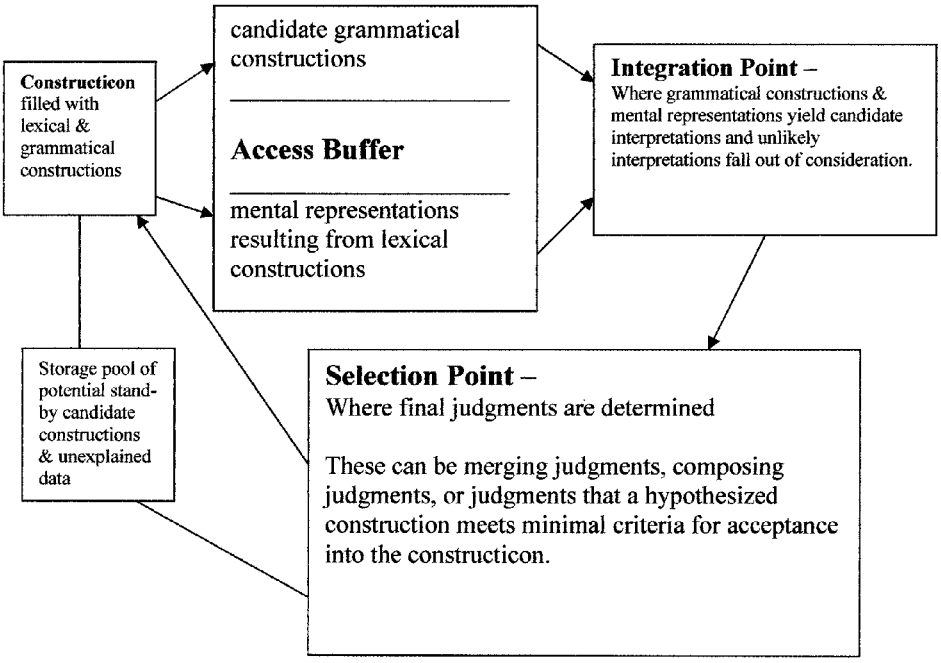


Figure 1. Schematic diagram of the combined model

4. Frequency effects

All three of the models discussed here rely immensely on the effects of frequently occurring linguistic input. In the case of the EALC Model, the comprehender is characterized as weighing the relative probability of the occurrence of constructions in determining which ones to access and prune. In the DIR Model, the frequency with which an idiomatic expression is encountered determines whether or not an idiom entry representation will eventually form and how quickly an existing one will be strengthened. Chang's model of construction learning grants frequency effects significant roles in both probabilistic hypothesizing and construction strengthening. Although large corpus databases of authentic language have enabled researchers like Jurafsky to obtain relative frequencies with which constructions occur and use this information in algorithms such as those employed in the EALC Model, this reliance on frequency information alone to predict linguistic judgments is, I contend, an over-simplification of the mechanisms involved in comprehension, especially when L2 processing and learning is considered.

As Ellis (2002) convincingly argues, the frequency with which forms are encountered does undoubtedly have an extremely powerful effect on the learning and comprehension process. It is an erroneous assumption, however, to view frequency of exposure as an absolute guarantee of eventual familiarity. Frequently occurring forms are very likely to become familiar ones, but the lexical items or grammatical structures that are familiar to learners are not always those that occur with great frequency. As Hulstijn (2002: 271) points out, "it sometimes happens that a single encounter suffices for a new word to be remembered forever." Likewise, forms that do occur with great frequency are sometimes never acquired by L2 learners. Frequency of input occurrences is clearly not the only force at play.

Giora (1999) lists frequency and familiarity as two separate factors that can influence the saliency of a form in the mind of the comprehender. The host of factors that, in turn, influence familiarity appear to vary considerably from individual to individual. The lexical or grammatical constructions that strike one learner as particularly useful or meaningful and, therefore, become permanent fixtures in that learner's *constructicon*, may not seem terribly relevant to another learner. The situation in which a form is learned could also play a role. A form learned in a novel situation may receive priority over those learned in a routine manner. When we also consider the fact that noticing must first occur before any of this is possible, individual differences in perception must also be taken into account. As Conzett (2000: 86) reminds us, "Those students often dubbed 'good at learning languages' are in fact unusually good observers of language."

This huge array of individual differences, of course, renders impossible any attempt to assign the sort of numerical values necessary for computational models. For such purposes, frequency information is arguably the best available evidence. The *constructicon* entries of forms that appear frequently in the input will, at the very least, not be weakened. Perhaps Bley-Vroman (2002: 213) makes the safest assumption in stating, "the more often something occurs in the input, the more opportunities there will be for it to be noticed." The bottom line is that frequency can only help.

5. Conclusion

Despite reservations about the possible overemphasis on frequency effects, I feel that the three models under discussion in this paper – Jurafsky's EALC Model, Abel's DIR Model, and Chang's model of construction learning – all have commendable features that can complement each other. The EALC Model, taking evidence from a large body of psycholinguistic research into account, proposes an on-line parallel processing system that operates in a uniform manner because it is retrieving and selecting from a uniform set of constructions. For the L2 learner who does not yet possess the necessary constructions, the DIR Model allows for smaller lexical constructions that are already known to trigger conceptual representations, which can then yield interpretations just as grammatical constructions do. Finally, Chang's model of construction learning allows for learning to occur concurrent with the on-line comprehension process through the processes of generalization, hypothesis formation, and the probabilistic consideration of all available evidence.

There are surely additional issues involved in conceptualizing a construction processing and acquisition model capable of accommodating L2 learners – complexities that I have no doubt neglected to discuss in this paper. Much remains to be explained regarding the precise manner in which humans extract meaningful interpretations from linguistic input. Acquisition of linguistic knowledge is arguably an even greater mystery. The characterization of the comprehension process as an ongoing series of educated guesses, however, present in all three of these models, but emphasized most in Chang's learning model, is undeniably accurate. All comprehenders, native speakers and L2 learners alike, are gamblers, perpetually weighing the odds of a candidate interpretation being an accurate one. For L2 learners, since they have considerably less knowledge of what the odds actually are, this game of chance is simply more of a gamble.

Notes

1. See Cieślicka (this volume) for a review of compositional and noncompositional models of idiom processing.

References

Abel, Beate (2003). English idioms in the first language and second language lexicon: A dual representation approach. *Second Language Research*, 19(4), 329–358.

Bernhardt, Elizabeth B. (1987). Cognitive processes in L2: An examination of reading behaviors. In J. Lantolf & A. Labarca (Eds.) *Research in second language learning: Focus on the classroom* (35–50). Norwood, NJ: Ablex.

Bley-Vroman, Robert (1991). Processing, constraints on acquisition, and the parsing of ungrammatical sentences. In L. Eubank (Ed.) *Point counterpoint: Universal grammar in the second language* (191–197). Amsterdam & Philadelphia: John Benjamins.

Bley-Vroman, Robert (2002). Frequency in production, comprehension, and acquisition. *Studies in Second Language Acquisition*, 24(2), 209–213.

Cacciari, Cristina & Patrizia Tabossi (1988). The comprehension of idioms. *Journal of Memory and Language*, 27, 668–683.

Chang, Nancy (2001). Learning grammatical constructions. Unpublished doctoral dissertation proposal, University of California at Berkeley. Retrieved April 2, 2005, from http://www.ICSI.Berkeley.EDU/~nchang/research/pubs/proposal.pdf

Cieślicka, Anna (2006). Literal salience in on-line processing of idiomatic expressions by L2 speakers. *Second Language Research*, 22(2), 115–144.

Cieślicka, Anna (this volume). Formulaic language in L2 – Storage, retrieval, and production of idioms by second language learners.

Conzett, Jane (2000). Integrating collocation into a reading & writing course. In M. Lewis (Ed.) *Teaching collocation: Further developments in the lexical approach* (70–87). Hove, England: Language Teaching Publications.

Ellis, Nick C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188.

Filatova, Ksenia L. (this volume). Third language acquisition, macro-categories and synonymy.

Fillmore, Charles J., Paul Kay, & Mary C. O'Connor (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language*, 64(3), 501–538.

Ford, Marilyn, Joan Bresnan, & Ronald M. Kaplan (1982). A competence-based theory of syntactic closure. In J. Bresnan (Ed.) *The mental representation of grammatical relations* (727–796). Cambridge, MA: MIT Press.

Gibbs, Raymond W., Nandini P. Nayak, & J. Cooper Cutting (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language*, 28, 576–593.

Giora, Rachel (1999). On the priority of salient meanings: Studies of literal and figurative language. *Journal of Pragmatics*, 31, 919–929.

Godfroid, Aline, Alex Housen & Frank Boers (this volume). A procedure for testing the Noticing Hypothesis in the context of vocabulary acquisition.

Goldberg, Adele E. (1995*). Constructions: A construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, Adele E. (1996). Construction grammar. In K. Brown & J. Miller (Eds.) *Concise encyclopedia of syntactic theories* (68–71). Oxford: Pergamon Press.

Hulstijn, Jan H. (2002). What does the impact of frequency tell us about the language acquisition device? *Studies in Second Language Acquisition*, 24(2), 269–273.

Jurafsky, Daniel (1993). A cognitive model of sentence interpretation: The construction grammar approach. Technical Report TR–93–077, International Computer Science Institute, University of California at Berkeley.

Jurafsky, Daniel (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20, 137–194.

Kay, Paul & Charles J. Fillmore (1999). Grammatical constructions and linguistic generalizations: The *What's x doing y?* construction. *Language*, 75(1), 1–33.

Lakoff, George (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.

Liontas, John I. (2002). Context and idiom understanding in second languages. *EUROSLA Yearbook*, 2(1), 155–185.

Marslen-Wilson, William D. (1987). Functional parallelism in spoken word recognition. *Cognition*, 25, 71–102.

Milne, Robert W. (1982). Predicting garden path sentences. *Cognitive Science*, 6, 349–374.

Paivio, Allan (1986). *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.

Pienemann, Manfred (1998). *Language processing and second language development*. Amsterdam & Philadelphia: John Benjamins.

Schmidt, Richard (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158.

Schmidt, Richard (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.) *Implicit and explicit learning of languages* (165–210). San Diego: Academic Press.

Swinney, David A. & Anne Cutler (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, 18, 523–534.

Taraban, Roman & James L. McClelland (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language*, 27, 597–632.

Titone, Debra A. & Cynthia M. Connine (1994). Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literality. *Metaphor and Symbolic Activity*, 9(4), 247–270.

Tomasello, Michael (1992). *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press.

CHAPTER 2

Construction learning as category learning

Nick C. Ellis
University of Michigan

This chapter presents a psycholinguistic analysis of constructions and their acquisition. It summarizes our recent research into the second language (L2) acquisition of English verb-argument constructions (VACs) and English Tense and Aspect (TA) morphology. We show that the Zipfian type/token frequency distribution of verbs in natural language optimizes construction learning by providing one very high frequency exemplar that is also prototypical in meaning: learners first acquire the most frequent, distinctive, prototypical, and generic exemplar (e.g. *go* in VL verb locative, *put* in VOL verb object locative, *give* in VOO ditransitive, etc.). VAC acquisition is further facilitated by the frequency and frequency distribution of exemplars within each island of VACs (e.g. [Subj V Obj $\text{Obl}_{\text{path/loc}}$]), by their prototypicality, and by their contingency of form-function mapping. Parallel findings in English L2 TA acquisition whereby learners first acquire the frequent, distinctive, and prototypical exemplars (e.g. telic verbs such as *steal* or *take* in perfective past, atelic verbs such as *run* or *think* in progressive) encourage the general conclusion that the acquisition of linguistic constructions can be understood in terms of the cognitive science of concept formation. It follows the general associative principles of the induction of categories from the experience of the features of their exemplars. In natural language, the type-token frequency distributions of the occupants of each of the features of the construction, their prototypicality and generality of function in these roles, and the reliability of mappings between these, together conspire to optimize learning.

Our theoretical framework is informed by cognitive linguistics, particularly constructionist perspectives (e.g., Bates & MacWhinney, 1987; Goldberg, 1995, 2003, 2006; Lakoff, 1987; Langacker, 1987; Ninio, 2006; Robinson & Ellis, 2008; Tomasello, 2003), corpus linguistics (Biber, Conrad, & Reppen, 1998; Sinclair, 1991, 2004), and psychological theories of cognitive and associative learning as they relate to the induction of psycholinguistic categories from experience (Ellis, 1998, 2002a,b, 2003, 2006a, 2006b, 2006c). The basic tenets are as follows: Language is intrinsically symbolic. It is constituted by a structured inventory of constructions as conventionalized form-meaning pairings used for communicative purposes. Usage leads to these becoming entrenched as grammatical knowledge in the speaker's mind. Constructions are of different levels of complexity and abstraction; they can comprise concrete and particular items (as in words and idioms), more abstract classes of items (as in word classes and abstract grammatical constructions), or complex combinations of concrete and abstract pieces of language (as mixed constructions). The acquisition of constructions is input-driven and depends upon the learner's experience of

these form-function relations. It develops following the same cognitive principles as the learning of other categories, schema and prototypes (Cohen & Lefebvre, 2005; Murphy, 2003). Creative linguistic competence emerges from the collaboration of the memories of all of the utterances in a learner's entire history of language use and the frequency-biased abstraction of regularities within them (Ellis, 2002a). Many cognitive linguists, corpus linguists, and psycholinguists share the realizations that we cannot separate grammar from lexis, form from function, meaning from context, nor structure from usage.

Constructions specify the morphological, syntactic and lexical form of language and the associated semantic, pragmatic, and discourse functions (Figure 1). Any utterance is comprised of a number of constructions that are nested. Thus the expression *Today he walks to town* is constituted of lexical constructions such as *today*, *he*, *walks*, etc., morphological constructions such as the verb inflection *s* signaling third person singular present tense, abstract grammatical constructions such as Subj, VP, and Prep, the intransitive motion Verb-Locative (VL: [Subj V $Obl_{path/loc}$]) verb-argument construction (VAC), etc. The function of each of these forms contributes in communicating the speaker's intention.

Psychological analyses of the learning of constructions as form-meaning pairs is informed by the literature on the associative learning of cue-outcome contingencies where the usual determinants include: factors relating to the form such as frequency and salience; factors relating to the interpretation such as significance in the comprehension of the overall utterance, prototypicality, generality, redundancy, and surprise value; factors relating to the contingency of form and function; and factors relating to learner attention, such as automaticity, transfer, overshadowing, and blocking (Ellis, 2002a, 2003, 2006b, 2008b). For example, as illustrated in Figure 1, some forms are more salient: '*today*' is a stronger psychophysical form in the input than is '*s*', thus while both provide cues to present time, *today* is much more likely to be perceived, and *s* can thus become overshadowed and blocked, making it difficult for second language learners of English to acquire (Ellis, 2006c, 2008a). These various psycholinguistic factors conspire in the acquisition and use of any linguistic construction.

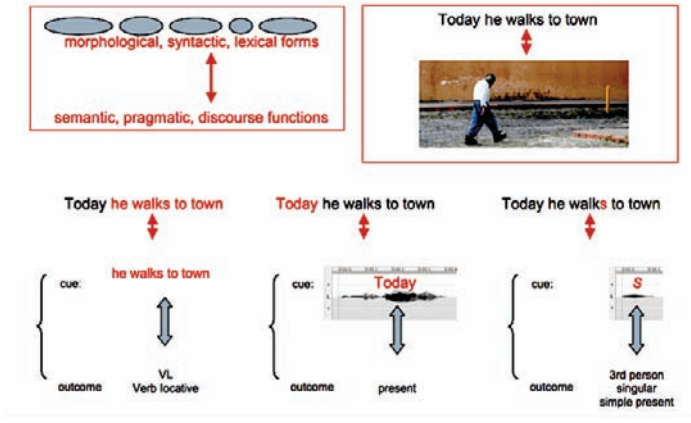


Figure 1. Constructions as form-function mappings. Any utterance comprises multiple nested constructions. Some aspects of form are more salient than others – the amount of energy in *today* far exceeds that in *s*

While some constructions, like *walk*, are quite concrete, imageable, and specific in their interpretation, others are more abstract and schematic. For example, the caused motion construction, (e.g. X causes Y to move $Z_{path/loc}$ [Subj V Obj $Obl_{path/loc}$]) exists independently of particular verbs, hence 'Tom sneezed the paper napkin across the table' is intelligible despite 'sneeze' being usually intransitive (Goldberg, 1995). How might verb-centered constructions develop these abstract properties? One suggestion is that they inherit their schematic meaning from the conspiracy of the particular types of verb that appear in their verb-island. The verb is a better predictor of sentence meaning than any other word in the sentence and plays a central role in determining the syntactic structure of a sentence (Tomasello, 1992). There is a close relationship between the types of verb that typically appear within constructions (in this case *put, move, push*, etc.), hence their meaning as a whole is inducible from the lexical items experienced within them. Ninio (1999) argues that in child language acquisition, individual "pathbreaking" semantically prototypic verbs form the seeds of verb-centered argument-structure patterns, with generalizations of the verb-centered instances emerging gradually as the verb-centered categories themselves are analyzed into more abstract argument structure constructions.

These are examples of *semantic bootstrapping* (Pinker, 1989) explanations of the acquisition of VACs whereby semantic categories are used to guide form-meaning correspondences – objects are nouns, actions are verbs, etc, and finer-grained action semantics guide particular VACs (Goldberg 1995: 39):

> Constructions which correspond to basic sentence types encode as their central senses event types that are basic to human experience... that of someone causing something, something moving, something being in a state, someone possessing something, something causing a change of state or location, something undergoing a change of state or location, and something having an effect on someone.

Learning grammatical constructions thus involves the distributional analysis of the language stream and the contingent analysis of perceptual activity following general psychological principles of category learning. Categories have graded structures, with some members being better exemplars than others. The prototype is the best example, the benchmark against which surrounding "poorer," more borderline instances are categorized. The greater the token frequency of an exemplar, the more it contributes to defining the category and the greater the likelihood it will be considered the prototype.

Frequency promotes learning, and psycholinguistics demonstrates that language learners are exquisitely sensitive to input frequencies of patterns at all levels (Ellis, 2002a). In the learning of categories from exemplars, acquisition is optimized by the introduction of an initial, low-variance sample centered upon prototypical exemplars (Elio & Anderson, 1981, 1984; Posner & Keele, 1968, 1970). This low variance sample allows learners to get a 'fix' on what will account for most of the category members. Then the bounds of the category can later be defined by experience of the full breadth of exemplars. Goldberg, Casenhiser & Sethuraman (2004) demonstrated that in samples of child language acquisition, for each VAC there is a strong tendency for one single verb to occur with very high frequency in comparison to other verbs used, a profile which closely mirrors that of the mothers' speech to these children. In natural language, Zipf's law (Zipf, 1935) describes

how the highest frequency words account for the most linguistic tokens. Goldberg et al. show that Zipf's law applied within VACs too, and they argue that this promotes acquisition: tokens of one particular verb account for the lion's share of instances of each particular argument frame, and this pathbreaking verb is also the one with the prototypical meaning from which that construction is derived

Consider language as it passes, utterance by utterance, as illustrated in Figure 2. Learners with a history of exposure to this profile of natural language might thus successfully categorize the different utterances as examples of different VAC categories on the basis of the occupants of the verb islands.

But if the verbs were the only cues that were available, then VACs could have no abstract meaning above that of the verb itself. For, 'Tom sneezed the napkin across the table' to make sense despite the intransitivity of *sneeze*, the hearer has to make use of additional information from the syntactic frame. In considering how children learn lexical semantics, Gleitman (1990) argued that they made use of clues from syntactic distributional information – nounlike things follow determiners, prepositions most often prepose a noun phrase in English, etc. The two alternatives of *semantic* and *syntactic bootstrapping* are by no means mutually exclusive, indeed, these two sources of information both reinforce and complement each other.

In the identification of the caused motion construction, (X causes Y to move $Z_{path/loc}$ [Subj V Obj $Obl_{path/loc}$]) the whole frame as an archipelago of islands is important. The Subj

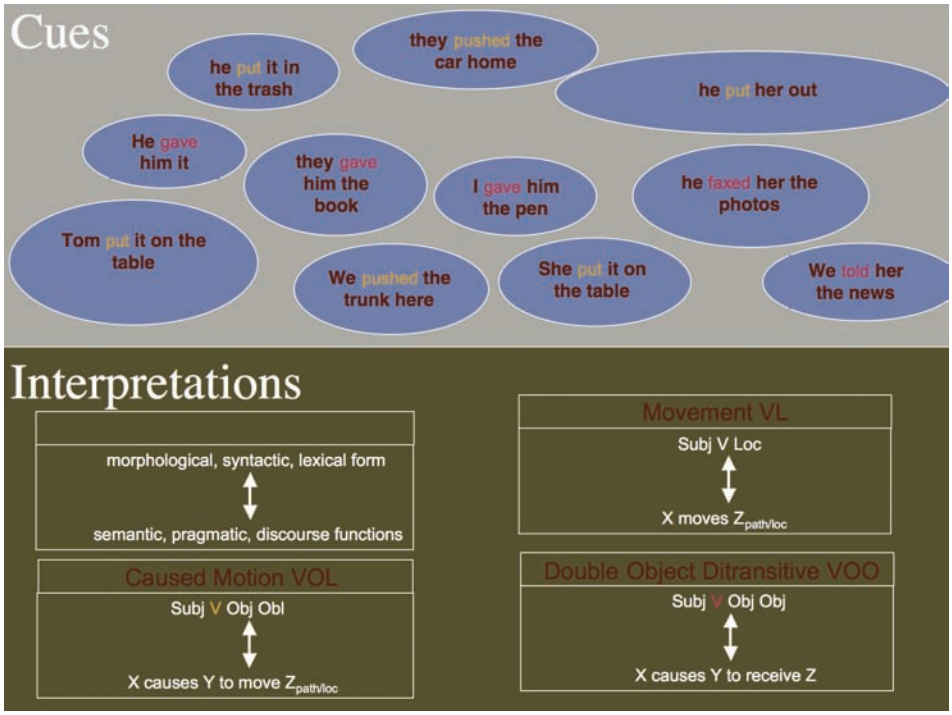


Figure 2. Verb island occupancy as cues to VAC membership

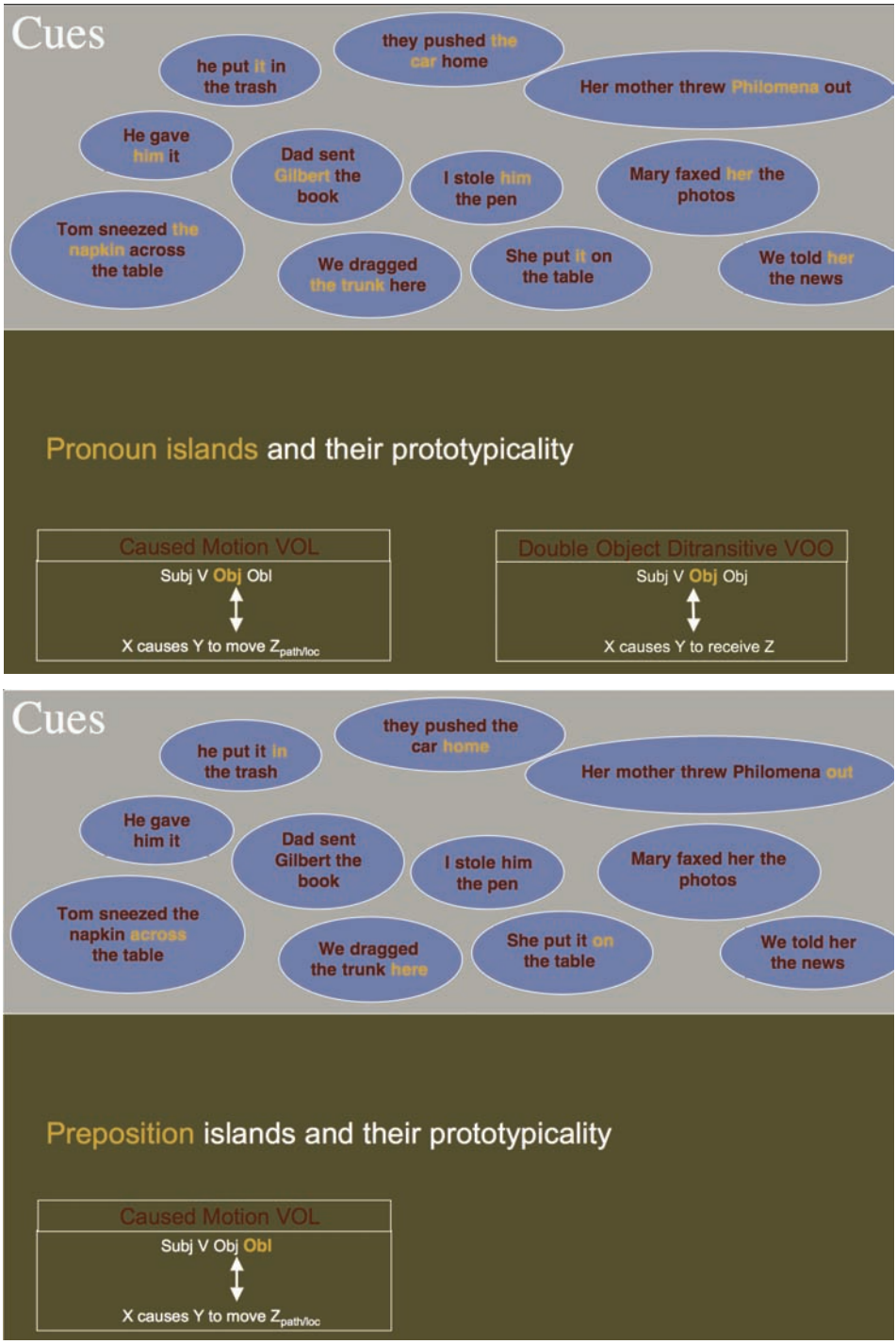


Figure 3. Other syntactic islands and their occupants as cues to VAC identity

island helps to identify the beginning bounds of the parse. More frequent, more generic, and more prototypical occupants will be more easily identified. Pronouns, particularly those that refer to animate entities, will more readily activate the schema. As illustrated in Figure 3, the Obj island too will be more readily identified when occupied by more frequent, more generic, and more prototypical lexical items (pronouns like *it* rather than nouns such as *serviette*). So too the locative will be activated more readily if opened by a prepositional island populated by a high frequency, prototypical exemplar such as *on* or *in*. Activation of the VAC schema arises from the conspiracy of all of these features, and arguments about Zipfian type/token distributions and prototypicality of membership extend to all of the islands of the construction.

The role of pronoun islands in child language acquisition has been demonstrated by Childers and Tomasello (2001) and by Wilson (2003), that of prepositional islands by Tomasello (2003: 153). Before Powerpoint, in the days when overhead transparencies provided the heights of embellishment for conference papers, Tomasello used to illustrate a putative schematic for the acquisition sequence of VACs by overlaying sequences of exemplars and considering how their cumulative experience results in entrenchment and generalization. As approximated in Figure 4, a high frequency prototype VOL seeds the VAC as a formulaic phrase. Subsequent experience of other VOLs with high frequency prototypical occupants of the different constituent islands leads to generalization of the schema, with the different slots becoming progressively more defined as attractors. The verb island must indeed play a key role in the schema, given its importance in defining the semantics of the sentence as a whole, but the other islands make important contributions too.

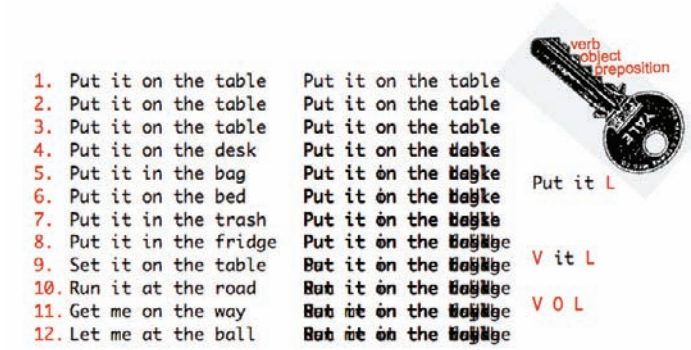


Figure 4. A schematic for the acquisition sequence of the VOL construction. Cumulative experience of VOL exemplars leads to entrenchment. A high frequency prototype VOL seeds the VAC as a formulaic phrase. Experience of other VOLs with high frequency prototypical occupants of the different islands leads to generalization of the schema, with the different slots becoming progressively defined as attractors

So frequency of usage defines construction categories. However, there is one additional qualification to be borne in mind. Some lexical types are very specific in the VACs which they occupy, the vast majority of their tokens occur in just one VAC, and so they are very reliable and distinctive cues to it. Other lexical types are more widely spread over a range

of constructions, and this promiscuity means that they are not faithful cues. *Put* occurs almost exclusively in VOL, it is defining in the acquisition of this VAC and a distinctive and reliable cue in its subsequent recognition. *Turn* however, occurs both in VL and VOL and is less distinctive in distinguishing between these two. Similarly, *send* is attracted to both the VOO and VOL constructions and so is a less discriminating cue for these categories. Think on the other islands too. It is clear that however useful they are at defining the beginning region of interest in the VAC parse, subject pronouns freely occupy any VAC with hardly any discrimination except that concerning animacy of agent. Prepositions are substantially selective for locatives, but as a class do not distinguish between the transitive and intransitive VACs. And so on.

The associative learning literature has long recognized that while frequency of form is important, so too is contingency of mapping. Consider how, in the learning of the category of birds, while eyes and wings are equally frequently experienced features in the exemplars, it is wings which are distinctive in differentiating birds from other animals. Wings are important features to learning the category of birds because they are reliably associated with class membership, eyes are neither. Raw frequency of occurrence is less important than the contingency between cue and interpretation. Distinctiveness or reliability of form-function mapping is a driving force of all associative learning, to the degree that the field of its study has been known as 'contingency learning' since Rescorla (1968) showed that for classical conditioning, if one removed the contingency between the conditioned stimulus (CS) and the unconditioned (US), preserving the temporal pairing between CS and US but adding additional trials where the US appeared on its own, then animals did not develop a conditioned response to the CS. This result was a milestone in the development of learning theory because it implied that it was contingency, not temporal pairing, that generated conditioned responding. Contingency, and its associated aspects of predictive value, information gain, and statistical association, have been at the core of learning theory ever since. It is central in psycholinguistic theories of language acquisition too (Ellis, 2006b, 2006c, 2008b; Gries & Wulff, 2005; MacWhinney, 1987; Wulff, Ellis, Römer, Bardovi-Harlig, & LeBlanc, 2009).

Our research therefore pursued a psycholinguistic analysis of constructions and their second language acquisition, focusing upon how acquisition is affected by the frequency and frequency distribution in natural language usage of exemplars within each island of the construction, by their prototypicality, and by their contingency of form-function mapping. In three studies we have focussed upon the contribution of the verb in VACs (Ellis & Ferreira-Junior, 2009a), upon the other islands in these constructions (Ellis & Ferreira-Junior, 2009b), and upon the frequency and prototypicality of lexical aspect in TA acquisition (Wulff, Ellis, Römer, Bardovi-Harlig, & LeBlanc, 2009). We will gather and summarize the major findings here. The reader is referred to these original publications for methodological detail, results, and analyses.

Study 1. The acquisition of VACs:

The role of the Verb

Goldberg, Casenhiser & Sethuraman (2004) demonstrated that in samples of child language acquisition, for a variety of constructions, there is a strong tendency for one single verb to occur with very high frequency in comparison to other verbs used:

- The Verb Object Locative (VOL) [Subj V Obj $Obl_{path/loc}$] construction was exemplified in children's speech by *put* 31% of the time, *get* 16%, *take* 10%, and *do/pick* 6%, a profile mirroring that of the mothers' speech to these children (with *put* appearing 38% of the time in this construction that was otherwise exemplified by 43 different verbs).
- The Verb Locative (VL) [Subj V $Obl_{path/loc}$] construction was used in children's speech with *go* 51% of the time, matching the mothers' 39%.
- The ditransitive (VOO) [Subj V Obj Obj_2] was filled by *give* between 53% and 29% of the time in five different children, with mothers' speech filling the verb slot in this frame by *give* 20% of the time.

Thus Goldberg (2006) argued that the constructions of natural language, like natural categories, are optimized for learning by providing one very high frequency exemplar that is also prototypical in meaning.

Since the same communicative and functional concerns motivate both first and second language (Robinson & Ellis, 2008), we expect a similar pattern for L2 acquisition. Ellis & Ferreira-Junior (2009a) therefore tested this proposal for naturalistic second language learners of English VACs in the European Science Foundation (ESF) corpus (Dietrich, Klein, & Noyau, 1995; Feldweg, 1991; Perdue, 1993). The ESF study collected the spontaneous second language of adult immigrants in France, Germany, Great Britain, The Netherlands and Sweden longitudinally with learners being recorded in interviews every 4 to 6 weeks for approximately 30 months. The corpus is available from the Max Planck Institute for Psycholinguistics (http://www.mpi.nl/world/tg/lapp/esf/esf.html). Our analysis is based on the data for seven ESL learners living in Britain whose native languages were Italian (n = 4) or Punjabi (n = 3). Data from 234 sessions were gathered and transcribed for these ESL learners and their native-speaker (NS) conversation partners from a range of activities. The NS language data is taken to be illustrative of the sorts of naturalistic input to which the learners were typically exposed, although we acknowledge some limitations in these extrapolations.

We performed semi-automated searches through the transcriptions to identify the VACs of interest and to tag them as VL, VOL or VOO following the operationalizations described in Goldberg, Casenhiser & Sethuraman (2004), e.g.

a. SLA: you come out of my house. [come] [VL]
b. SMA: charlie say # shopkeeper give me one cigar ## he give it ## he er # he smoking # [give] [VOO]
c. SRA: no put it in front # thats it # yeah [put] [VOL]

For the NS conversation partners, we identified 14,574 verb tokens (232 types) of which 900 tokens were identified to occur in VL (33 types), 303 in VOL (33 types), and 139 in

VOO constructions (12 types). For the NNS ESL learners, we identified 10,448 verb tokens (234 types) of which 436 tokens were found in VL (39 types), 224 in VOL (24 types), and 36 in VOO constructions (9 types).

Our specific hypotheses and the findings relating to them were as follows:

H1. *The frequency distribution for the types occupying the verb island of each VAC are Zipfian.*

The frequency distributions of the verb types in the VL, VOL and VOO constructions produced by the NS interviewers and the NNS learners are shown in Figure 5. For the NS interviewers *go* constituted 29% of the total tokens of VL, *put* constituted 32% of VOL use, and *give* constituted 47% of VOO. After this leading exemplar, subsequent verb types decline rapidly in frequency. For the NNS learners, again, for each construction there was one exemplar that accounted for the lion's share of total productions of that construction: *go* constituted 53% of VL, *put* 68% of VOL, and *give* 64% of VOO. Plots of these frequency distributions as log verb frequency against log verb rank produced straight line functions explaining in excess of 95% the variance thus confirming that Zipf's law is a good description of the frequency distributions with the frequency of any verb being inversely proportional to its rank in the frequency table for that construction, the relationship following a power function.

H2. *The first-learned verbs in each VAC are those which appear more frequently in that construction in the input.*

The rank order of emergence of verb types in the learner constructions followed the frequencies in the interviewer NS data. Correlational analyses across all 80 verb types which featured in any of the NS and/or NNS constructions confirmed this to be so. For the VL construction, frequency of lemma use by learner correlated with the frequency of lemma use by NS interviewer $r(78) = 0.97$, $p < .001$. The same analysis for VOL resulted in $r(78) = 0.89$, $p < .001$, and for VOO resulted in $r(78) = 0.93$, $p < .001$.

H3. *The pathbreaking verb for each VAC is much more frequent than the other members.*

Go was the first-learned verb for VL, *put* for VOL, and *give* for VOO. The Zipfian frequency profiles (Figure 5) for the types/tokens confirm H3. The acquisition functions (Ellis & Ferreira-Junior, 2009a, Figures 5–7) showed in each case that the first-learned verb seeded the construction and predominated in its cumulative usage, but thereafter the construction grew in membership as verbs similar in meaning to the pathbreaker joined one at a time.

Language use exhibits recency effects in the 'dance of dialogue' whereby the constructions used by one speaker affect the use and availability of the same constructions in their conversation partner (Pickering, 2006; Pickering & Garrod, 2006). This phenomenon, known as priming, can be observed across phonology, conceptual representations, lexical choice, and syntax. We thus asked to what extent the matching profiles of constructional verb use in NS and NNS result from priming. We used the CHIP routine in CLAN (MacWhinney, 2000) to look for this phenomenon in the highest frequency verbs in each

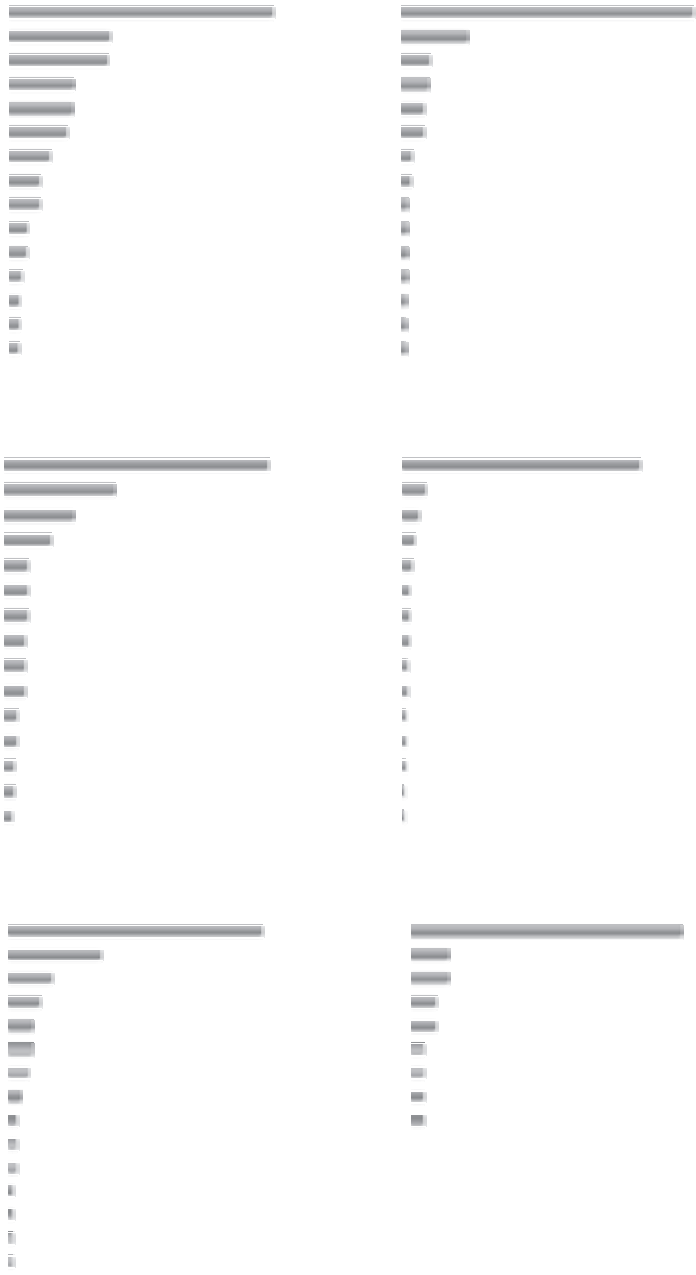


Figure 5. Zipfian type-token frequency distributions of the verbs populating the Interviewers' and Learners' VL, VOL, and VOO constructions. Note the similar rankings of verbs across Interviewers and Learners in each VAC

VAC: to what extent do the NNS uses of *go*, *put*, and *give* follow immediately from NS uses. Of 233 NNS uses of *go* in the VL construction, we could find 17 that seemed to result from priming from the NS interview interaction. Of 152 NNS uses of *put* in VOL, 8 appeared to be directly primed in this way. Of the 22 NNS uses of *give* in VOO, one appeared to be primed in this way. Interactional priming thus appears to account for some of the NNS use observed here, but by no means the majority.

H4. *The first-learned verbs in each VAC are prototypical of that construction's functional interpretation.*

In order to determine the degree to which different verbs matched the prototypical semantics of the three VACs, we had native English speakers rate the verbs on a 9 point scale for the degree to which they matched a VL schema (the movement of someone or something to a new place or in a new direction), a VOL schema (someone causes the movement of something to a new place or in a new direction), or a VOO schema (someone causes someone to receive something). We then assessed the association between verb-acquisition order and prototypicality so measured.

For the VL construction the most used verb, *go*, was rated as 7.4 out of 9 in terms of the degree to which it matched the prototypical schematic meaning. The correlation between prototypicality of verb meaning and log frequency of learner use was VL $rho(78) = 0.44$, $p < .001$. We had expected a higher correlation than this but realized that ten other verbs surpassed *go* in this rating: *walk* (9.0), *move* (8.8), *run* (8.8), *travel* (8.8), *come* (8.4), *drive* (8.2), *arrive* (8.0), *jump* (8.0), *return* (8.0), and *fall* (7.8). These match the schema very well, but their additional specific action semantics limit the generality of their use. What is special about *go* is that it is prototypical and generic – thus widely applicable. The same pattern held for the other constructions. For VOL, the most used verb *put* was rated 8.0 in terms of how well it described the construction schema. For the VOO construction, the most used verb *give* was rated 9.0 in terms of how well it described the VOO schema.

With regards to hypotheses 1–4, in sum, this study demonstrated that learner VAC acquisition is seeded by the highest frequency, prototypical, and generic exemplar across learners and VACs.

There is good evidence that these factors first play out in learning to comprehend the L2. The analyses of NNS here are done irrespective of total accuracy of form in production. While learner productions of the simpler VL construction are usually correct, the structurally more complex VOL and VOO constructions are often produced in a simplified form, i.e., the Basic Variety so clearly identified and analyzed in the original ESF project (Klein & Purdue, 1992; Perdue, 1993). This typically involves a pragmatic topic-comment word ordering, where old information goes first and new information follows. Examples for the VOL include:

yeah this television put it up the # book #
this bag <he put him> [/?] put in the st [/?] er floor # <bag> [>1]
a horse # put in there < > [$ laughs]
you know which block put down
yeah keep it money ## put the table [/?] # put in the table

Comprehending which verbs go with which arguments in which VACs is the start of the process. Learning to produce these arguments in their correct order is a slower process, one which in these data seems to start with highly generic formulaic phrases such as "put it there".

Our analysis of the degree to which NNS verb use resulted from priming from prior NS use resulted from comments upon an earlier draft that this could simply reflect priming. A perspective which attributes such use to either the cognitive phenomenon of priming, or the discourse phenomenon of shared topic, seems too black-and-white. Second language acquisition researchers who take an interactionist view of learning (Ellis, 2008c; Gass, 1997, 2002, 2003; Gass & Mackey, 2007; Gass, Mackey, & Pica, 1998; Gass & Varonis, 1994; Long, 1980; Mackey, 1999; Mackey & Gass, 2006) look to these interactions where scaffolding, negotiation, and priming of form occur while conversation partners focus upon shared meanings, as the very process of acquisition.

One other suggested explanation of these findings is that they are statistical artefacts because high frequency items will be sampled earlier than low frequency ones, and thus the fact that high frequency items are found in these samples for NS and NNS alike is uninformative with regards to NNS's knowledge of these forms (Tomasello & Stahl, 2004). Again, there is no denying these sampling phenomena as a potential force behind the data observed by the ESF researchers and analysed here. Yet at the same time, we are reluctant to relegate these observations as artifactual. This is because we believe these sampling phenomena are equally potent in the data observed and analysed by non-native speakers where their experience too is more likely to sample high frequency items which are tend to prototypical and generic uses.

Study 2. The acquisition of VACs:

The role of the other islands in the VAC archipelago

Ellis & Ferreira-Junior (2009b) extended these analyses in two ways, firstly to include the dimension of contingency / distinctiveness of form-meaning association, and secondly to investigate the contribution of the other islands in the VAC archipelago beyond the verb. Following the arguments introduced with Figure 4, we assume similar contributions from the other islands in each VAC, though perhaps to a lesser degree.

The particular hypotheses, and the results that pertained, were as follows:

H5. *The first-learned verbs in each TA construction are those which are more distinctively associated with that construction in the input.*

The determination of the reliability of association of a form-function pair involves contingency analysis of the type illustrated in Table 1.

Table 1. A contingency table showing the four possible combinations of events showing the presence or absence of a target Cue and an Outcome

	Outcome	**No outcome**
Cue	*a*	*b*
No cue	*c*	*d*

a, b, c, d represent frequencies, so, for example, *a* is the frequency of conjunctions of the cue and the outcome, and *c* is the number of times the outcome occurred without the cue.

A good cue is one where, whenever it is present the outcome pertains, and whenever absent the outcome does not, i.e. where observations load on the diagonal in cells *a* and *d* rather than being randomly distributed about the table. We used three different measures of contingency. The first, from collostructional analysis (Gries & Stefanowitsch, 2004; Stefanowitsch & Gries, 2003), was quantified by means of the log to the base of 10 of the *p*-value of the Fisher Yates exact test to assess the association strength between the verbs and the VACs they occur in, such that highly positive and highly negative values indicate a large degree of attraction and repulsion respectively, while 0 indicates random co-occurrence. The Fisher-Yates exact test is a measure of the two-way dependency between a pair of events. But associations are not necessarily reciprocal in strength. Recall how 'bird' cues 'eyes', but eyes are not distinctive cues for the category 'bird'. We therefore separately assessed these directional relations using the one-way dependency statistic ΔP which is a good predictor of cue learnability in the human associative learning literature (Allan, 1980; Ellis, 2006b; Shanks, 1995). ΔP approaches 1.0 as the presence of the cue increases the likelihood of the outcome and approaches –1.0 as the cue decreases the chance of the outcome – a negative association. ΔP (Word→Construction) determines the degree to which lexical types in islands are predictive of particular VACs, ΔP (Construction→Word) the degree to which particular VACs are predictive of particular lexical types in their various islands.

As already described under H2, learner uptake was strongly associated with frequency in the NS speech (over the 80 verbs, VL: $r = 0.97$; VOL $r = 0.89$; VOO $r = 0.93$). Our analyses under H5 showed that, if anything, learner uptake was predicted even more so by collexeme strength (Fisher-Yates) in the NS speech (over the 80 verbs, VL: $r = 0.96$; VOL $r = 0.97$; VOO $r = 0.97$), by contingency (ΔP Construction→Word) in the NS speech (over all 80 verbs, VL: $r = 0.95$; VOL $r = 0.89$; VOO $r = 0.93$) and, to a lesser degree, by contingency (ΔP Word→Construction) in the NS speech (over the 80 verbs, VL: $r = 0.26$; VOL $r = 18$; VOO $r = 0.75$).

These different measures of association are themselves highly correlated, and with such multicollinearity it is difficult to separate the predictor variables. However, it is clearly the case that NS collexeme strength (Fisher-Yates) is a very strong predictor of NNS acquisition, as is ΔP (Construction→Word). What is less predictive is ΔP (Word→Construction). When a construction cues a particular word, that word occurs very often in that construction and tends to be very generic. When a word cues a particular construction, it may be a lower frequency word, quite specific in its action semantics and thus very selective of that construction (e.g. *fell, turn*, and *stay* for VL, *hang,* and *drop* for VOL).

H6. *The frequency distribution for the types occupying each of the islands of each VAC is Zipfian.*

We determined the frequency distributions of the types occupying each (non-verb) island in the VL (Subj, Prep, Locative), VOL (Subj, Obj, Prep, Locative), and VOO (Subj, Obj_1, Obj_2) constructions produced by the NS interviewers and the NNS learners. For each construction, the frequency distribution for each island was Zipfian. In each case, for NS and NNS both, the lead exemplars took the lion's share of instances in that island, and the distribution was a power function as indexed by the regression of log frequency vs. log type rank being linear and explaining a very substantial part of the variance.

H7. *The first-learned types in each VAC island are those which appear more frequently in that construction island in the input.*

There was a clear correspondence between the types used in each island by the NNSs and the types that occupy them in the speech of the NS interviewers.

The NS interviewers filled the Subj island of VL with the following top 8 types, in decreasing order: *you, to* [verb in infinitive phrase], implied *you* [imperative], *I, he, they, we, us.* The corresponding list for the NNS learners was: implied *you* [imperative], *I, you, he, they, to* [verb in infinitive phrase], *she, we*. A similar profile was found for the Subj island for VOL: NS (*you*, implied *you* [imperative], *to, I, they, he, we, she*), NNS (implied *you* [imperative], *I, you, to* [verb in infinitive phrase], *he, the, bag, they*), and for VOO: top 4 NS (*I, you,* implied *you* [imperative], *to* [verb in infinitive phrase]), NNS (*they, I, she,* implied *you* [imperative]). Although a potentially infinite range of nouns could occupy the Subj islands in these different constructions, in NS and learner alike, they were populated by far by a few high frequency generic forms, the pronouns.

The top 8 occupants of the Prep island in VL were for the NS speakers (*to, in, at, there, from, into, out, back*), and for the NNS learners (*to, in, out, on, down, there, inside, up*). Similar profiles occurred for the Prep island of VOL: NS (*in, on, there, off, out, up, from, to*), NNS (*in, on, there, the table, up, from, the_bag, down*). Although a wide range of directions or places could occupy the post-verbal island in these two constructions, in NS and learner alike, it was occupied by far by a few high frequency generic prepositions.

Finally, the Obj islands of VOO. For Obj_1, the NS interviewers' top 5 occupants were (*you, me, him, her, it*), the NNS learners' the top 3 were (*me, you, him*). For Obj_2, the NS top 8 were (*AMOUNTMONEY* [like twenty pounds, three pounds, etc.], *the_names, a_bit, money, a_book, a_picture, something, the_test*), the NNS top 8 were (*money, a_letter, hand, something, the_money, a_bill, a_cheque, a_lot*).

The general pattern then, for each island of each VAC, is that there was high correspondence between the top types used in each island by the NNS learners and the types that occupy them in NS input typical of their experience.

H8. *The first-learned pathbreaking type for each VAC island are much more frequent than the other members.*

The qualitative patterns summarized under H7 demonstrates that, unlike for the verbs which centre the semantics of each VAC, there was no single pathbreaker that initially takes

over each of the other islands of the VAC exclusively. Nevertheless, for each construction, the frequency distributions for each island was Zipfian, and there was a high overlap between NS and NNS use of the top 5–10 occupant types which together make up the predominance of its inhabitation.

H9. *The first-learned types in each VAC island are prototypical of that island's contribution to the construction's functional interpretation.*

The 5–10 major occupant types for each island described under H7 do indeed seem to be prototypical in role. We did not have native speaker ratings for the prototypicality of meaning of the other island inhabitants as we had for the verbs, however, the qualitative data was highly consistent with this hypothesis.

Although a potentially infinite range of nouns could occupy the Subj islands in the VL, VOL and VOO constructions, in NS and NNS learner alike, these were occupied by the most frequent, prototypical and generic forms for this slot – pronouns such as *I, you, it, we*, etc. The Prep islands in VL and VOL were clearly identified with high frequency prototypical generic prepositions such as *in, on, there, to*, and *off*. Likewise the Objs in VOO are very stereotypic in their functional interpretations, and there is broad overlap between NS and NNS use: people (as pronouns) routinely give people (as pronouns) money, letters, bills, or books.

Indeed if we put all of these data together, and simply choose the two lead exemplars, the most popular / populating types in each island in each VAC and move left to right as in a finite state grammar for the NS and NNSs in turn, we compose the following utterances: such prototypical VL sequences as "come in", "I went to the shop", and "to go to [Country]", such prototypical VOL sequences as "you put it in it", "take them in there", "put it in the bag", and "I put it on the table", and such prototypical VOO sequences as "I gave you AmountMoney", "you tell me the names", "they wrote me a letter", and "I'll give you money".

H10. *The first-learned types in each VAC island will be those which are more distinctively associated with that construction island in the input.*

Our analyses showed that certain subjects were more significantly associated with certain VACs, for example *it* and *I* for VOO, and implied_*you* in the imperative for VOL. Nevertheless, comparison of the data under H5 showed that verbs are generally much more distinctively associated with these VACs than Subjs both in terms of Collocation Strength, and ΔP measures. Thus while the occupants of Subj do follow a Zipfian distribution lead by pronouns, and thus could indeed signal the beginning of a VAC parse, they tend not to be associated with any particular VAC. Prepositions were much more like the verbs in their selectivity: *to, back, in* and *out* were distinctively associated with VL, *on, off*, and *up* were strongly selective of VOL; and all of these prepositions were repulsed by VOO. For the Obj1 islands, any Obj1 repulsed VL, *it, money, them* and *that* were very significantly distinctive of VOL, and the object pronouns *you, me, him* and *her* were distinctive recipients in VOO.

Together, these analyses demonstrated that, while the verb island is most distinctive, the constituency of the other islands is by no means negligible in determining VAC identity. In particular, VL and VOL are highly selective in terms of their Prep occupancy, and Obj1 types clearly select between VOO, VOL and VL.

Study 2 therefore illustrates how each island in each VAC archipelago is thus a significant feature which makes a contribution to its identification and interpretation.

Study 3. The L2 acquisition of Tense & Aspect

Wulff, Ellis, Römer, Bardovi-Harlig, & LeBlanc (2009) explored the acquisition of tense-aspect morphology from this same constructionist perspective. Child language learners are initially influenced by the inherent semantic aspect of verbs in the acquisition of TA morphology affixed to these verbs. They start out by using the perfective past morpheme with telic verbs (achievements and accomplishments, with a clear endpoint) before they extend its use to dynamic atelics (activity and stative verbs). Conversely, progressive marking is first used with activity verbs before it spreads to telic verbs. The Aspect Hypothesis (Andersen & Shirai, 1994) thus describes how the abstract grammatical schema for perfective past generalizes from more concrete beginnings close to the prototypic centre in the clear exemplifications of telic achievements and accomplishments. Likewise abstract progressive morphology emerges from concrete exemplars in the semantics of activities and states.

Aspect-before-tense phenomena also prevail in second language acquisition (Andersen & Shirai, 1994; Bardovi-Harlig, 2000; Li & Shirai, 2000). Adult language learners too are sensitive to the lexical aspects of verbs, initially using combinations of lexical and grammatical aspect that are maximally compatible, with telicity being a particularly salient feature. Thus L2 learners from a wide variety of L1/L2 combinations first use perfective past marking on achievements and accomplishments, and only later extend this to activities and state. Similarly, in L2s that have progressive aspect, progressive marking begins with activities and only extend slowly thereafter to accomplishments and achievements.

The influence of input frequency on TA acquisition has also been investigated. Andersen (1990, the Distributional Bias Hypothesis) observed that the input available to learners exhibits distributional patterns similar to those observed in learners' productions: "Native speakers in interaction with other native speakers tend to use each verb morpheme with a specific class of verbs, also following the aspect hypothesis" (Andersen & Shirai, 1994: 137). Such input frequency biases should aid the statistical learning of TA constructions. In our study of L2 TA morphology we therefore expected likewise that the first-learned verbs in each TA construction would be those which appear more frequently in that construction in the input.

Our research was therefore designed to again test hypotheses relating to form frequency, form-function distinctiveness, and functional prototypicality in the acquisition of L2 TA constructions as cognitive categories. Our particular hypotheses, and the findings relating to them, were as follows:

H11. *Natural language data has a distributional bias whereby some verb types occupy each TA construction much more frequently than others, the distribution of the types constituting each construction being Zipfian.*

In order to examine frequency biases in the input, we retrieved verb form frequencies for all verbs from two native speaker corpora taken to represent the type of language input adult second language learners are exposed to: the 10 million word spoken section of the British National Corpus (BNC_{spoken}) and the 1.7 million word Michigan Corpus of Academic Spoken English (MICASE, Simpson, Briggs, Ovens, & Swales, 2002). All verb form frequencies were retrieved from CLAWS-tagged versions of BNC_{spoken} and MICASE, respectively. When we analysed the verbs tagged as simple past or progressive, it was clear that their frequency distributions across the different TA categories was Zipfian: the frequency with which verbs occur with a certain tense-aspect category is inversely proportional to their rank in the frequency table, with the most frequent verb types accounting for the lion share of all occurrences of any given TA morpheme. Unlike for the VAC data, however, the top ten most frequent verbs within each category were not typically distinctive of that category, because the very highest frequency verbs in the language (like *do*, *be*, *have*, and *get*) naturally occupy the top ranks across all TA categories.

H12. *The more frequent verbs in each TA construction are distinctively associated with that construction in the input.*

In order to determine which verbs are particularly associated with the progressive and the perfective more systematically we computed a multiple distinctive collexeme analysis (MDCA) for the BNC_{spoken} and MICASE data sets (Gries & Stefanowitsch, 2004). The association-based distributions showed that a small number of verbs are extremely highly associated with a particular TA category, and association strength drops exponentially thereafter. Ranking the top ten most distinctively associated verbs for each TA reflected intuitions about verbs that typically occur with the different TA categories: the past and perfect TA columns were occupied by highly telic verbs such as *die*, *crash*, *explode*, *lose*, or *finish*; the progressive preferred continuous action verbs like *sit*, *play*, *walk*, and *run*. These distinctively-associated verbs, while not *the* highest frequency in the language (H11), are frequently experienced in that construction.

H13. *The verbs most distinctively associated with each TA construction in the input are prototypical of the meaning of that construction.*

In order to investigate the prototypicality of the verbs, we obtained native speaker telicity ratings for a range of verbs selected from these analyses from 20 native speakers of American English. A questionnaire presented the verbs in isolation, without arguments, and in their base forms. Subjects were instructed to evaluate each verb with regard to how strongly it implies an endpoint expressed in values from 1 (if there is no endpoint implied) to 7 (if an endpoint is strongly implied). Three examples were given: *smash* as a highly telic verb, *continue* as an example of a verb that is located at the opposite, atelic end of the continuum, and *swim* as an example of a verb that falls somewhere in between.

The resulting Telicity Rating data demonstrated that those verbs distinctively associated with past tense in the input received significantly higher telicity ratings than verbs

associated with the progressive (MICASE data: $t = -2.107$; df = 18; $p = .049$; BNC_{spoken} data: $t = -4.356$; df = 18; $p < .001$).

H14. *The first-learned verbs in each TA construction are prototypical of that construction's functional interpretation in terms of their telicity / lexical aspect.*

We revisited oral production data collected by Bardovi-Harlig (2000) who had 37 English beginning L2 learners from 5 different L1 backgrounds watch an excerpt of *Modern Time* and then tell the story in their own words. The resulting narratives produced an average of 51 verb tokens. All verb forms were coded for TA morphology (that is, simple past, past progressive, pluperfect, present, present progressive, progressive without auxiliaries, present perfect, or "uninterpretable"). For the purpose of the present study, we selected from this data set verbs that occurred more than 10 times overall and which were distinctly associated with present, simple past, or progressive as determined by a chi-square test.

The 5 most frequently occurring past tense verbs in the learner production data (*say*, *see*, *steal*, *take*, *tell*) and the 5 most frequently occurring progressive verbs (*begin*, *eat*, *run*, *think*, *walk*) differed significantly in their mean telicity ratings ($t = -2.838$; df = 9; $p < .01$), with the past tense verbs being judged more telic and the progressive verbs more atelic.

In sum, these analyses of Wulff, Ellis, Römer, Bardovi-Harlig, & LeBlanc (2009) suggest that the verbs first learned by adults in the progressive are also frequent in the progressive in the input, distinctively associated with the progressive in the input, and highly atelic (i.e., significantly less telic than verbs frequent and associated with past tense in the input). Likewise, the verbs first learned in past tense are frequent in past tense in the input, highly distinctive for past tense in the input, and highly telic. We conclude, in terms of the general cognitive properties of schema learning: (1) The first-learned verbs in each TA construction are those which appear frequently in that construction in the input. (2) The first-learned pathbreaking verbs for each TA construction are distinctive of that construction – the contingency of forms and function is reliable. (3) The first-learned verbs in each TA construction are those which are prototypical of the construction's functional interpretation in terms of telicity / lexical aspect.

Conclusions

These three studies demonstrate the same range of influences in the acquisition of the linguistic constructions. For each island of a VAC there is:

1. the frequency, the frequency distribution, and the salience of the form types,
2. the frequency, the frequency distribution, the prototypicality and generality of the semantic types, their importance in interpreting the overall construction,
3. the reliabilities of the mapping between 1 and 2.

All of these factors affect VAC acquisition. They affect the acquisition of TA constructions likewise. Learning is driven by the frequency and frequency distribution of exemplars within construction and by the match of their meaning to the construction prototype.

Construction type/token frequency distribution in natural language might thus optimize learning by providing one very high frequency exemplar that is also prototypical in meaning and widely applicable.

The three factors of frequency, distinctiveness and prototypicality interact, and they are usually positively associated. Thus distinguishing any one factor as the root cause of category acquisition is problematic, and probably naïve. The perspective adopted here suggests instead that, as in the acquisition of other categories, it is the conspiracy of these several different factors working together that drives acquisition of linguistic constructions.

Perhaps indeed it is natural that they conspire in these ways, and that the functions of language in human communication have resulted in the evolution through usage of a system that optimally maps human cognition onto language form. In so doing, it results in a system that is readily acquired. Investigation of the ways in which language usage and language cognition result in learnable language structures is a much larger enterprise, and it is only at the beginnings (Christiansen & Chater, 2008; Ellis, 2008b; Ellis & Larsen-Freeman, 2006, 2009). We can only begin to outline here how frequency, distinctiveness and prototypicality have become associated in language and thus in learning.

Before learners can use constructions productively, they have to analyze them, to identify their linguistic form and to map it to meaning. Each construction has its form, its meaning, and its contingency of mapping between them. The current research shows that the input that learners get is biased so that they frequently experience forms that are distinctively associated with prototypical functions or construals. Language lines up with the world, or, better, with the way we construe it. Our understanding of the world lines up with our language. Our actions in the world, our categorization of the world, and our talk about these actions and classifications occur in broadly parallel relative frequencies. These parallels make constructions learnable. Interference with any of these aspects reduces learnability: constructions of low salience of form are hard to learn, constructions where there is low reliability/contingency between form and meaning are hard to learn, constructions with subtle construals yet to be discerned are hard to learn, constructions of low frequency of occurrence tend to be acquired later (Ellis, 2006c). As 'causes' or forces in language learning, it would be difficult therefore to put any of these factors first. Nevertheless, these dynamic interactions warrant serious exploration, and we believe that computer simulation is one fruitful avenue for investigation of the contributions of these factors to language leaning, processing, and use, and the ways that language as a complex adaptive system has evolved to be learnable. Ellis with Larsen-Freeman (2009) presents various connectionist (Emergent) simulations of the emergence of the VACs described here.

Meanwhile, the findings of these studies provide empirical support for the hypothesis that the learning of linguistic constructions can be understood according to psychological principles of category learning. Learning is sensitive to input frequency, reliabilities of form-function mapping, and prototypicality and generality of function. The structure of language reflects these principles too. It is doubtful that these parallels are accidental, more likely they emerge through usage. A consequence is that in natural language, the type-token frequency distributions of construction islands, their prototypicality and generality of function in these roles, and their reliability of mappings between these, together conspire to optimize learning.

Acknowledgements

I thank my colleagues in this research: Steffi Wulff, Ute Römer and Fernando Ferreira-Junior.

References

Allan, Lorraine G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society, 15,* 147–149.

Andersen, Roger W. (1990). Models, processes, principles and strategies: Second language acquisition inside and outside the classroom. In Bill VanPatten & James F. Lee (Eds.), *Second language acquisition – Foreign language learning* (45–66). Clevedon: Multilingual Matters.

Andersen, Roger W., & Shirai, Yas (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition, 16,* 133–156.

Bardovi-Harlig, Kathleen (2000). *Tense and aspect in second language acquisition: Form, meaning, and use.* Oxford: Blackwell.

Bates, Elizabeth, & MacWhinney, Brian (1987). Competition, variation, and language learning. In Brian MacWhinney (Ed.), *Mechanisms of language acquisition* (157–193).

Biber, Douglas, Conrad, Susan, & Reppen, Randi (1998). *Corpus linguistics: Investigating language structure and use.* New York: Cambridge University Press.

Childers, Jane B., & Tomasello, Michael (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology, 37,* 739–748.

Christiansen, Morten H., & Chater, Nick (2008). Language as shaped by the brain. *Behavioral & Brain Sciences [target article for multiple peer commentary]*, *31*, 489–509.

Cohen, Henri, & Lefebvre, Claire (Eds.) (2005). *Handbook of categorization in cognitive science.* Mahwah, NJ: Elsevier.

Dietrich, Rainer, Wolfgang Klein & Colette Noyau (Eds.). (1995). *The acquisition of temporality in a second language.* Amsterdam: John Benjamins.

Elio, Renee, & Anderson, John R. (1981). The effects of category generalizations and instance similarity on schema abstraction. *Journal of Experimental Psychology: Human Learning & Memory, 7 (6),* 397–417.

Elio, Renee, & Anderson, John R. (1984). The effects of information order and learning mode on schema abstraction. *Memory & Cognition, 12 (1),* 20–30.

Ellis, Nick C. (1998). Emergentism, connectionism and language learning. *Language Learning, 48 (4),* 631–664.

Ellis, Nick C. (2002a). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24 (2),* 143–188.

Ellis, Nick C. (2002b). Reflections on frequency effects in language processing. *Studies in Second Language Acquisition, 24 (2),* 297–339.

Ellis, Nick C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In Cathy Doughty & Michael H. Long (Eds.), *Handbook of second language acquisition* (33–68). Oxford: Blackwell.

Ellis, Nick C. (2006a). Cognitive perspectives on SLA: The Associative Cognitive CREED. *AILA Review, 19,* 100–121.

Ellis, Nick C. (2006b). Language acquisition as rational contingency learning. *Applied Linguistics, 27 (1),* 1–24.

Ellis, Nick C. (2006c). Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27 (2),* 1–31.

Ellis, Nick C. (2008a). The dynamics of language use, language change, and first and second language acquisition. *Modern Language Journal, 41 (3),* 232–249.

Ellis, Nick C. (2008b). Usage-based and form-focused language acquisition: The associative learning of constructions, learned-attention, and the limited L2 endstate. In Peter Robinson & Nick C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (372–405). London: Routledge.

Ellis, Nick C. (2008c). The psycholinguistics of the Interaction Hypothesis. In Alison Mackey & Charlene Polio (Eds.), *Multiple perspectives on interaction in SLA: Second language research in honor of Susan M. Gass* (11–40). New York: Routledge.

Ellis, Nick C., & Larsen-Freeman, D. (2009). Constructing a second language: Analyses and computational simulations of the emergence of linguistic constructions from usage. *Language Learning*, *59*, Supplement 1, 93–128.

Ellis, Nick C., & Ferreira-Junior, Fernando (2009a). Construction learning as a function of frequency, frequency distribution, and function. *Modern Language Journal*, *93*, 370–385.

Ellis, Nick C., & Ferreira-Junior, Fernando (2009b). Constructions and their acquisition: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics*, *7*, 188–221.

Ellis, Nick C. & Larsen-Freeman, Diane (Eds). (2006). Language emergence: Implications for Applied Linguistics. Special issue. *Applied Linguistics, 27 (4).*

Ellis, Nick C. & Larsen-Freeman, Diane (Eds). (2009). Language as a complex adaptive system. Special issue. *Language Learning, 59,* Supplement 1.

Feldweg, Helmut (1991). *The European Science Foundation second language database.* Nijmegen: Max-Planck-Institute for Psycholinguistics.

Gass, Susan (1997). *Input, interaction, and the development of second languages.* Mahwah, NJ: Erlbaum.

Gass, Susan (2002). Interactionist perspectives in SLA. In Robert Kaplan (Ed.), *Handbook of applied linguistics* (170–181). Oxford: Oxford University Press.

Gass, Susan (2003). Input and interaction. In Cathy Doughty & Mike Long (Eds.), *Handbook of second language acquisition* (224–255). Oxford: Blackwell Publishers.

Gass, Susan, & Mackey, Alison (2007). Input, interaction and output: An overview. *AILA Review, 19,* 3–17.

Gass, Susan, Mackey, Alison, & Pica, Terry (1998). The role of input and interaction in second language acquisition: Introduction to the special issue. *Modern Language Journal, 82,* 299–307.

Gass, Susan, & Varonis, Evangeline (1994). Input, interaction and second language production. *Studies in Second Language Acquisition, 16,* 283–302.

Gleitman, Leila (1990). The structural source of verb meaning. *Language Acquisition, 1,* 3–55.

Goldberg, Adele E. (1995). *Constructions: A construction grammar approach to argument structure.* Chicago: University of Chicago Press.

Goldberg, Adele E. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Science, 7,* 219–224.

Goldberg, Adele E. (2006). *Constructions at work: The nature of generalization in language.* Oxford: Oxford University Press.

Goldberg, Adele E., Casenhiser, Devin M., & Sethuraman, Nitya (2004). Learning argument structure generalizations. *Cognitive Linguistics, 15,* 289–316.

Gries, Stefan Th., & Stefanowitsch, Anatol (2004). Extending collostructional analysis: a corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics, 9,* 97–129.

Gries, Stefan Th., & Wulff, Stefanie (2005). Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora. *Annual Review of Cognitive Linguistics, 3,* 182–200.

Klein, Wolfgang, & Purdue, Clive (1992). *Utterance structure: developing grammars again.* Amsterdam: John Benjamins.

Lakoff, George (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Langacker, Ronald W. (1987). *Foundations of cognitive grammar: Vol. 1. Theoretical prerequisites.* Stanford, CA: Stanford University Press.

Li, Ping, & Shirai, Yas (2000). *The acquisition of lexical and grammatical aspect.* Berlin: Walter De Gruyter.

Mackey, Alison (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21,* 557–587.

Mackey, Alison, & Gass, Susan (2006). Pushing the methodological boundaries in interaction research: An Introduction to the special issue. *Studies in Second Language Acquisition, 28,* 169–178.

MacWhinney, Brian (1987). The Competition Model. In Brian MacWhinney (Ed.), *Mechanisms of language acquisition* (249–308). Hillsdale, NJ: Erlbaum.

MacWhinney, Brian (2000). *The CHILDES project: Tools for analyzing talk, Vol 1: Transcription format and programs (3rd ed.)*: (2000). xi, 366 pp.

Murphy, Gregory L. (2003). *The big book of concepts.* Boston, MA: MIT Press.

Ninio, Anat (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language, 26,* 619–653.

Ninio, Anat (2006). *Language and the learning curve: A new theory of syntactic development.* Oxford: Oxford University Press.

Perdue, Clive (Ed.). (1993). *Adult language acquisition: Crosslinguistic perspectives.* Cambridge: Cambridge University Press.

Pickering, Martin J. (2006). The dance of dialogue. *The Psychologist, 19,* 734–737.

Pickering, Martin, & Garrod, Simon C. (2006). Alignment as the basis for successful communication. *Research on Language and Computation, 4,* 203–228.

Pinker, Steven (1989). *Learnability and cognition: The acquisition of argument structure.* Cambridge, MA: Bradford Books.

Posner, Michael I., & Keele, Stephen W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353–363.

Posner, Michael I., & Keele, Stephen W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology, 83,* 304–308.

Rescorla, Robert A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology, 66,* 1–5.

Robinson, Peter, & Ellis, Nick C. (Eds.). (2008). *A handbook of cognitive linguistics and second language acquisition.* London: Routledge.

Shanks, David R. (1995). *The psychology of associative learning.* New York: Cambridge University Press.

Simpson, Rita, Briggs, Sarah, Ovens, Janine, & Swales, John M. (2002). The Michigan Corpus of Academic Spoken English. [Electronic Version]. The Regents of the University of Michigan from http://www.hti.umich.edu/m/micase.

Sinclair, John (1991). *Corpus, concordance, collocation.* Oxford: Oxford University Press.

Sinclair, John (2004). *Trust the text: Language, corpus and discourse.* London: Routledge.

Stefanowitsch, Anatol, & Gries, Stefan Th. (2003). Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics, 8,* 209–243.

Tomasello, Michael (1992). *First verbs: A case study of early grammatical development.* New York: Cambridge University Press.

Tomasello, Michael (2003). *Constructing a language.* Boston, MA: Harvard University Press.

Wilson, Stephen (2003). Lexically specific constructions in the acquisition of inflection in English. *Journal of Child Language, 30,* 75–115.

Wulff, Stefanie, Ellis, Nick C., Römer, Ute, Bardovi-Harlig, Kathleen, & LeBlanc, Chelsea (2009). The acquisition of tense-aspect: Converging evidence from corpora and telicity ratings. *Modern Language Journal*, *93*, 354–369.

Zipf, George K. (1935). *The psycho-biology of language: An introduction to dynamic philology.* Cambridge, MA: The M.I.T. Press.

CHAPTER 3

The role of relevance theory in SLA studies

Maria Jodłowiec
Uniwersytet Jagielloński

1. Introduction

Pragmalinguistic theories, which define as one of the major objectives of their inquiry the accounting for cognitive processes that are at play during verbal communication, can be expected to provide an important theoretical foundation for SLA studies, which aim at identifying and optimizing the parameters crucial to L2 learning. The major goal of the paper is to explore interfaces between a pragmalinguistic model, namely Sperber and Wilson's (1986/95, 1987, Wilson & Sperber 2004) relevance theory and theories elucidating the development of competence in L2. Taking as a point of departure the key concepts of the relevance-theoretic framework, I will attempt to show how the model affords significant insight into the psycholinguistic and metapsychological processes underlying the production and comprehension of language, which may help explain the mechanisms underpinning language learning. I will furthermore try to argue that relevance-theoretic thinking ties in with and reinforces the conceptual approach to L2 phenomena advanced along emergentist lines.

As many researchers have pointed out, second language studies need to be framed within a context of theories that explore the mental representations underlying linguistic performance (cf., among others, Ellis, this volume, Ellis, 2003, Ellis & Larsen-Freeman, 2006, Doughty & Long, 2003, Hall et al., 2006, Jourdenais, 2001). The model of human communication that Sperber and Wilson have developed is based on a set of assumptions related to the psychological factors underlying general cognitive processes; therefore it seems only logical to assume that it should offer useful theoretical support for SLA theory and practice.

The paper is broken into 5 sections. First, the fundamental assumptions about human cognition that relevance theory is based on are presented. The most essential relevance-theoretic principles of the model of overt intentional communication are discussed then. In Section 3 some important earlier applications of the relevance-theoretic apparatus in the field of second language studies are focused on. Metarepresentation as a vital aspect of the relevance-theoretic analyses of the comprehension process is presented next. Major emergentist ideas as advanced in SLA studies are briefly sketched in the subsection that follows. In this part similarities between relevance and the new tendencies

in theories of second language are also pointed out with a view to suggesting that emergentist theorising can be fruitfully supplemented by the relevance-theoretic principles, and the metarepresentational abilities that relevance theorists explore might provide the basis for scrutinising the adaptive nature of second language growth. The paper ends with brief conclusions.

2. Relevance theory: Basic assumptions and claims about human cognition

Relevance theory, as a post-Gricean pragmatic model (cf. Carston & Powell, 2005), while accepting some general assumptions of Grice (to be referred to below), reduces a set of maxims originally proposed by the philosopher of language to a single maxim named the Principle of Relevance. Before this pivotal notion can be introduced, a general background for Sperber and Wilson's pragmatics needs to be discussed.

Relevance theory is rooted in some important observations about how human beings function cognitively, how they interact with the environment, and how the short-term and long-term cognitive goals that they pursue constrain the complex language processing machinery that lies at the heart of verbal communication.

The underlying axiom in this framework is the assumption that there is an important property of inputs that the human mind chooses to attend to and this is relevance (Wilson & Sperber, 2002: 600–601): while there is a whole panoply of various stimuli impinging on the human senses at any given moment of wakefulness, only some are taken notice of by the individual. The originators of relevance theory contend (Wilson & Sperber, 2002: 601) that cognitive processing is programmed towards maximizing the cognitive effects that the individual can achieve by aiming at "the best expected cost/benefit ratio." They postulate that in the course of phylogenetic development, *homo sapiens* have developed a mechanism of cognitive efficiency (Sperber & Wilson, 2006: 178; cf. also Sperber & Hirschfeld, 1999), whereby

> the human cognitive system has developed in such a way that our perceptual mechanisms tend automatically to pick out potentially relevant stimuli, our memory retrieval mechanisms tend automatically to activate potentially relevant assumptions, and our inferential mechanisms tend spontaneously to process them in the most productive way (Wilson & Sperber, 2004: 610).

In other words, when attending to what one sees, hears, or smells, when focusing on some pieces of information stored in memory, and/or when drawing inferences from inputs that seem worthwhile, the organism functions in the relevance-oriented mode, which has to do with minimising effort and maximising cognitive gains (cf. Sperber & Wilson, 2006: 176–9).

This suggests that relevance theory rests on the premise that there is a cognitive efficiency principle underpinning the mental functioning of man. This premise is formalised within the relevance-theoretic framework (Wilson & Sperber, 2004: 610) as the Cognitive Principle of Relevance, in accordance with which "human cognition tends to be geared to the maximisation of relevance."

As hinted at above, relevance is a comparative notion, which as Carston (2002: 44) puts it, is "a positive function of cognitive effects and a negative function of the processing effort expended in deriving these effects." This means that the more effort the individual needs to expend in deriving certain cognitive effects, the lower the relevance, and, on the other hand, the bigger the cognitive effects that processing yields, the greater the relevance (cf. Sperber & Wilson, 1987: 697, Wilson & Sperber, 2002: 602). Thus the organism does not only have a tendency to attend to phenomena that are potentially relevant, but also, by trying to maximize relevance, it optimizes the mental processing to be employed in attending to them (cf. Sperber & Wilson, 1987: 703, 2006: 177).

3. The relevance-theoretic model of human communication

The Cognitive Principle of Relevance as presented above is the corner-stone of the model of communication that Sperber and Wilson have developed. Since in overt intentional communication, which is labelled on this approach ostensive communication, the communicator openly claims the addressee's attention, she creates expectations of relevance in the recipient of the message. In other words, as recovering the speaker meaning inevitably requires an expenditure of effort on the interpreter's part, and given that human cognition is relevance-oriented, the comprehender is entitled to expect that what is communicated comes with a guarantee that the message is believed by the communicator to be relevant to the recipient. This tacit expectation is formalised in relevance theory (Wilson & Sperber, 2004: 612) as the Communicative Principle of Relevance, which states that "every ostensive stimulus conveys a presumption of its own optimal relevance." The claim, as Wilson explicitly formulates it (2005: 1137), is then that in the ostensive communicative context, in which the speaker does not "merely intend to convey certain information, but must intend her audience to recognise that she has this intention", stimuli that are used to convey the communicator's meaning are supposed to be treated by the audience as optimally relevant to them.

What does it mean that in verbal communication utterances are assumed by the hearer to be optimally relevant? The answer to this question follows directly from how the presumption of optimal relevance is to be understood. The presumption of optimal relevance licenses the interpreter to treat a given ostensive stimulus produced by the communicator as relevant enough to be worth processing, and to approach it as the most relevant stimulus compatible with the communicator's abilities and preferences (Wilson & Sperber, 2004: 612). What is postulated then is that the recipient of an ostensively produced signal automatically accepts that the effort invested in the interpretation process will yield some cognitive effects and that a given stimulus should be identified as the most relevant one that the communicator can provide. Certainly, more often than not, communicators may not supply information that would be the most relevant from the interpreter's point of view: they may not have this kind of information or may not be willing to disclose it (cf. Wilson & Sperber, 2004: 611–13); the important thing is that the recipient is to take it for granted that his effort will be offset by the cognitive gains which, as the rational communicator predicts, will at least make it relevant enough for the comprehender. It must also

be stressed that the presumption of optimal relevance does not posit that speakers consciously explore a whole range of potential utterances and choose the maximally relevant one; all that is postulated is that the addressee should treat a given stimulus as the most relevant one that under the circumstances suited the speaker's preferences and abilities.

The Communicative Principle of Relevance and the presumption of optimal relevance provide the theoretical underpinnings for the relevance-theoretic utterance comprehension heuristics, which purports to explain how in overt intentional communication the hearer recovers the speaker intended meaning. As Sperber and Wilson (2005: 360) contend, in accordance with this heuristics, the hearer is assumed to "follow a path of least effort in constructing an interpretation of the utterance (and in particular in resolving ambiguities and referential indeterminacies, in going beyond linguistic meaning, in supplying contextual assumptions, computing implicatures, etc.) in order of accessibility, and … [to] stop when [the interpreter's] expectations of relevance are satisfied" (or abandoned). This heuristics predicts that the first interpretation found relevant enough by the interpreter will be recognised as the one intended by the communicator, to the exclusion of other possible interpretations that the utterance might be compatible with (Wilson & Sperber, 2002: 605, Žegarac, 2006: 1703).

Utterance interpretation embraces a number of inferential tasks, resulting in the hearer recovering the meaning which he finds relevant enough in a given communicative context and which accords with the expectation that the speaker has produced a maximally relevant verbal stimulus. These subtasks involve formulating adequate hypotheses about both the explicit and the implicit content communicated by the speaker. Thus it is postulated that the relevance-constrained interpretation process may embrace (Wilson & Sperber, 2004: 615):

a. Constructing an adequate hypothesis about the explicit content (EXPLICATURES) via decoding, disambiguation, reference resolution, and other pragmatic enrichment processes.
b. Constructing an appropriate hypothesis about the intended contextual assumptions (IMPLICATED PREMISES).
c. Constructing an appropriate hypothesis about the intended contextual implications (IMPLICATED CONCLUSIONS)

At the level of explicit import then, explicatures are formed. An explicature is defined by Carston (2002: 377) as "an ostensively communicated assumption which is inferentially developed from one of the incomplete conceptual representations (logical forms) encoded by the utterance." Explicatures result from the development of a linguistically encoded meaning to full propositionality (Ariel, 2002: 1005), and they are, as Carston (2004: 636) emphasises, "an amalgam of decoded linguistic meaning and pragmatically inferred meaning." At the level of implicitly communicated import, implicated premises and implicated conclusions will be recovered. If a certain background assumption is accessed to form the context in which what is said will be processed then it is recognised to function as an implicated premise. On the other hand, on this approach an inference derived from contextual assumptions combined with what is explicitly communicated will be identified as an implicated conclusion. Verbal comprehension is thus assumed to involve mutual

adjustment of the explicit and implicit content (cf. Carston, 2002, Wilson, 2005, Wilson & Sperber, 2002).

Sperber and Wilson argue that the process of on-line utterance interpretation, which is performed automatically and spontaneously, is carried out by a dedicated autonomous special-purpose module in the human mind (Sperber, 2000, Sperber & Wilson, 2002, Wilson, 1999, 2005), functioning as what Wilson (2005: 1140) calls "a fast and frugal heuristics" (cf. also Sperber & Wilson, 2002).

Summing up, relevance theory is a psycholinguistic model of human communicative behaviour, which is rooted in the general assumptions about and observations of human cognitive functioning. As a cognitive theory of language production and comprehension, it has proved instrumental in second language acquisition research, the issue to be addressed in the section that follows.

4. Second language studies and relevance: Earlier applications

Over the last decade many researchers have used the analytic tools of the relevance-theoretic framework to describe and explain a number of L2 learning/acquisition phenomena (cf. online bibliography http://www.ua.es/personal/francisco.yus/rt.html). Because of obvious space limitations, the presentation below will focus only on selected publications.

In an important paper about the development of pragmatic knowledge in L2, Paiva (2003) shows that as a model of communication deeply rooted in cognitive and psycholinguistic considerations, relevance theory offers important analytic tools to account for processes involved in taking in language input. These processes are claimed to be crucial for the development of interactive abilities in the second language learning context. Emphasising that input needs to be approached both as a product of interaction and part of the cognitive environment in which the non-native speaker functions, Paiva (2003) argues that Sperber and Wilson's framework has a predictive explanatory power, which may provide essential support for theories of second language acquisition focusing on the role of attention and inferencing. While Schmidt's (1990) noticing hypothesis and Bialystok's (1994) model of L2 proficiency development have an established status as influential theoretical proposals in the field of SLA, they lack the principle that might explain in a psychologically plausible way how selective attention works and what triggers the inferencing that language users perform. Paiva makes the case for the notion of relevance *per se* and for the Communicative Principle of Relevance as providing the pivotal theoretical axiom missing from Schmidt's and Bialystok's accounts.

Paiva's (2003) ideas on how relevance theory may fruitfully supplement theories of second language competence development are expanded in the later papers by Paiva and Foster-Cohen (2004) and Foster-Cohen (2004).

Thus Paiva's original reflections on how second language theories profit from incorporating and using notions introduced by relevance theory are pursued further by Paiva and Foster-Cohen (2004: 287), who argue that relevance "can complement information processing accounts by offering a plausible theory of cognition and communication which

operates with a notion of internal context … where inferencing processes are central". Foster-Cohen (2004) in turn compares Sperber and Wilson's (1986/95) framework and Clark's (1996) theory as models that might provide relevant insight into how second language users' communicative resources grow. In pinpointing a number of similarities between relevance theory and action theory, Foster-Cohen shows how the cognitive and individualistic approach typical of Sperber and Wilson (1986/95), as opposed to the social and dyadic stance taken by Clark (1996), offers a more realistic and revealing line of analysis for the communicative behaviour characteristic of language learners. Applying the relevance-theoretic tools to studying non-native speaker output, Foster-Cohen provides convincing arguments that this model is well-suited to analysing phenomena underlying L2 production and reception.

While the literature on the relevance of relevance theory to L2 models presented above explores a possible strengthening, support and deeper insight that relevance can provide for SLA studies at the level of theoretical claims and postulates, there have also been successful attempts to apply relevance theory to explain different aspects of interlanguage development. For instance, Liszka (2004) and Žegarac (2004) show that interpreting along relevance-theoretic lines some aspects of grammatical deficits in the L2 production of learners with a specific native language background can explain the nature of the problem. Analysing how German, Japanese and Chinese EFL learners recover explicatures while processing English utterances with different grammatical tenses, Liszka (2004) identifies and explains problems that non-native speakers face and pragmalinguistic strategies that they resort to in searching for optimal relevance. Her study demonstrates how the relevance-theoretic approach makes it possible to predict and elucidate L2 acquisition problems, but it also shows that L2 acquisition processes endorse relevance-theoretic assumptions.

Žegarac (2004) employs the framework of relevance theory to explore the issue of the acquisition of the definite article in English by learners from an article-less first language background. Žegarac brings in a number of convincing arguments which show that relevance theory affords a useful insight into what meanings the article may be used to convey in English. This notoriously difficult area of English grammar, often not adequately presented in many ELT textbooks, can thus be examined in a more fruitful way from the relevance-theoretic perspective. He also shows how some transfer effects may be due to the tacit and idiosyncratic hypotheses that relevance-searching language learners construct, test and reject.

Paiva, Foster-Cohen, Liszka, Žegarac as well as some other researchers (cf., among others, Garcés-Conejos & Bou-Franch, 2004, Sequeiros, 2004, Smith & Tsimpli, 1991) focus on showing how relevance can contribute to the analyses of SLA processes, though their investigations also corroborate relevance-theoretic assumptions. Likewise Taguchi's (2002) empirical findings collected in an L2 context provide support for inferential mechanisms that Sperber and Wilson postulate as typical of verbal comprehension. Taguchi demonstrates that regardless of their proficiency level in the target language, learners invariably process utterances for relevance. However, there is a difference in the strategies that more advanced learners employ in contrast with the less advanced ones: the former

rely more on their background knowledge in recovering the speaker intended meaning whereas the latter tend to rely more on attending to the speaker's intentions.

One of the latest analyses of classroom processes through the microscope of relevance is Niżegorodcew's (2007) treatment of L2 classroom input in terms of optimal relevance. Niżegorodcew (2007: 18–21) argues that the instructional input that L2 students receive will be optimally relevant to them in different ways. She points out that depending on such factors as focus on accuracy, focus on fluency, focus on both fluency and accuracy, or focus on the metalinguistic aspects of utterances produced in class at different stages of the language teaching and learning process, different levels of optimal relevance will be called for. Assuming that in the formal second/foreign language learning setting the teacher provided input comes with a guarantee of its own optimal relevance, the scholar examines the issue of input interpretation along relevance-theoretic lines. Her idea is that in their search for the optimal interpretation of the teacher's input, defined by the author (Niżegorodcew, 2007: 93) as "the language intentionally presented by the teacher to facilitate the process of L2 learning", learners may understand in various ways the language addressed to them by the teacher, depending on how they assess the input for optimal relevance. So, for instance, in some contexts students may be led to engage in cognitive macroprocessing, which involves the learners in focusing on form, which may fulfil an important facilitative function in their language development. This macroprocessing affords them access to what Doughty (2001: 214–224) calls cognitive windows of opportunity, which are postulated to play a vital role in the internalisation of the target language code (cf. also comments on attention, noticing and incidental learning by Godfroid, Housen and Boers, this volume) .

As hinted at above though, language teachers may use instructional input to perform various functions; the intended function, as Niżegorodcew emphasises will be performed if the learner will relevantly identify the teacher's intention. A mismatch between the teacher intended function (ranging from the purely communicative to the corrective to the metalinguistic) to be fulfilled by supplying a certain instructional input and the learner's recognition of her intentions may undermine – if not altogether annihilate – the desirable pedagogic effect, the more so that this function can be realised in an implicit manner.

Both Taguchi (2004) and Niżegorodcew (2007, cf. also 2004) then consider processing geared to optimising the relevance of the verbal input and the intentionality behind the communicative acts performed and interpreted in the L2 context, be it the classroom or a natural discourse situation, as intrinsic properties of language production and reception underlying the development of communicative competence in the target language. In fact, the metapsychological abilities in general and the metacommunicative capabilities in particular that speakers and hearers rely on while communicating are a pivot of Sperber and Wilson's model. In the remaining part of the paper I would like to argue that the newest theoretical developments in SLA studies can be substantiated by the key relevance-theoretic principles.

5. Relevance theory and emergentism in SLA research

The newest theoretical developments in L2 studies referred to above have to do with the emergentist program, which – to introduce it in very general terms first as O'Grady (2008: 459) does –"is based on the simple thesis that the core properties of language are best understood with reference to more fundamental non-linguistic … factors and their interaction." Emergentist thinking in linguistic and applied linguistic investigations encompasses a broad and diverse research agenda that looks at language phenomena from the perspective of complex, non-linear and dynamic systems, trying to account for the language mechanisms under scrutiny by exploring general human cognitive capacities, universal pragmatic principles, and essential social parameters of language, to mention but the most important factors to be considered (cf. MacWhinney, 2006).

Relevance theory and emergentism converge in certain important aspects. Even more importantly, the major concern of emergentism, that is ensuring that SLA theory accounts for the mental representations underlying language processing, can find support from a relevance-theoretic treatment of the metarepresentational abilities that speakers and hearers use in communication.

5.1 Metarepresentational abilities and expectations of relevance

Already the pragmalinguistic framework created by Grice (1957, 1969, 1989) postulated that the inferential processing essential for verbal comprehension is rooted in the metacommunicative abilities that the communicator and the audience necessarily need to possess in order to be competent language users (for a brief discussion of Gricean maxims of verbal comprehension cf. Bromberek-Dyzman & Ewert, this volume). These abilities have to do with the communicator being able to represent the thoughts she wants to convey, which initiates the verbal exchange, and the addressee being able to represent the communicator's intentions, which underpin the interpretation process (Wilson, 1999: 129–130). Since such representations are often representations of representations (thoughts and intentions are mental representations *per se*), strictly speaking they are metarepresentations, with one representation being embedded in the other (Wilson, 1999: 127).

Criticizing Grice's original comprehension schema as psycholinguistically unfeasible (cf., among others, Sperber & Wilson, 1986/95, Wilson, 1999, 2005, Wilson & Sperber, 2002, 2004), relevance theorists fully agree with his idea that metarepresentation constitutes an important part of the spontaneous intuitive reasoning typical of utterance comprehension (Sperber, 1994, Wilson, 1999). They go on to show that interpreting ostensive stimuli in general, and utterances in particular, may involve different levels of metarepresentation. Metarepresentational analyses, as Wilson (1999) proposes, can be used to explain in a uniform and plausible way a whole range of communicative phenomena, in particular those which help to resolve indeterminacies of meaning. It seems that they are also relevant to the processes underlying L2 acquisition/learning, so their nature will be briefly described here.

On the relevance-theoretic approach, communication is defined by Sperber and Wilson (1986/95: 1) as the process "involving two information-processing devices [in

which] one device modifies the cognitive environment of the other." As a result of successful communication, the mental representations in the interpreter's mind are supposed to resemble those in the communicator's mind. Thus the initial state and the final state in communicative exchanges are mental states of the communicator and the comprehender respectively.

With reference to verbal communication, Sperber and Wilson (1986/95, 2002, 2004) claim that the speaker produces an utterance which is a piece of evidence as to the meaning she wants to convey in a given context, and the hearer infers the speaker meaning from this evidence, following the path of least effort and assuming that the utterance is intended to be maximally relevant, as the Principle of Relevance together with the presumption of optimal relevance posit. This means that in communicative situations, by making their thoughts public, speakers reveal their communicative intentions while addressees, by attributing intentions to speakers, infer the intended meaning, automatically searching for an optimally relevant interpretation. In this framework then, utterance interpretation is an exercise in mind-reading (Wilson, 2005, Wilson & Sperber, 2004): language users necessarily need to be able to foresee mental states of interlocutors and, if the need arises, to represent these states and pursue a strategy that will override meanings that cannot have been intended by the communicator (cf. Yus, 2006: 512–13).

What has just been said suggests that recognising the communicator's intentions is an important part of the comprehension procedure. As Sperber (1994) and Wilson (1999, 2005) emphasise, identifying intentions in communicative situations differs in an important way from how intentions are attributed to agents in non-communicative contexts though. The problem is that when deciding what intentions a person has in doing something, we can usually predict the result that they hope to achieve when performing a particular act and in this way we identify their intentions, whereas in interpreting ostensively produced utterances we cannot proceed in this manner. The point is that the speaker's intentions in saying something provide important guidelines for the interlocutor to interpret adequately what has been said. That is why, the hearer cannot rely on the meaning that is intentionally conveyed by the speaker in a given exchange, because it is precisely on the basis of the intentions attributable to the sender of the message that the meaning needs to be worked out. So when Mary takes out a mobile phone from her pocket, if in the circumstances there is nothing that would provide evidence to the contrary or make a competing hypothesis plausible, it can be safely inferred that she is going to make a phone call. On the other hand, when Mary, who happens to be the teacher, utters (1), the meaning she intends to convey must be based on the communicative intention that will be manifest to the audience in a given communicative context. The word *manifest* in relevance theory is a technical term, referring to assumptions which can be mentally represented by an individual and accepted as true or probably true in a given communicative situation (cf. Carston, 2002: 378):[1]

(1) I will phone your parents and tell them everything.

In the conversational setting in which it is mutually manifest to Mary and the audience that Peter has misbehaved the optimally relevant interpretation will yield explicature (1a):[2]

(1) a. THE TEACHER$_i$ HAS THREATENED THAT THE TEACHER$_i$ WILL PHONE PETER$_j$'S PARENTS AND TELL PETER$_j$'S PARENTS EVERYTHING $_{\text{ABOUT PETER}j\text{'S MISBEHAVIOUR}}$.

Observe though that if in the context in which (1) is uttered, it is manifest that Peter has just won the school Mathematics competition and he will be taking part in the Mathematics championship, by uttering (1) his teacher may be trying to reassure the boy and dispel his doubts about whether the parents will approve of his going to the capital city for a week. The optimally relevant interpretation will be completely different and the explicature recovered might instead be something like (1b):

(1) b. THE TEACHER$_i$ HAS REASSURED PETER$_j$ THAT THE TEACHER$_i$ WILL PHONE PETER$_j$'S PARENTS AND TELL PETER$_j$'S PARENTS EVERYTHING $_{\text{ABOUT PETER}j\text{'S PARTICIPATION IN THE MATHEMATICS CHAMPIONSHIP}}$.

This example illustrates how contextual clues, background assumptions, and intentions that can be attributed to communicators contribute to recovering the intended interpretation. Necessarily then the speaker's intentions are part and parcel of the interpretation process and contribute to the recovery of the intended meaning, that is the end product of communication, so they are quite different from non-communicative intentions, which are inferable from the foreseeable results of human actions. If verbal comprehension followed a regular course of intention attribution, the hearer would have to know in the first place what effect the speaker wants to achieve by saying something, and assume that this is her intention, yet usually it is the other way round: the communicator's meaning is recovered by attributing particular intentions to the speaker in a given communicative situation. Thus, in principle, any utterance, just as shown by example (1), may give rise to a number of quite unrelated interpretations.

The discussion above indicates furthermore that utterance interpretation is indeed a mind-reading procedure, even though it should not be understood in terms of a deliberate and reflective process, but rather treated as an automatic and unreflective mental function performed by a specialised dedicated pragmatic module in the individual's mind (cf., among others, Wilson, 1999, 2005, Wilson & Sperber, 2004).

Even though on-line verbal comprehension is a spontaneous and automatic process, there may be various levels of sophistication in interpretation that comprehenders will poise to reach. As Wilson (1999: 137) aptly puts it, "the expectations of relevance created (and adjusted) in the course of the comprehension process may be more or less sophisticated." Relevance theorists argue that while interpreting ostensive stimuli the addressee may ascend to different levels of metarepresentation, which will allow him to arrive at the optimally relevant interpretation. Empirical evidence demonstrates that people with some mental deficiencies, for instance, individuals afflicted with autism or suffering from Asperger's syndrome reveal impairment in the ability to metarepresent in conversational contexts, which severely undermines their communicative functioning (cf. Wilson, 1999, 2005). Metarepresentational capacities seem to be psychologically feasible then.

On the relevance-theoretic approach it is assumed that when deriving optimally relevant meanings, interpreters may take into account some information which is available

to them through metacommunicative insight. In other words, it is postulated within this framework that comprehenders will not blindly fall for the first interpretation that the presumption of optimal relevance predetermines, but on certain occasions, will discard this interpretation as unlikely to be intended as optimally relevant by the communicator and will search further for another optimally relevant meaning. This may happen when an utterance will yield an accidentally optimally relevant interpretation, an accidentally irrelevant interpretation or an interpretation that will merely seem optimally relevant even though it is genuinely not optimally relevant as it appears to the interpreter. All the three instances are cases of a mismatch between the first relevant interpretation accessed and the interpretation that is or might be assumed to be optimally relevant due to metacommunicative considerations, that is thanks to the fact that the interpreter will engage in doing what is referred to in the literature as reading the speaker's mind.

Metacommunicative representation is assumed to be intuitive and, unless an interpreter for some reason (for instance, in an experimental situation) needs to do so, performed subconsciously. The mechanism is triggered when the first interpretation that surfaces in the hearer's mind is identified by him as the interpretation that cannot have been intended as optimally relevant by the speaker. Under such conditions the search for the relevant interpretation continues until a new candidate is accessed. Some examples seem to be in place to illustrate what is involved. Let us assume that (2) is again uttered by Peter's teacher:

(2) I know what you have done.

If Peter, who is the addressee of (2), happens to be thinking of the money that he took from his mum's purse the previous evening, recovers as the first interpretation the teacher's announcing that she knows about his misdemeanour, he is likely to discard this meaning as unfeasible to have been intended by the speaker. He will continue searching for another accessible optimally relevant meaning, which may be – considering the background assumptions highly available to him and taking into account the speaker's communicative intention manifest in the context – the meaning that the teacher knows Peter cheated during the Maths test (somebody must have reported on him).

However, utterances may turn out to be not merely accidentally relevant (as the first interpretation of (2) that Peter accessed was), but accidentally irrelevant, as (3) below illustrates. Let us consider a communicative situation in which after the students have been shown a film on video in a dark classroom, while turning off the video, apparently affected by the action she is performing, the teacher asks Peter:

(3) Turn off the light please, Peter.

The teacher may not even notice her slip of the tongue, but Peter will most probably interpret her utterance correctly, as he will immediately access the interpretation under which the speaker may have intended the utterance to be optimally relevant: the interpretation which will be about turning *on* rather than turning *off* the light.

A more sophisticated strategy of metarepresentation may be called for in communicative contexts in which the speaker will intend the hearer to recover the interpretation which, in the latter's estimation, will not be truly optimally relevant but will merely *seem*

to be so. This kind of metarepresentation will occur in situations in which it will be obvious to the hearer that the first interpretation he derives is a meaning at odds with his knowledge of the world, the information available to him in the environment, contextual assumptions made accessible by previously processed discourse, etc., which will make this interpretation, as if by definition, irrelevant to the recipient. Nevertheless, in the circumstances, a sophisticated interpreter will not abandon this interpretation as unintended, as he will be capable of hypothesising about what interpretation the speaker might have thought the hearer would think was relevant enough. For instance, if Peter misbehaved during the Maths lesson, and he knows that the head teacher and the Maths teacher are barely on speaking terms with each other, he will treat (4) as not true, however, will easily recover the interpretation under which the speaker might have thought he would think the utterance would be relevant.

(4) I am angry because of what you have done.

To recapitulate the postulates of relevance theory with reference to different levels of metarepresentations, it needs to be pointed out that depending on what a comprehender takes to be relevant enough in a given communicative context, the interpretation process may involve more or less complex metarepresentation. The comprehension strategies briefly presented above, which have to do with different levels of metarepresentation that the interpreter will be ready to reach, are triggered by what he considers to be relevant enough in a given communicative situation. At certain times what is accepted as relevant enough will yield an interpretation that is not intended by the speaker: this is how misunderstandings arise – an interesting issue in its own right, which merits a lengthy discussion, exceeding the scope of this paper.

5.2 Metarepresentation and emergentism

I would like to suggest that the theoretical assumptions and claims of the relevance-theoretic framework seem not only to harmonise with but also provide essential support for recent developments in SLA theory. These newest tendencies in the theories of L2 development, as hinted at earlier, appear to be following an emergentist thinking. In the field of cognitive psychology, emergentism is defined in Hollich et al. (2000: 2) as "a process-oriented trend towards more fluid analyses and towards integrative approaches that do not parcel out .. [different kind of] influences, but rather seem to borrow the best from each of the prior theories." Advancing the view that the time is ripe to abandon false dichotomising[3] and to take a more open-minded attitude at the level of metatheoretical claims and postulates, emergentists postulate a convergent, interactionist approach in scientific exploration, which instead of following rigidly one line of theorising will allow for a more flexible, dynamic and eclectically-oriented modelling. Hollich et al. (2000) consider this new approach to herald a paradigm shift in science. Larsen-Freeman (2006: 591) seems to be of a similar opinion: the scholar maintains that emergentism opens novel theoretical perspectives for scientific investigations as it adopts a different supra-theory.

As Hollich et al. emphasise, emergentist theories may come in different guises (cf. also O'Grady, 2008), but they all share what might be called three metatheoretical assumptions (Hollich et al., 2000: 12):

> (1) Simple regularities when iterated can produce extraordinarily varied and complex behavior that is emergent from the interaction at the lower levels (upward causation). (2) Each problem space has its own set of constraints or boundary conditions which serve to limit the behavior of the system and which can produce discontinuous patterns of behavior, or phase shifts, from a single nonlinear process. (3) Finally, there is a beginning realization that often the emergent whole may affect the lower levels as well: downward causation.

These three postulates mirror the postmodern approach to the role of science: the reality is too complex to be reduced to a set of principles, too dynamic to be reduced to a set of states and too unpredictable to be explained through neat algorithms (cf. Larsen-Freeman, 2000).

In the context of SLA, emergentist thinking manifests itself in accepting that language itself is a complex dynamic system (in line with the chaos-theoretic approach), and that language growth and use differs among individuals, and even for the same individual changes unevenly over time, so uniformity and a stage-like character of language development can hardly be expected (cf. Larsen-Freeman, 1997). It is widely acknowledged that the learner's interlanguage is constantly evolving, and its development is affected both by the unique individual characteristics of the student and the universal principles of learning in general and of language learning in particular (cf. discussion on the dynamism of interlanguage phenomena in papers by de Bot & Lowie and Filipović & Vidaković, this volume). This suggests that, as Larsen-Freeman (2006: 592) convincingly argues, "there are no discrete stages at which learners' performance is invariant [though some] stages in the acquisition of certain grammatical structures can be traced." Language learners in fact can be seen – to a greater or lesser extent – as following their own agenda. In a nutshell then, it is believed that learners, to cite Larsen-Freeman's (2006: 594) influential words, "not only determine which aspects of the outside world are relevant to them, but they actively construct a world around themselves and are constantly altering it."

Looking at L2 learning from the vantage point of complex adaptive systems brings to light the non-linearity of language development (with progress being interspersed with regression), the fuzziness of the target language system that the learner has at his/her disposal in different contexts of use, the explicit and implicit teaching/learning that intermingle in language learning environments, the cognitive and the social dimensions of language, which necessarily overlap in communicative contexts, to mention but a few parameters that seem to become more amenable to scrutiny and exploration within the emergentist approach.

The theoretical emergentist assumptions sketched above overlap with some theoretical premises and postulates of relevance theorists. In the first place, relevance researchers, just like applied linguists adopting an emergentist stance, view language as an important cognitive resource, used for communicative purposes.

Some striking parallels can be traced in how emergentism and relevance look at context: both approaches postulate that context should be treated as a dynamic and changing parameter of discourse. In relevance-theoretic analyses, verbal comprehension involves accessing background assumptions that form context for interpretation, and context is not given but chosen and modified as interpretation proceeds (Sperber & Wilson, 1986/95, Wilson & Sperber 2004); in emergentist theory context affects linguistic performance of the individual who is in turn influenced by context (Larsen-Freeman, 2006). Relevance as a property of the inputs to cognitive processing lies at the heart of the cognitive model propagated by Sperber and Wilson; echoes of the importance of what is relevant to an individual for L2 processes reverberate in the emergentist approach, with Larsen-Freeman (2006: 594) herself contending that "individuals … determine what aspects of the outside world are relevant to them."

Yet some affinities between the two theoretical strands like those mentioned above (and possibly some others) are not the only reason for which SLA researchers might find relevance theory worth their while. It seems that the theoretical apparatus of relevance framework as well as the principles and analytic tools that this model affords could be used to support emergentist ideas. While, as it has been hinted at above, emergentism in SLA by definition defies rigidity and uniformity, it necessarily needs to be based on psychologically plausible principles of cognitive functioning, which will help explain in a homogenous way the very individual and idiosyncratic paths that learners follow in their L2 development (possibly also in their L3 acquisition, some intricacies of which are explored in Filatova's paper in this volume). I would like to suggest that relevance theory, as a psychologically plausible model of human cognition and communication, may provide a sound foundation and methodological support for emergentist theories of SLA, which – as MacWhinney (2006: 733) remarks – they need. As Paiva (2003) and Foster-Cohen (2004) rightly emphasise, and as emergentists also endorse (Ellis, 2003, Ellis & Larsen-Freeman, 2006, Larsen-Freeman, 2006), SLA theories need to be anchored in models of general human cognitive and communicative capacities: relevance theory is precisely this kind of framework, so it should be seen as relevant to emergentist pursuits. To be straightforward about it, whereas relevance theory has developed a theoretical apparatus suitable to explore mental reality and useful in explaining the cognitive processes going on in speakers' and hearers' minds, emergentists appear to lack such analytic tools.

It seems in particular that the metarepresentational abilities postulated and applied to scrutinising comprehension procedure, which have been presented in an earlier section, might be fruitfully incorporated into analyses of how awareness of second language form and function emerges from the processing of target language input. In adopting a relevance perspective, it can be speculated that the target language input that L2 learners are exposed to will be processed for relevance, and the level of relevance that will meet the individual's expectations will vitally influence the mental representations to be generated.

Plausibly, the metapsychological capabilities that Sperber and Wilson posit as underlying communicative behaviour can be traced in classroom discourse contexts (cf. Robinson's remarks on task complexity and learners' abilities to theorise other minds, this volume). After all, mechanisms underpinning inferential processing will necessarily be evoked in second language production and comprehension. That is why recognition of the teacher's or another student's or the native-speaker interlocutor's intention in

saying something, and adjusting the level of relevance to the individual expectation of the learner, may help explain why the same input that the students are exposed to in virtually the same conditions will not result in the same intake across these different students, as emergentists are quick to point out (cf. Larsen-Freeman, 2006). Likewise conditions that optimise learning outcomes might probably be fruitfully studied from the relevance-theoretic perspective: it seems that the facilitative conditions that obtain when learners function in the zone of proximal development might be explained by exploring the relevance concerns that may be postulated to determine non-native speaker's functioning in talking to native speakers.

6. Concluding remarks

As a model that brings cognition and communication together, relevance theory has been recognised to offer valuable support for SLA models and it has been used successfully by some second language researchers to account for different aspects of L2 development. Thus the range and originality of the relevance-theoretic model of human communication has been appreciated not only in opening interesting lines of investigation into various theoretical issues pertaining to second language development, but also in illuminating some important practical aspects directly related to target language instruction.

More recent tendencies in SLA theorising which are rooted in emergentism may also profit from turning to the relevance-theoretic framework. Relevance-orientation, postulated by Sperber and Wilson to underlie and constrain human cognition, might be incorporated into emergentist SLA models and deployed in investigating dynamic, non-linear, erratic and chaotic (in the chaos-theoretic sense) L2 growth. The fundamental affinities between relevance theory and emergentist thinking could be capitalised on at a more general level and at a more practical level; in particular the metapsychological and metarepresentational capacities that language users possess might constitute theoretically useful concepts to be incorporated into applied linguistic empirical investigations.

To conclude with an apt quotation from Hollich et al. (2000: 26), "[t]he real change in cognition and language development comes in the form of increased attention to multiple and mutually reinforcing sources of linguistic information that act in concert to ensure development." It seems that relevance theory could aid in reading the score for the different instruments playing this concert.

Notes

1. These assumptions may be accessible to the individual from an immediate physical context (linguistic and/or extralinguistic), from information stored in their mental encyclopaedia, from the utterance they are processing, etc.

2. In the model under discussion it is a convention to use capitals in spelling out explicatures as conveyed by utterances.

3. Mirrored, for example, in the ungrounded postulates favouring either input or output in the L2 classroom.

References

Ariel, Mira (2002). Privileged interactional interpretations. *Journal of Pragmatics, 34,* 1003–1044.

Bialystok, Ellen (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition, 16 (2),* 157–168.

Bromberek-Dyzman, Katarzyna & Anna Ewert (this volume). *Figurative competence is better developed in L1 than in L2, or is it? Understanding conversational implicatures in L1 and L2.*

Carston, Robyn (2002). *Thoughts and utterances.* Oxford: Blackwell.

Carston, Robyn (2004). Relevance theory and saying/implicating distinction. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (633–656). Oxford: Blackwell.

Carston, Robyn & George Powell (2005). Relevance theory – new directions and developments. *UCL Working Papers in Linguistics, 17,* 279–299.

Clark, Herbert (1996). *Using language.* Cambridge: Cambridge University Press.

de Bot, Kees & Wander Lowie (this volume). *Stability of representation in the multilingual lexicon: Cattell (1886) revisited.*

Doughty, Catherine (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (206–257). Cambridge: Cambridge University Press.

Doughty, Catherine & Michael Long (2003). The scope of enquiry and goals of SLA. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (3–16). Oxford: Blackwell.

Ellis, Nick (1998). Emergentism, connectionism and language learning. *Language Learning, 48,* 631–64.

Ellis, Nick (2003). Construction, chunking and connectionism: The emergence of second language structure. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (63–103). Oxford: Blackwell.

Ellis, Nick (this volume). *Constructing a second language: Learning as a function of frequency, frequency distribution, form and function.*

Ellis, Nick & Diane Larsen-Freeman (2006). Language emergence: Implications for applied linguistics. *Applied Linguistics, 27 (4),* 558–589.

Filatova, Ksenia L. (this volume). *Third language acquisition, macro-categories and synonymy.*

Filipović, Luna & Ivana Vidaković (this volume). *Typology in the L2 classroom: Second language acquisition from a typological perspective.*

Foster-Cohen, Susan (2004). Relevance Theory, Action Theory and second language communication strategies. *Second Language Research, 20 (3),* 289–302.

Garcés-Conejos, Pilar & Patricia Bou-Franch (2004). A pragmatic account of listenership: Implications for foreign/second language teaching. *Revista Alicantina de Estudios Ingleses, 17,* 81–102. <http://www.ua.es/personal/francisco.yus/rt.html#second.> accessed July 2008.

Godfroid, Aline, Alex Housen & Frank Boers (this volume). *A procedure for testing the Noticing Hypothesis in the context of vocabulary acquisition.*

Grice, Herbert Paul (1957). Meaning. *Philosophical Review, 66,* 377–388.

Grice, Herbert Paul (1969). Utterer's meaning and intentions. *Philosophical Review, 78,* 147–177.

Grice, Herbert Paul (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Hall, Joan, An Cheng, & Matthew Carlson, M. (2006). Reconceptualising multicompetence as a theory of language knowledge. *Applied Linguistics, 27 (2),* 220–240.

Hollich, George, Kathy Hirsh-Pasek, Michael Tucker, & Roberta Golinkoff (2000). The change is afoot: Emergentist thinking in language acquisition. In P. B. Anderson, C. Emmeche, N. O. Finnemann, & P. V. Christiansen (Eds.), *Downward causation* (143–178). Oxford, England: Aarhus University Press. <http://hincapie.psych.purdue.edu/HHPTTuckerG00-DC.pdf > accessed November 2007.

Jourdenais, Renee (2001). Cognition, instruction and protocol analysis. In P. Robinson (Ed.), *Cognition and second language instruction* (354–375). Cambridge: Cambridge University Press.

Larsen-Freeman, Diane (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics, 18 (2),* 141–165.

Larsen-Freeman, Diane (2000). An attitude of inquiry: TESOL as a science. *The Journal of the Imagination in Language Learning and Teaching.* <http://www.njcu.edu/CILL/vol5/larsen-freeman.html> accessed November 2005.

Larsen-Freeman, Diane (2006). Emergence of complexity, fluency, and accuracy in the oral production of five Chinese Learners of English. *Applied Linguistics, 27 (4),* 590–619.

Liszka, Sarah (2004). Exploring the effects of first language influence on second language pragmatic processes from a syntactic deficit perspective. *Second Language Research, 20 (3),* 213–231.

MacWhinney, Brian (2006). Emergentism – use often and with care. *Applied Linguistics, 27 (4),* 729–740.

Niżegorodcew, Anna (2004). L2 classroom discourse in the light of input theory and relevance theory. In E. Mioduszewska (Ed.), *Relevance studies in Poland* (307–315). vol. 1. Warszawa: Uniwersytet Warszawski.

Niżegorodcew, Anna (2007). *Input for instructed L2 learners: The relevance of relevance.* Clevedon: Multilingual Matters.

O'Grady, William (2008). The emergentist program. *Lingua, 118,* 447–464.

Paiva, Beatriz Mariz de (2003). Pragmatic interaction in a second language. In C. Grant (Ed.), *Rethinking communicative interaction: New interdisciplinary horizons* (187–206). Amsterdam, Philadelphia: John Benjamins.

Paiva, Beatriz Mariz de & Susan Foster-Cohen (2004). Exploring the relationships between theories of second language acquisition and relevance theory. *Second Language Research, 20 (3),* 281–288.

Robinson, Peter (this volume). *Task complexity, intentional reasoning demands, L2 speech production, learning and syllabus design.*

Schmidt, Richard (1990). The role of consciousness in second language learning. *Applied Linguistics, 11,* 129–158.

Sequeiros, Xose Rosales (2004). Interpretation of reflexive anaphora in second language VP-ellipsis: Relevance theory and paradigms of explanation. *Second Language Research, 20 (3),* 256–280.

Smith, Neil & Ianthi-Maria Tsimpli (1991). Second-language learning: Evidence from a polyglot savant. *UCL Working Papers in Linguistics, 3,* 171–183.

Sperber, Dan (1994). Understanding verbal understanding. In J. Khalfa (Ed.), *What is intelligence?* (179–198). Cambridge: Cambridge University Press.

Sperber, Dan (2000). Metarepresentations in an evolutionary perspective. In D. Sperber (Ed.), *Metarepresentations: A multidisciplinary perspective* (117–137). Oxford University Press, <http://www.dan.sperber.com/metarep. > accessed September 2005.

Sperber, Dan & Lawrence Hirschfeld (1999). Evolution, cognition and culture. In R. Wilson & F. Keil (Eds.), *MIT encyclopedia of the cognitive science (cxi–cxxxii.).* Cambridge, Mass.: MIT Press.

Sperber, Dan & Deirdre Wilson (1986/1995). *Relevance: Communication and cognition.* Oxford: Blackwell.

Sperber, Dan & Deirdre Wilson (1987). Précis of *Relevance: Communication and cognition*. *Behavioral and Brain Sciences, 10,* 697–754.

Sperber, Dan & Deirdre Wilson (2002). Pragmatics, modularity and mind-reading. *Mind and Language, 17,* 3–23.

Sperber, Dan & Deirdre Wilson (2005). Pragmatics. *UCL Working Papers in Linguistics, 17,* 353–388.

Sperber, Dan & Deirdre Wilson (2006). A deflationary account of metaphor. *UCL Working Papers in Linguistics, 18,* 171–203.

Taguchi, Naoko (2002) An application of relevance theory to the analysis of L2 interpretation processes: The comprehension of indirect replies. *IRAL, 40,* 151–176.

Wilson, Deirdre (1999). Metarepresentation in linguistic communication. *UCL Working Papers in Linguistics, 11,* 127–161.

Wilson, Deirdre (2005). New directions for research on pragmatics and modularity. *Lingua, 115,* 1129–1146.

Wilson, Deirdre & Dan Sperber (2002). Truthfulness and relevance. *Mind 111,* 583–632.

Wilson, Deirdre & Dan Sperber (2004). Relevance theory. In L. Horn & G. Ward (Eds.), *The handbook of pragmatic* (607–632), Oxford: Blackwell.

Yus, Francisco (2006). Relevance theory. In K. Brown (Ed.), *Encyclopedia of language and linguistics (2nd edition)* (512–519). Amsterdam: Elsevier.

Žegarac, Vlad (2004). Relevance Theory and the in second language acquisition. *Second Language Research, 20 (3),* 193–211.

Žegarac, Vlad (2006). Believing in: A pragmatic account. *Lingua, 116,* 1703–1721.

CHAPTER 4

Distinct mechanisms in the processing of English past tense morphology

A view from L2 processing*

Bilal Kırkıcı
University of Oxford

1. Introduction

Theories of the mental representation and processing of language fall, broadly speaking, into two classes: *associationist theories* like the connectionist models of Rumelhart and McClelland (1986) and Plunkett and Marchman (1991, 1993, 1996), and *rule-and-representation theories* such as the dual-mechanism model (Pinker & Prince, 1988, 1991; Pinker, 1991; Marcus et al., 1995). Proponents of connectionist models basically claim that linguistic processing can be accounted for by a single associative mechanism that is responsive to properties of the stimulus, such as frequency of occurrence, without resorting to linguistic rules and symbolic representations (Alegre & Gordon, 1999). Advocates of rule-and-representation theories like the dual-mechanism model, on the other hand, mainly argue for the psychological reality of rules, which manipulate symbols, in addition to an associative memory (Zobl, 1998).

The bulk of empirical and theoretical research surrounding the above stated 'connectionist-symbolist debate' (Clahsen, 1995), which is also known as the 'past tense debate' (Pinker & Ullman, 2002), has focused on inflectional morphology, and particularly on the representation and (L1) acquisition of English past tense morphology. The main reason for this prominence of the English past tense is that the inflectional processes within the English past tense include two descriptively distinct systems (regular and irregular past) that compute independently of other linguistic subsystems like syntax, semantics, and phonology (Pinker, 1991). Hence, they lend themselves particularly well for investigating whether or not two distinct mechanisms are at work, as predicted by the dual-mechanism theory. In acquiring the English past tense morphological system, children characteristically begin by correctly producing a small number of both regular and irregular forms, then produce typically 'over-regularized' forms (like *breaked* or *bringed*) for a small but noteworthy rate of their verb forms.[1] They then appear to re-learn the correct forms (or unlearn the over-regularized forms), producing the classic 'U-shaped developmental profile' (Berko, 1958; Marcus et al., 1992).

The dual-mechanism model takes the above-mentioned emergence of over-regularized forms as an indication of the psychological reality of the *-ed* suffixation rule hypothesized by grammarians (Marcus et al., 1992). From a dual-mechanism perspective, the first language (L1) acquisition of English past tense morphology reveals that two psychological processes are at work, one for regular and one for irregular morphology. It is proposed that irregular past tense forms (like *grew* and *taught*) are stored undecomposed in the mental lexicon but are connected to their stems over an associative memory. The production of irregular past tense forms is executed by retrieving the necessary form from memory. Regular past tense forms (like *looked* or *walked*), on the other hand, are stored in the mental lexicon in their stem forms (i.e., *look* and *walk*) and are produced in real-time via a symbol-manipulation system that applies the regular inflection (the *-ed* suffix) to the verb stem (Pinker & Prince, 1988; Pinker, 1991; Marcus et al., 1992).

In other words, it is argued that irregular verb forms are directly retrieved from an associative memory, while regular verb forms are produced by means of a general symbolic rule that operates on verbs and reads as "add /d/ to V (verb)." This rule can, in principle, be productively applied to any verb, no matter if familiar or unfamiliar, similar or not to remembered regular verbs (Marcus et al., 1995). The problem of over-regularization, on the other hand, is prevented by means of a *blocking* mechanism, which blocks the application of the general rule to the stem when memory already contains a more specific form. Thus, in terms of the English past tense, if memory (or the mental lexicon) already contains an irregular past tense form, the blocking mechanism blocks the application of the *default*, general rule "add /d/ to V", thus preventing the production of over-regularized forms like *breaked* or *goed*. However, if no entry exists for a past tense form in memory, the general rule "add /d/ to V" is applied (Marcus et al., 1992).

Evidence in support of the above-stated theoretical tenets of the dual-mechanism model has been sought in a wide range of studies like research focusing on child language acquisition data (e.g., Marcus et al., 1992), elicitation and priming experiments with adults and children (e.g., Ullman, 1999), neurolinguistic research like lesion studies (e.g., Aphasia), studies focusing on developmental disorders (e.g., Specific Language Impairment or Williams Syndrome), and neuroimaging studies (Ullman, Corkin, Pinker, Coppola, Locascio & Growden, 1993, 1994; Clahsen & Almazan, 1998; Say, 2000; Ullman, 2001a). Most of these psycholinguistic and neurolinguistic studies of the L1 processing of the English past tense have largely obtained results that independently support the view of the dual-mechanism model that distinct mental mechanisms underlie the production of regular and irregular past tense forms.[2]

2. A Dual-Mechanism in L2 processing?

Needless to say, if the tenets of the dual-mechanism model are meant to carry universal validity, they should also hold for the acquisition and processing of a second language. Thus, just as in L1 processing, it should be possible to encounter similar manifestations of two distinct mechanisms in the linguistic processing of an L2. It is a well-known fact that adult L2 learners usually fail to reach a final state that matches that of native language

learners and especially show deficits in certain areas of implicit grammatical competence rather than in the use of lexical items (Johnson & Newport, 1989; Beck, 1997, 1998). It has been proposed that this is due to the fact that grammatical processing is more age-of-exposure sensitive than lexical processing in L2 and, thus, results in a greater reliance on associative memory in adult L2 learners (Ullman, 2001b). If this hypothesis, as it stands, is correct, it implies that adult L2 learners to a certain extent eschew symbol manipulation of the kind applied in regular verb formation and compute regular as well as irregular verb forms over an associative memory (Beck, 1997). In other words, findings supporting the above view would constitute counter-evidence for the theory that a dual-mechanism is also at work in L2 processing.

Beck (1997) and Zobl (1998) investigated this hypothesis for L2 English and obtained results largely speaking against this view and for the presence of a dual-mechanism in L2.[3] Beck (1997) conducted a series of reaction time experiments in which adult native speakers and high-proficiency adult L2 speakers of English from various L1 backgrounds were mainly asked to produce verbally the simple past tense forms of verb stems presented on a computer. To be able to measure potential frequency effects, subjects were presented with high-frequency and low-frequency regular and irregular verbs. It was found that neither L1 nor L2 subjects showed significant differences in the reaction times for high-frequency and low-frequency regular verbs. This finding supports the view that regular verb forms are frequency-insensitive and, thus, produced by the application of a general rule in L1 as well as L2 processing.

Interestingly, however, Beck (1997) did not find any significant differences in the reaction times for high-frequency and low-frequency *irregular* verb forms with L2 subjects, either. Beck attributes this rather unexpected result to the fact that all L2 subjects in her study had gone through years of formal L2 instruction, where students are usually asked to memorize lists of irregular verb forms, which do not take natural input frequency into account. Thus, from the outset of formal instruction, classroom L2 learners are exposed to high-frequency as well as low-frequency irregular verb forms (like *stood* and *spoke*, respectively) to equal degrees. This practice apparently eliminated the frequency-differences in irregular forms for the L2 subjects and caused equal reaction times for the 'high-frequency' and 'low-frequency' irregular forms used in the experiment (Beck, 1997).

Zobl (1998) proposes a model of L2 processing that, similar to the dual-mechanism model, rests on listings and computations, but develops in two stages: an early *listing stage* and a subsequently evolving *computational stage*. As the names suggest, it is proposed that in the early stages of L2 acquisition the learner goes through a stage in which forms are simply listed in the lexicon, while later on in the developmental process productive rules evolve. In other words, Zobl's theory offers an extended view by capturing the two distinct mechanisms proposed by the dual-mechanism theory in a sequential dimension for L2 processing.

In his analysis of spontaneous speech production recordings of three L1 Russian adult immigrants in (English-speaking) Canada, Zobl (1998) indeed found differences between the L2 proficiency-wise different subjects. The two higher-proficiency subjects behaved like L1 adults in that it was possible to attest frequency effects[4] for irregular past tense forms, but not for regular past tense forms. With the lower-proficiency subject, on the

other hand, no differences were found in the frequency-sensitivity of regular and irregular forms, supporting Zobl's view that L2 learners do not initially possess two distinct mechanisms, but build them up gradually. This view was further sustained by the finding that over-regularization errors, indications of the presence of a productive rule, were exclusively found in the output of the most advanced subject.

In summary, then, Beck (1997) and Zobl (1998) present supportive evidence for the theory that the two distinct mechanisms predicted by the dual-mechanism model are functional in L2 processing, distinguishing regular and irregular morphology. In addition, Zobl's (1998) findings point to developmental changes in that L2 learners possibly set off with a single listing mechanism and eventually 'build up' a second computational mechanism as they advance linguistically.

Unfortunately, the number of relevant L2 studies has so far been limited and, it seems, many more studies are required to come up with a clearer picture of L2 morphological processing. The empirical study to be reported in the remainder of this study thus aims to constitute a contribution to the above stated discussions and empirical endeavors by investigating the morphological processing of L2 English. For this purpose, the elicited past tense production task previously conducted by Ullman & Gopnik (1999) and van der Lely & Ullman (2001) for native English speakers was run with L1 Turkish learners of L2 English.

3. The present study

3.1 Aim

The main aim of the present study was to investigate the validity of the dual-mechanism model for the mental representation of the English past tense morphological system in Turkish learners of L2 English. In line with this aim, one of the research questions was whether or not the L2 subjects in this study would perform as L1 and L2 subjects in previous studies and display asymmetries in the production of regular and irregular past tense forms. If so, it was predicted that L1 as well L2 subjects would exhibit frequency-effects for irregular past tense forms, but not for regular past tense forms. This prediction was formulated on the basis of the assumption that irregular past tense forms are retrieved from memory, and are thus frequency-sensitive, while regular past tense forms are rule-produced (Pinker, 1991). Furthermore, it was predicted that L1 and L2 subjects would produce over-regularization errors for irregular verbs, but especially for low-frequency irregular verbs, since the latter were expected to have weaker or no memory traces and, thus, to trigger the application of the default past tense inflection rule (Marcus et al., 1995).

Novel regular and irregular forms constituted a suitable testing-ground for the question to what degree L2 subjects would apply the default rule to verbs with non-existing or weak memory traces. It was predicted that novel regular verbs, which carry no phonological similarity to existing verbs, would be regularized by L1 and L2 subjects on the same grounds as real low-frequency irregular forms; i.e., the subjects would apply the default rule since access to a stored form in memory would not be possible. With irregular novel

verb forms, which carry phonological similarities to existing irregular verbs, on the other hand, it was expected that regularizations would occur to a lesser extent and that irregularizations in analogy to existing irregular inflectional patterns would be observed (Xu & Pinker, 1995).

A second purpose of this study was to scrutinize the claim that less advanced and more advanced L2 learners exhibit asymmetries in the processing and representation of L2 morphological information as suggested by Zobl (1998). Thus, in reference to Zobl's (1998) findings, a further research question was whether it would be possible to sustain the view that, broadly speaking, less advanced L2 learners lack a rule-mechanism and compute both regular and irregular verb forms over an associative memory. If so, it was predicted that less advanced L2 subjects would exhibit frequency-effects for regular as well as irregular past tense forms and would very rarely, if ever, produce over-regularizations.

3.2 Subjects

The subjects for the present study consisted of 49 L1 Turkish learners of L2 English and 8 native English speakers. The L2 subject-group was comprised of 27 female and 22 male undergraduate students (mean age: ~19) at Middle East Technical University (METU) in Ankara/Turkey. The L1 speakers were five (5) female and three (3) male postgraduate students (mean age: ~24) at the University of Essex / UK and constituted the control group. All subjects participated on a voluntary basis.

Since one of the aims of the present study was to examine potential developmental differences in L2 learners, the L2 subject-group was sub-divided: 25 of the L2 subjects were randomly selected advanced learners of English while the remaining 24 L2 subjects were randomly selected low-level English proficiency students. This assignment of L2 the learners into proficiency groups was conducted on the basis of the scores they had obtained from the METU English proficiency exam, which they had taken a month prior to the present study (mean scores: 21.9 and 85.0 out of 100 for low level and advanced learners, respectively). Apart from 6 students who stated that they had a very limited knowledge of French and/or German and 2 students who defined themselves as Turkish-Kurdish bilinguals, none of the 49 L2 subjects reported to have had (natural or tutored) contact with any other second language but English.

3.3 Materials

Each subject was presented with 56 verbs, which were drawn from the stimuli developed and used by Ullman (1999), Ullman & Gopnik (1999), and van der Lely & Ullman (2001). The verbs belonged to four classes:

1. 14 irregular verbs (*sing-sang*), which take only an irregular past tense form. Similar to the procedure applied in the above-mentioned studies, doublet words, which can take both an irregular and a regular paste tense form (e.g., *dive-dove-dived*), were not included.

2. 16 regular verbs (*talk-talked*), which take only a regular past tense form. The stems of the verbs in this group were dissimilar to the stems of irregular verbs.
3. 14 novel irregular verbs, whose stems carried phonological similarities to the stems of irregular verbs and which can be inflected for the past tense regularly or irregularly (e.g., *crive-crove/crived*).
4. 12 novel regular verbs, whose stems were phonologically dissimilar to the stems of all irregular verbs (*plam-plammed*). All regular and irregular novel verbs had acceptable English spellings.

Half of the real regulars and irregulars had high past tense frequencies (e.g., *gave, robbed)* and half had low past tense frequencies (e.g., *dug, stalked)*. Van der Lely and Ullman (2001) report the frequency counts to be drawn from the 17.9 million-word British English COBUILD corpus of the University of Birmingham.[5]

The verbs were presented in the context of two written sentences, which were drawn from Ullman & Gopnik (1999) and required the subjects to orally complete the elicitation sentence using the verb used in the introductory sentence, such as "Every day I *go* to work. Just like every day, yesterday I _____ to work." The introductory and elicitation sentences for each verb had the same two-word complement or adjunct (e.g., *to work*). Furthermore, every introductory sentence began with "Every day I", and every past tense eliciting sentence was written in the completive aspect and began with "Just like every day, yesterday I" (Ullman, 1999; Ullman & Gopnik, 1999; van der Lely & Ullman, 2001).

3.4 Procedures

Every subject was tested individually in a quiet classroom. The experimenter and the subject, who each had a printed version of the test-material, sat facing each other and the experimenter read out a standard instruction, which was also printed on the cover of the test-booklet.

For the coding of the responses, a procedure similar to Ullman & Gopnik (1999) and van der Lely & Ullman (2001) was employed. Accordingly, responses were coded as one of the following response types.

1. *Unmarked:* unmarked verb forms like *go-go, talk-talk* etc.
2. *Past marked:* Responses were coded as *past marked* when the expected past tense form was provided as an answer. For real words, this meant the provision of a correct past tense inflection for real regular and irregular forms (*want-wanted*, *sing-sang*). Novel regular verb forms were coded as *past marked* if they were inflected regularly (*plam-plammed*). Novel irregular verb forms, on the other hand, were coded as *past marked* if they were regularized (*crive-crived*) or if they were irregularized by the application of a vowel change (*crive-crove*).
3. *Irregularization Errors:* Responses were coded as *irregularization errors* if real or novel regular verbs were inflected irregularly (i.e., *over-irregularizations* like *squeeze-squoze, spuff-spaff*), and if real irregular forms were incorrectly irregularized by the overapplication of an inappropriate irregular inflectional pattern (*bring-brang* in reference to *sing-sang*, *ring-rang*; Xu & Pinker, 1995).

4. *Over-regularization Errors:* The overapplications of the regular inflectional suffix "ed" to real irregular verb forms (*bring*-bringed, *dig*-digged) were coded as *over-regularization errors.*
5. *Other responses:* Responses that included the use of an irregular past participle (*sing-sung*) were coded as *other responses*.

After the completion of all experiments and the coding procedure, the coded responses were analyzed using SPSS (Statistical Package for the Social Science).

4. Results

4.1 Native control group

A summary of responses for real-word stimuli provided by the 8 native subjects (overall response rate: 100%) is presented in Table 1. As can be seen, native control subjects correctly inflected all high-frequency irregular verb forms and all low- and high-frequency regular verb forms (100%). As expected, native subjects produced a very small number of over-regularization (3.57%) and irregularization (5.36%) errors exclusively for low-frequency irregular verb forms.

Table 1. Native subjects' mean response rates (as % of items) for high and low frequency regular and irregular verbs

	Mean %	S.D.
HF Regular Verbs		
Past Marked	100.00	–
LF Regular Verbs		
Past Marked	100.00	–
HF Irregular Verbs		
Past Marked	100.00	–
LF Irregular Verbs		
Past Marked	91.07	15.15
Over-regular. Errors	3.57	6.61
Irregularization Errors	5.36	10.63

The only verb that was over-regularized by native subjects was *bend* (*bended* instead of *bent*); irregularization errors, on the other hand, were produced for the low-frequency irregular verbs *wring* (*wrang* instead of *wrung*) and *swing* (*swang* instead of *swung*) – both probably in reference to the high-frequency irregular pattern present in *sing-sang, ring-rang* etc.

A correlational analysis for real irregular verbs revealed that their past tense frequencies were correlated with the rate of correctly past marked forms ($r = .49$; $p = .074$). For regular verbs, on the other hand, a correlational analysis was unnecessary since, as mentioned

above, both low and high frequency regular verbs were correctly inflected at a rate of 100%. Although the correlation obtained for irregular verb forms was not statistically significant, it clearly indicated a frequency-effect trend for irregular past tense forms as predicted by the dual-mechanism model.

Considering the nonce-stimuli, as can be seen in Table 2, native control subjects regularized an overwhelming majority of novel regular verbs (97.92%). This was an expected result since novel regular verbs carried no phonological similarity to existing verbs, and thus it was predicted that the default rule would be applied. Irregularizations, on the other hand, were produced for only 2.08% of novel regular forms (*plam-plum*, *trab-trobe*).

Table 2. Native subjects' mean response rates (as % of items) for novel regular and irregular verbs

	Mean %	**S.D.**
Novel Regular Verbs		
Past Marked	97.92	3.86
Irregularization Errors	2.08	3.86
Novel Irregular Verbs		
Unmarked	1.79	5.05
Regularized	59.82	16.61
Irregularized	38.39	15.24

For novel irregular verb forms, it was found that native controls past-marked 98.21% of the items[6] and left only 1.79% of the items unmarked (*cleed-cleed, prend-prend*). Out of the past-marked novel irregular verbs, 59.82% were produced by regularization (e.g., *shreep-shreeped*, *shrell-shrelled*, *blide-blided*) and 38.39% constituted irregularizations of various kinds (e.g., *strink-strank*, *drite-drote*, *sheel-shole*).

4.2 L2 subject group

Table 3 presents a summary of L2 subjects' responses to real-word stimuli (overall response rate: 100%).

4.2.1 *Correctly past-marked forms*

Table 3 reveals that, overall, both L2 groups produced quite high rates of correctly past marked forms for high frequency verbs, but comparatively less for low frequency verbs. For LP L2 subjects, the mean rate of correctly past marked forms was 92.19% for high frequency regular verbs and 84.90% for low frequency regular verbs. HP L2 subjects, on the other hand, produced a mean rate of 97.50% correctly past marked forms for high frequency regular verbs and 82.50% for low frequency regular verbs. Both differences were found to be statistically significant ($t(23) = 3.077$, $p = .005$ and $t(24) = 4.536$, $t = .000$, respectively), thus indicating that both L2 groups produced significantly more correct forms for high frequency than for low frequency regular verbs.

Table 3. Mean response rates (as % of items) for high and low frequency regular and irregular verbs by subject groups

	Low proficiency L2 subjects (n = 24)		High proficiency L2 subjects (n = 25)	
	%	S.D.	%	S.D.
HF Regular Verbs				
Unmarked	1.04	5.10	1.00	3.46
Past Marked	92.19	12.12	97.50	6.25
Irregularization Errors	4.69	9.62	1.00	5.00
Other Responses	2.08	4.76	0.50	2.50
LF Regular Verbs				
Unmarked	8.85	9.38	7.50	11.97
Past Marked	84.90	15.63	82.50	16.93
Irregularization Errors	6.25	14.28	10.00	12.50
HF Irregular Verbs				
Unmarked	0.60	2.92	0.57	2.86
Past Marked	92.26	10.30	96.00	6.55
Over-regular. Errors	3.57	7.59	2.86	5.83
Irregularization Errors	2.98	5.93	0.57	2.87
Other Responses	0.60	2.93	0	-
LF Irregular Verbs				
Unmarked	2.38	5.44	1.14	3.96
Past Marked	47.02	12.27	60.00	17.98
Over-regular. Errors	23.80	13.10	11.43	14.29
Irregularization Errors	25.00	12.10	26.86	13.25
Other Responses	1.79	3.96	0.57	2.86

For irregular verb forms, this observed difference between the mean rates of correctly past marked high and low frequency verbs was noticeably higher. LP L2 subjects correctly past marked 92.26% of high frequency irregulars, but only 47.02% of low frequency irregulars. Similarly, HP L2 subjects produced 96% of correctly past marked HF irregulars but only 60% of low frequency irregular verbs. Both differences were found to be significant, $t(23) = 19.000$, $p = .000$ and $t(24) = 9.498$, $p = .000$, respectively.

Correlational analyses between past tense frequencies and the production rates of correctly past marked forms revealed low and insignificant correlations for both L2 groups on regular verb forms (LP L2: $r = .27$, $p = .312$; HP L2: $r = .38$, $p = .146$). For irregular verb forms, on the other hand, statistically significant and high correlations were found between past tense frequencies and production rate of correctly past marked forms (LP L2: $r = .68$, $p = .007$; HP L2: $r = .73$, $p = .003$). Thus, for both L2 groups past tense frequency had a strong effect on the production of correct irregular past tense forms but not on the production of regular past tense forms.

In order to compare the L2 groups' and L1 group's relative performance on the four verb types, a between-groups ANOVA with the four verb types as dependent variables was

run. The results (Table 4) revealed significant differences between the three groups for low frequency verbs, but only marginally significant differences for high frequency verbs.

Table 4. Between-Groups ANOVA results for correct responses to real regular and irregular stimuli

	df	F	Sig.
Correct HF Regulars	2	3.281	.045
Correct LF Regulars	2	4.114	.022
Correct HF Irregulars	2	3.152	.051
Correct LF Irregulars	2	24.625	.000

Post-hoc, pairwise, multiple comparisons showed that the three subject groups had behaved more or less similarly on HF stimuli. Thus, no significant mean differences for HF regular verbs were found between the native group and LP L2 subjects (p = .091), the native group and HP L2 subjects (p = .771), and between the L2 groups (p = .104). Similarly, only marginally or no significant mean differences were obtained for HF irregular verbs between the native group and LP L2 subjects (p = .056), the native group and HP L2 subjects (p = .442), and the L2 groups (p = .241). For LF verbs, however, the picture looked different in that post-hoc comparisons pointed at considerable differences between the native group and the L2 groups. The native group was found to have produced significantly more correct forms than both the LP L2 group and the HP L2 group on LF regulars (p = .047 and p = .018, respectively) as well as on LF irregulars (p = .000 in both cases). The comparison of the two L2 groups revealed no significant differences in the performance on LF regular verbs (p = .846), but showed that HP L2 subjects had produced more correctly past-marked LF irregular verbs (p = .013).

In sum, the analyses of correctly past marked forms showed that L2 subjects overall produced more HF than LF correctly past marked forms for both regular and irregular verb forms. Correlational analyses, however, clearly revealed that only irregular past tense forms were frequency-sensitive, thus supporting the dual-mechanism model and suggesting that there is strong reason to assume that only irregular forms were stored in memory by L1 as well as L2 subjects. In addition, it was found that all three subject groups showed similarities in the rates of correctly past marked high frequency verbs, but displayed differences on low frequency verb forms, native subjects performing clearly better than the L2 groups on both types of low frequency verb forms. The comparison within the two L2 groups pointed to probable L2 developmental changes in the processing of irregular forms since high proficiency L2 subjects produced more correct forms than low proficiency L2 subjects for irregular, but not for regular, LF forms.

4.2.2 *Over-regularization errors*

As expected, both L2 groups produced over-regularization errors especially for low frequency irregular forms. While the mean rate of over-regularized HF irregular forms was 3.57% for LP L2 subjects and 2.86% for HP L2 subjects, over-regularizations of LF irregular verbs accounted for 23.80% and 11.43% of the L2 groups' total responses, respectively (Table 3). For both L2 groups, this observed difference between the mean rates of over-regularized

HF and LF irregular verb forms was significant (LP L2: t(23) = −7.474, p = .000; HP L2: t(24) = −2.882, p = .008). Thus, both L2 groups over-regularized significantly more low frequency irregular forms than high frequency irregular forms – a trend similar to that of the L1 subjects, who did not over-regularize any HF irregular verbs at all, but produced 3.57% of over-regularizations for LF irregular verbs.

A between-groups ANOVA comparing the three subject groups' production rates of over-regularization errors revealed that they did not differ in the rates of over-generalization errors on high frequency irregular verbs, F(2) = .970, p = .386. However, significant differences between the subject groups were found for the rate of over-regularization errors on low frequency irregular verbs, F(2) = 9.472, p= .000. Post-hoc, pairwise, multiple comparisons showed that these differences stemmed from the production rates of the LP L2 subjects, who were found to have produced significantly more over-regularizations on LF irregular forms than both HP L2 subjects (p = .004) and native subjects (p = .001). No significant differences were found between HP L2 subjects and native subjects (p = .306).

The results clearly demonstrated that both L2 subject-groups were able to fall back on the application of the default rule in cases where the retrieval of a correct irregular past tense form failed, although to different degrees. Furthermore, similar to the previous analysis of correct responses, the results also indicated L2 developmental changes in the processing of irregular forms since LP L2 subjects evidently resorted more often to the regularization of LF irregular forms than HP L2 subjects.

4.2.3 *Irregularization errors*

L2 subjects produced irregularization errors for both regular and irregular stimuli, though to varying degrees. Similar to over-regularization errors, the rate of irregularization errors was comparatively higher for low frequency verbs than for high frequency verbs (Table 3). Statistical analyses revealed that HP L2 subjects produced significantly more irregularization errors with LF verbs for both regular (t = −3.674, p =.001) and irregular (t = −9.751, t = .000) past tense forms. For LP L2 subjects, on the other hand, the mean difference was significant for irregular verb forms (t = −8.547, t = .000) but not for regular past tense forms (t = −.681, p = .503).

Overall, L2 subjects produced 44 different irregularization-error forms (comprising 148 tokens and 19 different verbs). These consisted of 32 errors involving an incorrect vowel change (119 tokens) and 12 errors involving incorrect *-t* suffixation (29 tokens). 75 out of the 148 error-tokens were produced by LP L2 subjects (12 HF regular, 12 LF regular, 5 HF irregular, 46 LF irregular) and the remaining 73 tokens were produced by HP L2 subjects (3 HF regular, 20 LF regular, 1 HF irregular, 49 LF irregular). A qualitative analysis of the irregularization errors revealed that 36 out of the 44 error types (81.8%) were direct analogies of existing irregular pairs, while only 8 error types reflected distortions of existing patterns.

Thus, in sum, L2 subjects produced a number of irregularization errors for both regular and irregular stimuli, the great majority of which were built upon existing irregular patterns. Although L2 subjects also produced some rather 'weird' forms that carried gross distortions, these were very few in number. Furthermore, L1 subjects and L2 subjects showed major differences only in the irregularization rates of LF irregular verb forms.

4.2.4 *Novel regular and irregular verbs*

A summary of the responses produced for novel-word stimuli by L2 subjects is presented in Table 5.

Table 5. Mean response rates (as % of items) for regular and irregular novel stimuli by L2 subject groups

	LP L2 subjects (n = 24)		**HP L2 subjects (n = 25)**	
	%	S.D.	%	S.D.
Novel Regular Verbs				
Unmarked	4.86	5.45	5.33	5.38
Past Marked	84.38	17.77	89.00	10.68
Irregularization Errors	10.76	15.24	5.67	9.23
Novel Irregular Verbs				
"Unmarked	0.89	3.20	1.43	3.57
Regularized	48.21	23.19	40.57	16.33
Irregularized	50.89	22.62	58.00	15.47

For novel regular forms, both LP and HP L2 groups produced significantly higher rates of regularized forms (84.38% and 89%, respectively) than irregularized forms (10.76% and 5.67%, respectively). For both groups, these differences were significant at the $p = .001$ level.

For novel irregular forms, on the other hand, the comparison of past tense production rates of regularized and irregularized responses within each L2 group showed that HP L2 subjects had produced significantly more irregularized than regularized forms for novel irregulars, $t(24) = -2.756$, $p = .011$. For LP L2 subjects, no significant mean difference was found between the rates of regularization and irregularization, $t(23) = -.287$, $p = .777$.

5. Discussion and conclusion

The obtained results can, by and large, be accommodated in the theoretical framework of the dual-mechanism model and leave little space for alternative accounts like single-mechanism approaches. As has been shown, L2 subjects overall behaved very similar to L1 subjects and displayed the quantitative and qualitative production patterns predicted by the dual-mechanism model.

As mentioned at various points throughout this study, one of the major theoretical tenets of the dual-mechanism model is the view that one of the two basic mechanisms in language processing is a default rule-system that computes 'on-line' and attaches the default suffix *-/ed/* to regular stems. One of the research questions at the outset of this study was whether it would be possible to find evidence for the presence of a default rule in L2 subjects, and especially for LP L2 subjects since Zobl (1998) presented empirical evidence for the lack of a rule-mechanism in low proficiency L2 subjects. The results obtained present clear evidence from various perspectives for the view that the rule-system was indeed intact in low proficiency as well as high proficiency L2 learners.

First of all, the results clearly reflected that the production rates of correctly past marked regular forms were independent of the verbs' past tense frequencies for both low proficiency and high proficiency L2 subjects. The correlational analyses conducted between the past tense frequencies and the rates of correctly past marked regular forms revealed no significant correlations between the two variables for either L2 group, constituting the first evidence for the view that regular past marked forms were computed on-line by means of a rule and were probably not retrieved from memory by either low proficiency or high proficiency L2 subjects.

That the rule-mechanism in L2 subjects was intact was also clearly reflected in the high rates of correct responses they provided for real and novel regular stimuli. Both L2 groups correctly past marked an overwhelming majority of both high frequency and low frequency real regular verbs and novel regular verbs. That these were instances of a rule-application procedure and that the subjects had not simply memorized the inflected forms was clear from the fact that novel regular verbs, which were completely unknown to the subjects and furthermore carried no phonological similarities to existing words, were regularly inflected to high degrees (84.38%–89.00%) as expected.

Further evidence for the fact that L2 subjects were able to make use of the default suffix and did not implement all inflectional processes via the lexical memory came from the over-regularization errors they produced. According to the dual-mechanism view, it is expected that the default rule be applied in instances where the retrieval of an irregular past tense form is not possible because no entry exists or because memory traces are very weak (Marcus et al., 1992; Pinker, 1991, 1999). On this view, the lack of over-regularization errors would be an indication of a dysfunctional or even completely lacking rule-system. Such a failure to over-regularize was observed by Ullman & Gopnik (1999) and van der Lely & Ullman (2001) for SLI subjects, which was interpreted as the manifestation of a dysfunctional *-ed* suffixation rule in both instances. The L2 subjects in the present study, however, over-regularized 2.86%–3.57% of high frequency irregular stimuli and 11.43%–23.80% of low frequency irregular stimuli, clearly showing that they were able to over-regularize, i.e., to apply the default rule.

Similarly, the L2 subjects also regularized an important proportion of novel irregular verbs (mean rates: 40.57%–48.21%), whose stems carried phonological similarities to the stems of existing irregular verbs. A dysfunction in the rule-system would have manifested itself in very low, if any, regularization rates for novel irregular verbs, which was clearly not the case here. For comparison, the SLI-children in van der Lely & Ullman (2001), who were found to have a disability in the rule-system, regularised only 7.1% of the novel irregular stimuli. Furthermore, they were also found to be regularizing novel irregular forms significantly less than normally developing control children. Such significant differences in regularization rates of novel irregular forms between the L2 subjects and the native subjects in the present study were not found, either.

In contrast to regular verb forms, the subjects' productions of irregular verb forms were evidently influenced by lexical effects, underscoring the dual-mechanism view of distinct mechanisms serving for regular and irregular forms. First of all, it was found that the production of irregular past tense forms was significantly affected by their past tense frequencies, which was captured in the positive and significant correlations ($r = .68$ and

r = .73, both significant at p = .01) between the rates of correctly past marked irregular forms and their past tense frequencies. In addition, the L2 subjects also produced high rates of irregularizations for irregular rhyming novel verbs (50.89%–58%), showing that the phonological characteristics of the novel verbs also determined their irregular past tense use. These frequency and phonological effects for irregular verb forms were very much in line with the view that irregular forms are retrieved from a pattern associator memory.

Similar to over-regularization errors, the rates of irregularizations turned out to be rather high when compared to the irregularization rate of 0.19% reported by Xu & Pinker (1995) for the spontaneous speech of L1 children. This might partly have been due to the fact that irregularizations are more frequent in elicitation experiments than in spontaneous speech (Xu & Pinker, 1995). Another possible interpretation, however, is that L2 learners show a relatively higher reliance on the associative memory when compared to native speakers. It has been attested in studies on L1 subjects that the pattern associator memory yields some productivity, resulting in the production of irregular past tense novel forms and occasional irregularization errors (Prasada & Pinker, 1993; Xu & Pinker, 1995). It may well be possible that this attested productivity is less limited in L2 subjects and that L2 learners use it as an alternative to fall back on in addition to the default rule, which is certainly not lacking or impaired as the high rates of regularizations and over-regularizations have shown. This interpretation would constitute a weak version of Ullman (2001b) and Zobl (1998), who claim that in adult L2 learners, or in less proficient adult L2 learners, the ability to implement grammatical rule computations may be impaired and may result in a complete reliance on the associative memory.

It is of course possible to speculate on the reason(s) of this hypothesized higher productivity of the associative memory in L2 learners in various ways. One possibility is to tie this observed phenomenon to the common practice of making L2 learners in classroom settings memorize irregular pairs from lists, as has also been underlined by Beck (1997). In a sense, this practice of memorizing irregular pairs pushes L2 learners to focus more on the exceptions than on the rule, since what is focused on and tested in school environments is more often the irregular past tense forms than the regular forms. Thus, this practice may cause L2 learners to considerably strengthen the links between pairs of memorized irregular forms, which may boost the 'gravity' of these pairs and lead to an increased reliance on such associations. It is also well known from research on the effect of input-frequency modifications that the frequent repetition of items in the input may easily strengthen their representations. Considering that L2 learners, especially less advanced L2 learners in classroom settings, have a rather limited vocabulary size, it would be only natural that these well-memorized and frequently focused-on irregular pairs have a relatively higher impact on them than they have on L1 speakers, who do not boost such irregular connections artificially by reinforced memorization. As speculative as it may be, such an explanation also bears certain parallelisms to the arguments of Ullman (2001b), who claims that in later exposed L2 learners the tendency to rely on the productivity of the associative lexical memory may be quite high.

In conclusion, the present study has arrived at results supporting the view that regular and irregular verb forms are stored and computed distinctively by L2 learners. To be

more specific, the results indicated that in L2 subjects regular past tense forms were computed *on-line* by the combination of the verb stem with the default suffix /ed/ and were not stored as whole words in the memory, as was captured in the lack of frequency and phonological effects for regular forms and in the over-regularization errors produced. Irregular past tense forms, on the other hand, appeared to be stored in and retrieved from the associative lexical memory as whole words since they clearly exhibited both frequency and phonological effects. In this sense, the findings constitute counter-evidence for the views of Zobl (1998) and Ullman (2001b), who claim that in L2 speakers, or less-proficient L2 speakers, the rule-mechanism may be impaired and all (or most) regular and irregular computations are conducted over the associative lexical memory.

However, the L2 subjects seemed to be more prone to producing irregularizations and irregularization errors when compared to results obtained with native child and adult subjects in previous studies. Although the dual-mechanism model does not necessarily posit a dominant rule for the rule-based mechanism but mainly hypothesizes it as a back-up procedure (Marcus et al., 1995; Say, 2000), the rates of produced irregularizations still appeared to be highly marked, especially in comparison to L1 speakers' rates of irregularizations. This phenomenon was tied to the possibility that the lexical memory in L2 learners may be more productive since irregular forms cover a larger lexical space in the lexical memory of L2 learners when compared to L1 speakers due to certain practices related to the learning of these forms; e.g., frequent focus on and reinforcement of the memorization and retrieval of irregular forms. Nevertheless, this speculation is clearly premature and in need of further support.

In sum, then, this study showed that the dual-mechanism model was overall successful in predicting the past tense production-patterns of the L2 subjects in this study. Developmental changes of the kind found by Zobl (1998) where less proficient L2 subjects lack a rule-mechanism and compute regular as well as irregular forms over the associative memory, however, were not encountered.

In addition, what the results of the present study definitely show is, above all, that more empirical studies are needed to further illuminate the storage and computations of linguistic information in L2 speakers. Considering the shortcomings of elicitations tasks, it appears absolutely necessary to dwell on spontaneous speech productions of L2 speakers to promote a better understanding of the psycholinguistic processes in L2. Particularly the hypothesis that the associative memory in L2 speakers is less limited when compared to L1 speakers and accounts for a large amount of productions *in addition to the rule-mechanism* certainly needs further verification from different psycholinguistic angles and with a focus on a greater variety of linguistic forms and languages.

Furthermore, it may be fruitful to target future investigations also at L2 learners in natural environments, since, as is well-known, pedagogical settings and practices may certainly cause some artificial by-products like the over-exposure to certain linguistic forms as has been observed in the present study. Thus, considering the fact that the L2 subjects in the present study had been/were receiving intensive training on academic English, exhaustively focusing on the grammatical features of English, it may well be possible that a similar study conducted with a group of L2 learners in a natural L2 setting may yield partly different results.

Notes

* The study reported here was supported through a Jean Monnet Scholarship. The author wishes to thank Prof. Dr. Harald Clahsen (University of Essex) for his academic support, Prof. Dr. Hüsnü Enginarlar (Middle East Technical University) for permission to involve students from the Department of Basic English as subjects in this study and the audience at LAUD 2008 for helpful comments and suggestions.

1. Marcus et al. (1992) report this rate of over-regularization as 4.2% of opportunities.

2. See, however, Ellis and Schmidt (1998) for an interesting, opposing view, which the authors build upon the findings they obtain from a connectionist system and in which the need for two separate processing systems is firmly denied.

3. See Clahsen (1995) for similar results in L2 German.

4. Zobl made use of a measure he labels "text-internal frequency". See Zobl (1998) for details.

5. See van der Lely & Ullman (2001) and Ullman & Gopnik (1999) for statistical details concerning the selection of high and low frequency verbs.

6. It should be remembered that both regularizations and irregularizations that included a vowel-change were counted as *past-marked*.

References

Alegre, Maria & Peter Gordon (1999). Rule-based versus associative processes in derivational morphology. *Brain and Language, 68*, 347–354.

Beck, Marie-Luise (1997). Regular verbs, past tense and frequency: Tracking down a potential source of NS/NNS competence differences. *Second Language Research, 13*(2), 93–115.

Beck, Marie-Luise (1998). L2 acquisition and obligatory head movement: English-speaking learners of German and the local impairment hypothesis. *Studies in Second Language Acquisition, 20*(3), 311–348.

Berko, Jean (1958). The child's learning of morphology. *Word*, *14*, 150–177.

Clahsen, Harald (1995). German plurals in adult second language development. In Lynn Eubank, Larry Selinker & Michael Sharwood Smith (Eds.), *The current state of interlanguage: Studies in honor of William E. Rutherford* (123–137). Amsterdam: John Benjamins Publishing.

Clahsen, Harald & Mayella Almazan (1998). Syntax and morphology in Williams syndrome. *Cognition*, *68*, 167–198.

Ellis, Nick C. & Richard Schmidt (1998.) Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language and Cognitive Processes, 13*, 307–336.

Johnson, Jacqueline S. & Elissa L. Newport (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60–99.

Marcus, Gary F., Ursula Brinkmann, Harald Clahsen, Richard Wiese, & Steven Pinker (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, *29*, 189–256.

Marcus, Gary F., Steven Pinker, Michael Ullman, Michelle Hollander, T. John Rosen, & Fei Xu (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, *57*(4), Serial No. 228.

Pinker, Steven (1991). Rules of language. *Science*, *253*, 530–535.

Pinker, Steven (1999). *Words and rules: The ingredients of language*. New York: Basic Books.

Pinker, Steven & Alan Prince (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28*, 73–193.

Pinker, Steven & Alan Prince (1991). Regular and irregular morphology and the psychological status of rules of grammar. In Laurel A. Sutton, Christopher Johnson, & Ruth Shields (Eds.), *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society* (230–251). Berkeley, CA: Berkeley Linguistics Society.

Pinker, Steven & Michael T. Ullman (2002). The past tense debate: The past and future of the past tense. *Trends in Cognitive Science*, *6*(11), 456–463.

Plunkett, Kim & Virginia Marchman (1991). U-shaped learning and frequency effects in a multilayered perception: Implications for child language acquisition. *Cognition*, *38*, 43–102.

Plunkett, Kim & Virginia Marchman (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, *48*, 21–69.

Plunkett, Kim & Virginia Marchman (1996). Learning from a connectionist model of the acquisition of the English past tense. *Cognition*, *61*, 299–308.

Prasada, Sandeep & Steven Pinker (1993). Similarity-based and rule-based generalizations in inflectional morphology. *Language and Cognitive Processes*, *8*, 1–56.

Rumelhart, David E. & James L. McClelland (1986). On learning the past tenses of English verbs. In James L. McClelland, David E. Rumelhart, & PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. 2, 216–271). Cambridge, MA: Bradford/MIT Press.

Say, Tessa (2000). *The mental representation of Italian morphology: Evidence for the Dual-Mechanism Model.* Unpublished Ph.D. dissertation, Colchester: University of Essex.

Ullman, Michael T. (1999). Acceptability ratings of regular and irregular past-tense forms: Evidence for a dual-system model of language from word frequency and phonological neighbourhood effects. *Language and Cognitive Processes, 14*(1), 47–67.

Ullman, Michael T. (2001a). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research, 30*(1), 37–69.

Ullman, Michael T. (2001b). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism, 4*(2), 105–122.

Ullman, Michael T., Suzanne Corkin, Steven Pinker, Marie Coppola, Joseph Locascio, & John Growdon (1993). Neural modularity in language: Evidence from Alzheimer's and Parkinson's diseases. *Society for Neuroscience Abstracts*, *19*, 1806.

Ullman, Michael T., Suzanne Corkin, Steven Pinker, Marie Coppola, Joseph Locascio, & John Growdon (1994). The neural structures subserving language: Evidence from inflectional morphology. Paper presented at the *1st Annual Meeting of the Cognitive Neuroscience Society*, San Francisco.

Ullman, Michael T. & Myrna Gopnik (1999). Inflectional morphology in a family with inherited specific language impairment. *Applied Psycholinguistics*, *20*, 51–117.

van der Lely, Heather K. J. & Michael T. Ullman (2001). Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes, 16*(2/3), 177–217.

Xu, Fei & Steven Pinker (1995). Weird past tense forms. *Journal of Child Language*, *22*, 531–556.

Zobl, Helmut (1998). Representational changes: From listed representations to independent representations of verbal affixes. In Marie-Luise Beck (Ed.), *Morphology and its interfaces in second language knowledge* (339–372). Amsterdam: John Benjamins Publishing.

CHAPTER 5

Third language acquisition, macrocategories and synonymy

Ksenya Filatova
Ural State University A. M. Gorki

1. Introduction

Saying that the majority of the world population is multilingual is a trite way, indeed, to begin an article on language acquisition. However, in some works there still remains a contradictory assumption that cerebral circuits are normally meant to be equipped with a single language system (Dehaene, 1999: 2207). Thus, language acquisition is conceived as a process of adaptation to multiple languages, a process made possible due to deep functional alterations in the brain. Which seems a paradox, once we come to think of it, if at least one half of the humankind speaks two languages on a daily basis (Dictionnaire des Sciences Cognitives, 2002: 51).

The second paradox, of much humbler scale, concerns the state of the art in language acquisition studies. During the first decades of the research development, the attention was mostly focused on the second language (L2) acquisition – precisely on getting used to the idea that multilingualism is "not reducible to multiple monolingualism" (Herdina & Jessner, 2002: 19), and it was apparently taken for granted that the third and all the consecutive languages (L3 and L(2+n)) acquisition was following the same pattern.

The problem of L(2+n) acquisition and functioning has recently moved to the fore of research, with new journal titles (e.g. International Journal of Multilingualism) or special volumes dedicated to the topic (e.g. Acquisition et Interaction en Langue Étrangère, № 24 – 2006). Multilinguals – those who acquire, process and use two or more languages – have been proclaimed "unique learners and speakers" (De Angelis, 2005) rather than bilinguals who acquire additional languages; multilingualism was described as leading to the development of proficiencies not to be found in monolingual speakers (Herdina & Jessner, 2002: 17). It suddenly became obvious that the dynamic processes underlying multilingual organization differ substantially from the ones involving only one foreign language.

This paper applies a cognitive linguistics framework to theories about multilingual mental lexicons. I am convinced that any valid linguistic theory should be able to account for a continuum of multilingual functioning. Beginning with a case study of Russian

students learning their L(2+n) Spanish, I present my hypothesis about a multilinguistic continuum, which is based on simple reasoning: in principle, any language may potentially describe one and the same referent. A functional unity grounded in referential logic allows us to consider languages as ***highly complex systems of synonyms.*** Mechanisms of segregation and co-ordination that help avoid the cross-talk and condition the choice of the language are largely the same as the ones that orient the stylistic variation between registers of speech; that is, they are primarily social.

2. L3 acquisition: Lexicon specificity

Due to the explicit nature of the phenomenon, much work on multilingualism has always been centered on the lexicon, this "central module of a natural language processing" (Handke, 1995: 50), its structure, and functioning (de Bot, 2004: 17). Among the problems which become particularly evident for any teacher of 'additional foreign language' one should name, first of all, lexical interference, or cross-talk: non-target units in the linguistic performance that occur because the needed words and constructions are being systematically borrowed from other languages – native or foreign – known to the student.

There is no place for guessing; in L2 acquisition, the one and only source of transfer – both positive and negative – is the student's L1. However, it has been extensively shown that in L(2+n) acquisition the L1 intrusions are drastically reduced, and the major influence on the nascent inter-language comes from the L2 (see Hammamberg, 2001: 22–23 for review).

The results I present in my case study are consistent with the already described tendency: in L(2+n) production the accidental intrusion of L2 elements is recurrent, whereas L1 is practically deactivated.

2.1 Method and participants

My observations are based on a year of L(2+n) Spanish for Russian-speaking university students. The group of 14 students – 13 female, 1 male – were doing their second year of International Relations studies at Ural State University in Russia, 2006–2007. The students received Spanish instruction for ten hours per week during the first academic year, and eight hours during the second year. All but one of the students were advanced in L2 English. In addition, for the majority there was the additional L(2+n) (French and/or German). The conversations that we had in class contained a lot of examples of language mixing; students used English words rather often in their Spanish speech, which provoked my inevitable reflection. Unfortunately, these conversations were not recorded. Material for the study was taken from the online forum that the students engaged in during the whole year.

2.2 Results

A total of 87 messages were posted by the students, generally in rather good and understandable Spanish. However, there were many cases of interference, for example:

- *la arquitectura impresa la gente (*< English *to impress*, instead of *impresionar)*
- *ciudad nativa (*< English *native,* instead of *natal)*
- **factos de historia (*< English *fact,* instead of *hecho)*
- *edificios son familiares (*< English *familiar,* instead of *conocido)*
- **tremendoso (*< English *tremendous,* instead of *grandioso, enorme).*

2.3 Discussion

The choice in favor of an L2 non-target unit is especially interesting when there is more similarity between the L1 form and the target L3 one; for example, Spanish *arquitectura* sounds much more like Russian *архитектура* [arhitektúra], than English architecture. However, the student used the form **architecture*, reflecting, to make the matters worse, the English spelling. Intuitively any language teacher would agree to say that a student, however proficient in English he is, would hardly ever incorporate the word **architecture* in his Spanish discourse using the perfectly English-like pronunciation. He would indeed make his best to *sound* Spanish-like in order to make a phonological adaptation to the target language. In cases of such non-target units as *nativa* or **tremendoso* one also witnesses a morphological adaptation. That very simple observation speaks for the intrusion regularity: the forms that are likely to be mixed up are lexical. Vocabulary is our melting pot par excellence. These findings speak in favor of the dual-mechanism model (Pinker, Ullman) described in Kirkici (this volume). This model is based on the theoretical assumption that human language is characterized by distinct mental representations for lexicon – "mental lexicon" and grammar – "mental grammar".

Thus, the cases of interference can be divided into two categories. First, ***direct borrowings***, when an authentic L1 or L2 form is used instead of the target one – what B. Hammarberg calls *non-adapted language switches*, "expressions in languages other than L3 that were not phonologically or morphologically adapted to L3" (Hammarberg, 2001: 25). Second, ***influences***, when the target form is distorted under the influence of the implicitly present form in the learner's other languages. There must be a reason for such an error segregation typology, and I return to this issue below.

3. Languages in interaction: Cluster of theories

There is a whole cluster of theories that attempt to explain the phenomenon of multi-competence, each of them possessing a certain appeal. I address three theories that correlate with each other in certain respects, and through this correlation contribute to our explanation of a lexical continuum.

1. *Default supplier theory*

Björn Hammarberg concentrates his attention on the language that eventually becomes the source where non-target interfering lexical units originate. This language, called the default supplier, is activated either consciously or nonconsciously whenever a speaker needs lexical compensation. Four criteria account for the choice of a language for this role, namely,

- typological similarity between the potential default supplier and the language it compensates for;
- proficiency of the speaker in this language;
- recency of activation; and finally,
- its status, whether the language in question is native or foreign (Hammarberg, 2001: 36).

Williams and Hammarberg (1998) and Hammarberg (2001) propose a number of plausible explanations for the fact that L2 is much more likely to be drawn upon in production. First, they speak of a different acquisition mechanism that is used for learning foreign languages in contrast to learning L1. This mechanism becomes reactivated during L3 acquisition. Second, they advocate a "desire to suppress the L1 as 'non-foreign' and instead orientate oneself towards a prior foreign language when approaching the target language" (Hammarberg, 2001: 36–37).

The first explanation seems perfectly obvious: indeed, if L1 emerges to give the world its name, L2 comes into an *already worded world*. That is, it serves as a *system of synonyms*. So does L3. The second explanation seems rather psychological: Williams and Hammarberg (1998) state that a certain metalinguistic caution towards reliance on the L1 is systematically perceived as a better acquisition strategy. Speakers may have 'a desire to suppress L1 in the belief that this is inherently "non-foreign" and thus that using a non-L1 and hence "foreign" language would be a better strategy in acquiring another language' (Williams & Hammarberg, 1998: 323). For example, a Russian-speaking student of Spanish subconsciously presumes that recurring to his native language resources is not advisable when he is supposed to produce a phrase in a 'foreign language', therefore, when lacking proper Spanish language units, he inserts English ones. This point becomes even more pronounced in the writings of Gessica De Angelis (2005).

2. *System shift theory*

A particularly interesting approach to the problem of L(2+n) lexis acquisition is presented in the works of Gessica De Angelis (2005), who postulates the existence of a separate cognitive process involved in lexical transfer from a "source to a guest system." She calls this transfer a *system shift*. Due to this process, De Angelis claims that multilinguals who produce non-target lexical items in L(1+n) using the patterns of L(2+n) may "come to identify the lexical item transferred from a source to a guest system as belonging to the guest system and fail to recognise the source of their knowledge in the original linguistic system" (De Angelis, 2005).

The system shift, according to De Angelis, is largely determined by two factors, namely ***perception of correctness*** and ***association of foreignness***. These two are closely interwoven:

the speaker feels that what is 'native' for him cannot be 'correct' when he speaks 'foreign', thus, the association of foreignness parallels the L2 status notion of Williams and Hammarberg.

3. *Language mode theory*

Another explanation to the intriguing fact of L2/L3 interaction stems from François Grosjean's language mode theory. Starting with anecdotal stories of introspection in bilinguals – whether they "feel the same" when communicating to monoglots or to persons who share both of their languages – then proceeding to the linguistic behavior description, he introduces the notion of a ***language mode*** defined as "the state of activation of the bilingual's languages and language processing mechanisms at a given point in time" (Grosjean, 1999: 3). There is always (1) a base language, which is normally defined as base for the context or topic given, and (2) the activation scale, which presents three distinct positions: monolingual mode, intermediate mode, and bilingual mode.

In order to apply this theory to the L3 processing, we simply have to assume that 'guest language' is actually a fuzzy set of all the guest languages known to the person, which gives form to a very rigid conceptual opposition "native language" versus "foreign language." This opposition might serve to explain the "talk foreign" theory: there is a subconscious cognitive strategy implying that you have more chances of coming up with a target form when you search through all your "foreign languages" then when you address your native one. There is an option for "native languages" versus "foreign languages," though for the sake of the presentation I subsume this option under the previous dichotomy.

To sum up, there is a rich theoretical background for multilingual language representations, based on both introspection and logic. There is a general consensus on the existence of some 'foreign talk.' However, this provokes the inevitable question: is there a separate or an integrated lexicon for every language known to the individual (Cenoz et al., 2003; Kroll & Dijkstra, 2002; Singleton, 2001, 2003)? Or are there subsets in the lexicon (Paradis, 1985, 1987, 1997)?

This very problem is the focus of attention for neuroscientists studying language. It is beyond the scope of this article to present a detailed meta-analysis of the relevant neuropsychological data. What follows, however, is an outline of several facts that should be taken into consideration when addressing the problem of language macrocategories in the multilingual mind.

4. Neuropsychology of bilingualism

The studies of several languages localized in one brain result in the two major questions to be answered: (1) the nature of language representation in the brain and (2) the neural basis of language switching. There is little doubt that at least some degree of separation between the linguistic systems does exist. It becomes even more clear when we consider the patterns of language recovery in bilingual and multilingual aphasic patients. Not only do particular lesions provoke more or less serious loss of one or several languages the patient

could speak, a whole number of pathologic linguistic mixtures was reported in literature. An accent from one language appears in the other as a post-traumatic symptom, anomalous language switches take place in the same sentence, as in *Je parle de ça to mes amis*, where an English preposition is inserted into a French phrase, or even in the same word, as in *consciousheit*, where the root *conscious* comes from English, and the German suffix *-heit* replaces the English *-ness*) (Paradis, 2003: 539).

Experimental studies of the bilingual brain include electrophysiological investigations (electrocorticostimulation during brain surgery and event-related potentials) and functional neuroanatomy studies (positron emission tomography, PET; functional magnetic resonance imaging, fMRI; magnetoencephalography, MEG). In the latter, healthy subjects are presented with discrete lexical or syntactical stimuli such as word completion tasks.

Most current studies on bilingualism demonstrate that both languages a person speaks access a common semantic system (Illes et al., 1999: 347) and even more specifically, that both lexicons are "macroscopically represented in the same brain areas" (Fabbro, 2001: 217), both in early and late bilinguals. At the same time, grammatical aspects demonstrate different representation for the L2 acquired after the age of seven, which leads Fabbro to conclude that bilingual lexicons are "stored" in the declarative memory systems, whereas morphological and syntactic aspects vary in their localization. That interpretation correlates with the clear indications that both overlapping and separate components are involved in cortical language representations for bilinguals (Halsband, 2006: 359). Moreover, neuronal responses in bilinguals are sensitive to changes of the words meaning and still differently sensitive to the combination of language and meaning change (Crinion et al., 2006).

The differences in cortical activation do happen, though, with the lower proficiency of the subjects. Whereas there is hardly any language-specific brain activity in highly proficient bilinguals, low proficiency in one of the languages is accompanied by "generally less activation, less participation of language relevant areas in the left temporal lobe, and more widespread contribution of other areas" (De Bleser et al., 2003: 440).

Fully aware of the fact that I am treading a dangerous ground of neuroscience, I suggest that there is a common semantic system being verbalized in different languages. Distinct grammatical and phonological systems are adjacent, but separate. A theoretical explanation accounting for this segregation is presented in the next section.

5. Synonymic fields: A cognitive linguistics response to multilingualism?

Presumably, all the languages known are potentially equal in their functions, that is, we can imagine a situation when one and the same pragmatic task is being solved with the help of them all, when one and the same referent is being described in all of them.

Bearing this assumption in mind, I propose to consider all the languages known to a person as ***systems of synonyms***. On the lexicon level these systems are bundled together through ***synonymic fields*** that replace the linguistic conceptualizations of L1. Grosso modo, the content words *mela – manzana – pomme – Apfel – яблоко* are nothing but ***cross-linguistic synonyms*** for the real-world *apple*.

A brief explanation of the term is needed here. If we consider some classical definitions of a **lexical field**, e.g., "a collection of words that all name things in the same conceptual domain" (Dirven & Verspoor, 1998: 37), we see that there is absolutely no restriction concerning the language these words belong to. I coin the term "synonymic field" to emphasize the non-linear character of the phenomenon and to distinguish it from the traditional "synonymic chain" for one language only. It comes closest to the notion of the **semantic cluster**, as in Cruse, "clusters are essentially groups of synonyms. The name is intended to indicate that the sharpness and complexity of structuring is much less than in other types of fields: they are somewhat informal groups" (2000:193). What is especially interesting about the cluster concept is that clusters are subdivided into two types, a centered and a non-centered one. A centered cluster "has a more-or-less clear core of one or two items, and a penumbra of more peripheral items" (Cruse, 2000: 193), whereas a non-centered cluster has "the items spread over a spectrum of sense, but there is no superordinate item" (Cruse, 2000: 194). Should a multilingual synonymic field be considered as a cluster, it would apparently be centered, the core item belonging to one of the languages in which the conceptualization is essentially primed.

A model of multimapping lexicon is therefore proposed, in which the famous semiotic triangle (Figure 1)

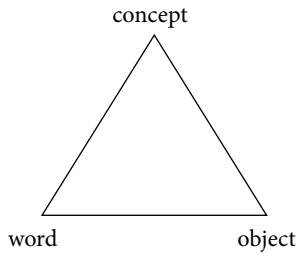


Figure 1. Semiotic triangle for monolingual speakers

is easily transformed into (Figure 2):

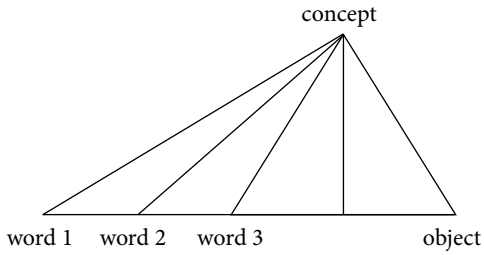


Figure 2. Semiotic triangle for multilingual speakers

One of the basic tenets of semantics is that the words in the lexicon entertain either substitutional, ***paradygmatic*** sense relations, or combinational, ***syntagmatic*** sense relations (Lyons, cited by Handke, 1995: 88). As the L1, L2 and L(2+n) words have no intrinsic

incompatibility *per se,* the mechanisms to block the combinational relations between the words belonging to different languages must be very efficient, indeed.

One possible solution to the problem of linguistic segregation is the notion of ***language tags*** (cf. de Bot & Schreuder, 1993; Poulisse & Bongaerts, 1994; Green, 1986, 1998). These markers, attached to the words, help retrieve them in speech production. For example, having a tag «French» attached to the words «ce», «être», «une», «maison» in a set «ce», «it», «être», «to be», «une», «a», «house», «maison» we construct a sentence *C'est une maison* and inhibit the analogous *It is a house*. Faced with the necessity to account for cross-talk and interlinguistic confusion, some authors put forward possible explanations of the so-called ***retagging problem*** (e.g., De Angelis, 2005): language tags are being changed over time along with the development or withering of certain language proficiency.

In contrast, I claim that the notions of both the "sense of foreignness" so aptly described in De Angelis (2005) and the "language tags" elaborated by Green (1986, 1998), predict the existence of a huge ***foreign language macrocategory***. In terms of cognitive categorization theories dating back to Eleanor Rosch (1973, 1977), there are certain categories that are considered to be more "basic" than others: that would be the case of "dog" in the series of categories "mammal" – "dog" – "terrier." These basic categories are recognized more quickly, learned earlier, and used in a wider range of frequent situations. As Handke puts its, "they are processed more easily than other categories. These *basic-level categories* are set theoretically in the middle" (Handke, 1995: 99).

In terms of *macrocategories,* which function nearly like conceptual categories but on a reduced scale, the same theoretical division might be assumed. Indeed, there is a basic-level category, quickly learned, easily accessible, often used: obviously, it is L1. From this perspective, the "basicness" of a category is directly conditioned by the number of projections from the conceptual field: the number of concepts that can currently be verbalized with the help of this language (Figure 3):

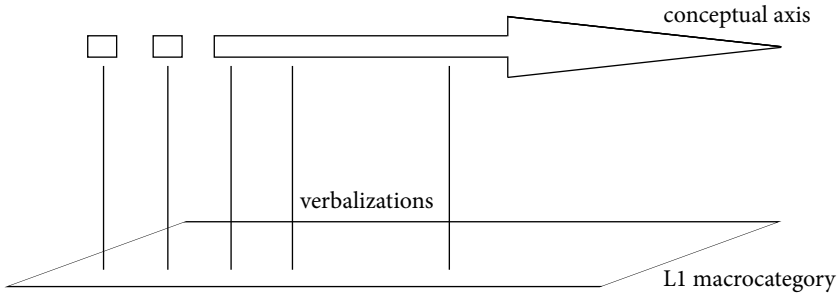


Figure 3. Wording the conceptual world in L1

When L2 acquisition starts, the second macrocategory is constituted. The projections that have already been engrammed on the L1 lexicon find their respective correlates in the L2 lexicon; moreover, due to the multidimensional nature of human categorization, they belong simultaneously to the synonymic field built around the concept (syntagmatic sense relations) and are tagged to a certain lexicon (paradygmatic relations), as shown in Figure 4:

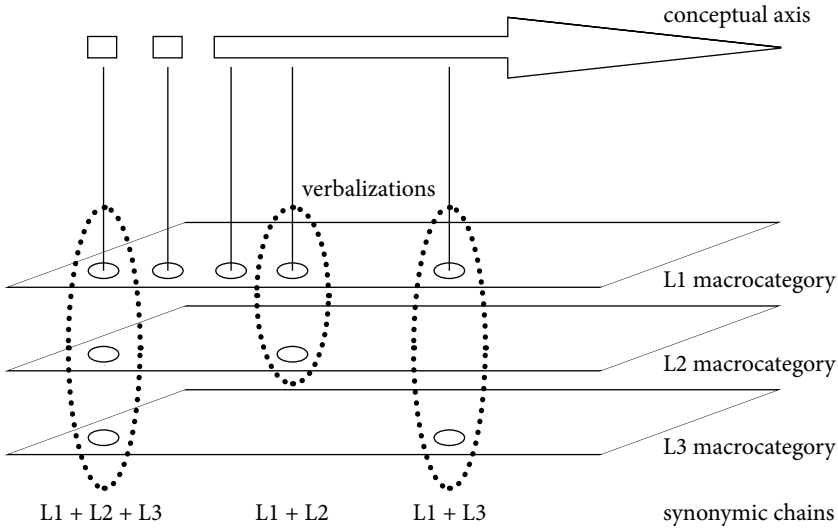


Figure 4. Wording the conceptual world in L(1+n)

The synonymic fields verbalizing the concepts are very often incomplete, as the languages known to a person may differ drastically in currently assumed functions, in social settings, therefore, in richness of lexicons. For instance, a person listening to explanations on how to cook a typical dish in a foreign country in his L3 might ignore the respective words in his L1, but still get the message and form a conceptual category where lexemes in his L1 and L2 would not be present. This fact is well accounted for in the approximative systems theory: multilinguals often work with approximative, inherently incomplete systems of their target languages (Herdina & Jessner, 2002: 13).

As with any normal-scale category, our macrocategories have their ***prototypical members*** and ***peripheral*** ones. Also similar to any normal-scale category, macrocategory boundaries are fuzzy and they quite often overlap (Dirven & Verspoor, 1998: 18). The overlapping sectors between the languages are filled with the ***internationalisms***, or otherwise, with the words that sound similar for this very linguistic combination. To put it another way, their ***language tags*** are not firmly labeled, and a word search in the linguistic field may lead to indiscriminate triggering of its foreign neighbors.

Consider the idea of lexical item placement in different categories, or "subsets" as expressed in De Bot & Lowie (this volume):

> The same lexical item can simultaneously be part of several subsets, like a register, a language variety or a language. Some of these subsets can be expected to be more stable than others. The subset of lexical items associated with a language that is not someone's native language can be expected to be less stable than lexical items associated with the mother tongue assuming that more frequent use and more frequent co-occurrence lead to more stable connections in the network.

This notion of stability seems very promising in terms of metaphorical explanation of retagging.

In addition, De Bot & Lowie's suggestion allows us to answer the most evident question my hypothesis might raise: is there a fundamental distinction between the languages known to a person? A positive answer, according to some scholars, would be a "simple" and "basically uninteresting" solution (Meara, 2006: 631). I argue that a multilingual lexicon does have many of the properties of a large monolingual lexicon. Yet this does not prevent it from exhibiting "other interesting properties" (ibid), namely, tagging, retagging and stability.

Returning to the L2 intrusion problem, I suggest that it benefits from being considered in the light of the cognitive theory of meaning (cf. Sihna, 1999). Apart from bilingual children for whom the world is being voiced simultaneously in two languages, throughout the acquisition process ***L2 is always mediated by L1***, a base language through which the learner strives to transform, to rephrase, to reformulate, an additional grounding. Originally, ***L2 is a synonymous doubling of L1, quite deliberately created as such*** (just think of *"Now, how would you say it in Spanish?"* – sort of talk). Therefore, L2 units that ***were meant to be synonyms*** from the very beginning, to pop up when the other words are asking to be said, become quite easy to evoke whenever there is a knowledge gap interfering with L3 production (deficient multilingual mapping) – even though awareness thereof might be totally absent.

Conscious activation of a language in a multilingual mind should start with the socially determined choice, that is, choice of a language to speak, setting of grammar and phonetics, a syntagmatic grid that triggers off the selection of appropriate vocabulary. Indeed, these two – grammar and phonetics – constitute a watershed between languages in the FL macrocategory because they help to tease the synonyms apart. Grammar codings and phonological articulations are much harder to mix up after a certain level of proficiency.

6. Conclusion

I have presented cognitive and neurophysiological evidence to argue for viewing multilingualism as an intricate interplay of macrocategories. Every theoretical model implies considerable simplification, and mine is no exception. The design for future experiments necessarily implies the psychological measurements – eye-tracking, reaction time changes – in multilinguals stimulated for verbal synonymic production in one versus several languages. If my prediction is correct, paradygmatic synonymic fields should be more rapidly processed when all the languages are activated at the same time. These should occur outside of the syntagmatic contexts governed by cross-talk inhibiting mechanisms.

As a finishing touch to this paper, I leave a hypothesis – first of all, for myself – that remains at the moment too bold an idea. A possible corollary of considering L(2+n) as synonymic doubling might be an attempt to ground the L3 acquisition in L2 interface, thus replacing L1 in its role of the intermediate language in the learning process. It might have positive effects on the inner restructuring of the foreign language macrocategory, though I dare not claim anything until future research resolves my doubts.

Acknowledgments

Many thanks to David Kydd and to the anonymous reviewer for helpful comments on an earlier draft. I am also grateful to Martin Pütz and Laura Sicola for their support. Finally, I would like to thank my students who constantly provoke my reflection on processes inside the learner's mind.

References

Crinion, J., Turner, R., Usui, K., Green, D., Price, C., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J., Aso, T., Urayama, S., Fukuyama, H., & Stockton, K. (2006). Language control in the bilingual brain. *Science*, *312*, 1537–1540.

Cruse, A. (2004). *Meaning in language: An introduction to semantics and pragmatics (Oxford Textbooks in Linguistics)*. New York: Oxford University Press, USA.

De Angelis, G. (2005). Multilingualism and non-native lexical transfer: An identification problem. *International Journal of Multilingualism*, *2*(1), 1–25.

De Bleser, R., Dupont, P., Postler, J., Bormans, G., Speelman, D., Mortelmans, L., & Debrock, M. (2003). The organisation of the bilingual lexicon: a PET study. *Journal of Neurolinguistics*, *16*, 439–456.

De Bot, K. (2004). The multilingual lexicon: Modelling selection and control. *International Journal of Multilingualism*, *1*(1), 17–32.

De Bot, K. & Lowie, W. (This volume). On the stability of representations in the multilingual lexicon.

Dehaene, S. (1999). Fitting two languages into one brain. *Brain*, *122*(12), 2207–08.

Dirven, R., & Verspoor, M. (2004). *Cognitive exploration of language and linguistics (Cognitive Linguistics in Practice)*. Amsterdam: John Benjamins Publishing Co.

Fabbro, F. (2001). The bilingual brain: cerebral representation of languages. *Brain and Language*, *79*(2), 211–222.

Grosjean, F. (1999). *The bilingual's language modes*. Retrieved February 7, 2009, from www.unine.ch/ltlp/pub/langmode.pdf

Halsband, U. (2006). Bilingual and multilingual language processing. *Journal of Physiology, 99,* 355–369.

Hammarberg, B. (2001). Roles of L1 and L2 in L3 production and acquisition. In Cenoz, J., Hufesein, B. & U. Jessner (Eds.), *Cross-linguistic influence in third language aquisition: Psycholinguistic perspectives* (21–41). Clevedon: Multilingual Matters.

Handke, J. (1995). *The structure of the lexicon. Human versus machine*. Berlin: Mouton de Gruyter.

Herdina, Ph. & Jessner, U. (2002). *A dynamic model of multilingualism: Perspectives of change in psycholinguistics*. Clevedon: Multilingual Matters.

Illes, J., Francis, W. S., Desmond, J. E. et al (1999). Convergent cortical representation of semantic processing in bilinguals. *Brain and Language, 70,* 347–63.

Kirkici, B. (This volume). Distinct mechanisms in the processing of English past tense morphology: A view from L2 processing.

Lowie, W., Verspoor, M., & Seton, B. (This volume). Conceptual representations in the multilingual mind: A study of advanced Dutch students of English.

Meara, P. (2006). Emergent properties of multilingual lexicons. *Applied Linguistics*, *27*(4), 620–644.

Miertsch, B., Meisel, J., & Isel, F. (2009). Non-treated languages in aphasia therapy of polyglots benefit from improvement in the treated language. *Journal of Neurolinguistics*, *22*(2), 135–150.

Paradis, M. (2003). Aphasie chez les bilingues et les multilingues. In Rondal, J. A., & Seron, X. (Eds.) *Troubles du langage. Bases théoriques, diagnostic et rééducation* (529–49). Liège: Mardaga.

Sinha, Ch. (1999). Grounding, mapping and acts of meaning. In Janssen, T., & Redeker, G. (Eds.) *Cognitive Linguistics: Foundations, Scope and Methodology* (223–55). Berlin & New York: Mouton de Gruyter. Retrieved February 7, 2009, from http://formes-symboliques.org/IMG/pdf/doc-41.pdf

Tiberghien, G. (Ed.) (2002). *Dictionnaire des sciences cognitives*. Paris: Armand Colin.

Williams, S. & Hammarberg, B. (1998). Language switches in L3 production: Implications for a polyglot speaking model. *Applied Linguistics, 19*(3), 295–333.

SECTION II

Mental processes and acquisition procedures followed by language learners

CHAPTER 6

First exposure

Converting target language input to intake*

Rebekah Rast
The American University of Paris and UMR 7023 (CNRS)

Introduction

One could argue that our power of perception goes beyond the limits of what there is to perceive. One could also argue that what there is to perceive provides the basis for our perceptual activity. These arguments could extend as well to comprehension. In second language acquisition (SLA) research, debate ensues over the extent to which what we hear and see in the form of aural input, the written word, or signs provides the basis for what we learn. One way to move forward in this debate is to investigate this perceptual activity of hearing and seeing language, that is, to examine the language input to second language learners and to observe what they *do* with the language items they hear or see. Another way of thinking about this is to consider the aspects of the target language (TL) input that learners "take in" and the means by which they do this. This paper presents a study in which *all* the input provided to our learners (French learners of Polish) from the moment of first exposure was controlled. This methodology allows for a quantitative comparison of our learners' performance in Polish on a given task with the total input they had received up to and including the task. Such a comparison sheds light on the issue of how some TL input becomes part of the learner's language system and some does not. Before we proceed with this discussion, however, we need to revisit the notions of input and intake.

Input and intake

The term *intake* in the field of second language acquisition refers to a variety of phenomena ranging from intake as a product, that which has been "integrated" into the learner's language system, to intake as a process, beginning with perception and moving towards assimilation (cf. Wong & Simard, 2000 for an overview). If we are to operationalize this notion, we must define intake as something that is measurable.

As a first step, let us take a look at how SLA researchers have understood the term *intake* in the past (see also Schmidt, 1990 for a similar discussion). Corder (1967: 165)

distinguishes between input and intake in the following quote: "The simple fact of presenting a certain linguistic form to a learner in the classroom does not necessarily qualify it for the status of input, for the reason that input is 'what goes in' not what is *available* for going in, and we may reasonably suppose that it is the learner who controls this input, or more properly his intake". In essence, in this quote Corder comes to define intake as "what goes in" and implies that input therefore is "what is *available* for going in". The work presented in this article is based partially on this understanding of input and intake; however, we extend Corder's definition of input and explore how other researchers have defined or described intake. With regard to input, in the present article the term refers to that which is in the environment that the learner can hear or see. It is the linguistic phenomena that are available for being taken in by means of aural systems (hearing) or visual systems (reading and interpreting signs and gestures). Using this definition, we can measure input under certain conditions.

Intake, on the other hand, is that which is "taken in" and as such poses more problems for the researcher than does input. First of all, what does it mean to "take in" and how do we know whether a learner has "taken in"? Hatch (1983: 81) throws a wrench into our discussion by commenting, "I am not sure why we have created two terms, input and intake, except to make it clear that learning from language that is *not* addressed to the learner and that is *not* negotiated by the learner to the appropriate level is difficult". This thinking set the stage for the focus-on-form approach in SLA in which studies seek to identify how redirecting learners' attention when they fail to notice a particular structure in the input may aid the acquisition of that structure (cf. Doughty, 2003). In other words, what can help facilitate the process of input becoming intake so that the form in question can be learned? According to VanPatten (2000: 295–296), intake is the result of input processing that occurs while the learner is focusing on form. VanPatten breaks intake down into three types: (1) "intake as incorporated data", i.e. "what goes in" following Corder; (2) "intake as process", i.e. intake as the process of successively incorporating grammatical features into the linguistic system; (3) "intake as a filtered subset of input before incorporation", i.e. intake as the product of input processing – it is stored in working memory and available for subsequent processing. Sharwood Smith (1996: 1) poses the following related questions: "[…] can focusing the learner's attention on the formal properties of language have some sort of beneficial effect after all? And, more crucially, if there *is* an observed effect, how are we to explain this?" If some sort of "beneficial effect" is observed, this would be a clear sign that some sort of intake process has occurred. The second question takes us back to where we started; how do we describe, explain or measure this effect?

Hatch (1983: 81) recognizes the need for a common understanding of this terminology:

> If we wish to keep both terms, we may say that input is what the learner hears and attempts to process. That part that learners process only partially is still input, though traces of it may remain and help in building the internal representation of the language. The part the learner actually successfully and completely processed is a subset called intake. That part, then, is the language that is *already* part of the internal representation.

Here Hatch identifies different "parts" of input that become intake. Before entering into an analysis of her proposal for intake, however, we will first take issue with her definition of input, being "what the learner hears and attempts to process". With regard to the first part of her definition, the process of "hearing" may imply some element of intake if perception is considered as part of the hearing process (the view adhered to in this paper). Concerning the second part of her definition, what a learner "attempts to process", we propose that this is already a partial definition of intake. If the learner has "attempted" to process language in the environment, the intake process is already at work. By "attempted", we imagine that Hatch is referring to an initial segmentation or parsing of the speech stream. As mentioned above, we propose that "input" refer to that which is available for the learner to hear or see, that is, the learner's TL environment, exterior to the learner. This definition has the advantage of focusing only on the linguistic environment regardless of the learner, allowing for the possibility of recording input and measuring it. Intake, on the other hand, is everything that belongs to the learner's internal processes including that which the learner attempts to process. Hatch's intake, the part of the input that "the learner actually successfully and completely processed", would be included in this definition. This discussion leads to a logical question: What does it mean for a learner to "attempt to process", or to "successfully and completely process", or more simply, to "process" language?

Psycholinguistic literature has provided much research on "language processing", which is gradually infiltrating into SLA research. Carroll (2001, 2004) for one has incorporated this rich research tradition into her development of the Autonomous Induction Theory. She begins with an examination of how we "hear words" in the L2 speech stream: "Hearing words is merely a first step in a series of processes which take the speech signal as their input and culminate in an interpretation" (2004: 227–8). She further breaks down "hearing words" into the processes of "segmentation", "word activation" and "word selection", tacitly acknowledging that "hearing words" involves processes of perception and comprehension. Much research has been conducted in an attempt to identify how this process of segmentation works in a monolingual, let alone a bilingual or a second language learner. In studies on monolinguals, one fundamental question is whether a hearer accesses and then recognizes a word, starting by means of a phonemic representation, as in the Cohort Model (cf. Marslen-Wilson, 1987) or rather, if lexical access and word recognition are indistinguishable, as in the Trace Model (cf. Forster, 1976).

Also of interest here is the actual structure of the language processing system. Autonomous models claim the existence of strong constraints on how contextual information affects the bottom-up analysis. In interactive models, on the other hand, different types of information interact with each other. Contextual and sensory information, for example, contribute to lexical processing (cf. Frauenfelder & Tyler, 1987).

One way in which SLA researchers have investigated language processing is by observing the effects of feedback and correction. Some researchers have made great strides to explain how input becomes intake by studying input processing through analysis of the restructuring that occurs or does not occur in learners' responses to feedback and correction, where the immediate input is controlled. Havraneck and Cesnick (2001) investigate the success of "corrective feedback" with regard to the type of feedback, the type of error corrected, and certain qualities specific to the individual. Carroll (2001) investigates what

makes linguistic feedback useable, with a view to contributing to a constrained theory of SLA. She distinguishes between positive feedback, which confirms that a given form, string or interpretation is possible in the language, and negative feedback, which signals that it is not possible.

Matthey (1996), published in French, differentiates between "taking in" (*prise*) and "assimilating" (*saisie*), noting that one does not necessarily imply the other. A *prise* could be, for example, a simple repetition of something said with no comprehension involved. Matthey's use of *saisie*, on the other hand, refers to a phenomenon linked to processing and storing information. Both *prise* and *saisie* are traditionally translated as "intake" in English SLA texts, but a distinction between these two processes needs to be made.

A related discussion can be found in the Interaction Research program conducted primarily in North America (cf. Mackey & Gass, 2006, special issue of *Studies in Second Language Acquisition*). This program examines the relationship between interaction and learning, where Matthey's use of *saisie* could be viewed as the equivalent of "uptake" in the interaction literature. Lyster & Ranta (1997: 49) define uptake in their model as "[...] a student's utterance that immediately follows the teacher's feedback and that constitutes a reaction in some way to the teacher's intention to draw attention to some aspect of the student's initial utterance". Generally, uptake refers to learners' responses to feedback. Some type of intake clearly must take place for uptake to occur. Interaction research treats a variety of questions related to intake, such as detection, noticing, awareness, and attention, as do other research programs (cf. Godfroid & Housen, 2008 for a recent overview of these factors in SLA research). The present article will refrain from discussing the questions of whether or not the process of input becoming intake occurs during a conscious or unconscious state (cf. Hulstijn, 2005; DeKeyser, 2003) or whether or not intake leads to learning (cf. Ellis, 2006b). Given the existing research on the latter issue, we will assume that intake of all forms can potentially lead to learning, but possibly to different degrees. Our focus here is, rather, to move towards an operationalization of intake. The following section introduces an approach to investigating intake and its relationship to input by means of collecting data from the moment of first exposure to the target language.

First exposure studies

Some researchers have concluded that the best way to study what learners "take in" (or their "intake") is to compare the input learners receive with their performance in the TL. To do this, one must *measure* the input provided to the learner. One way to measure input is to begin at the beginning, that is, control the input provided to learners at the very beginning of the acquisition process. As natural language provides methodological challenges, artificial languages have often been used (cf. Reber, 1967; Hudson Kam & Newport, 2005; Hulstijn & DeKeyser, 1997, special issue of *Studies in Second Language Acquisition*). SLA researchers, however, are not always in agreement about the application of theories based on artificial languages to explain the acquisition of natural languages (cf. Schmidt, 1994). More recently, a small number of researchers have attempted to gather data from subjects upon first exposure to a novel natural language or within the first seconds, min-

utes or hours of acquisition (cf. Dimroth et al., 2006; Rast, 2008; Zwitserlood et al., 2000; Hendriks & Prodeau, 1999).

Although generalizations cannot yet be made given the limited number of first exposure studies in existence, some interesting findings have surfaced. Dimroth et al. (2006), for instance, found that frequency (the number of occurrences of a lexical item in the input) and exposure alone had no effect on lexical acquisition in the first minutes of exposure; however, frequency and visual or gestural highlighting affected the ability of subjects to recognize words and their meaning. In other words, visual or gestural highlighting appears to help learners at the initial stage of L2 acquisition "take in" the linguistic input they receive. Rast (2008) reports that lexical transparency (as measured against the learner's L1) and the position of a lexical item in a sentence (initial, middle or final) affected learners' ability to repeat the item in a sentence repetition task. These results suggest that intake is facilitated by one's prior linguistic knowledge (cross-linguistic influence) *and* by the positioning of a lexical item in an utterance (independent of cross-linguistic influence). Hendriks and Prodeau (1999) found that their learners used not only their L1 but also their L2 (in the case of L3 acquisition) when they perceived a similarity in the L2 and L3 items and structures in both oral and written tasks, confirming that prior linguistic knowledge, not limited to the L1, aids a learner's intake. What follows provides additional information about what aids learners' intake and about how first exposure studies can contribute to this important research area of SLA.

The study

Methodology

The data collection for this study was conducted at the Université Paris 8 in Saint-Denis, France. Participants, native French speakers exposed to Polish for the first time at the onset of the study, formed two groups: (1) the "learner group" – French learners of Polish attending a Polish course; (2) the "first exposure group" – French participants for whom the only Polish input was that provided during the language task.

With regard to the learner group, we report here on data collected from 8 members of this group. These learners are all monolingual (they use only French on a daily basis), and they attended all of the classes, an obvious requirement for any input study. The Polish course ran once a week for six weeks. During the course everything the Polish instructor said was recorded, and the tapes were transcribed in their entirety. These transcriptions represent the "input" we refer to in this article. The Polish instructor used the communicative approach in the classroom, avoided metalanguage, and asked students not to consult dictionaries, grammar books, or any outside input for the duration of the project. Once the data collection period had ended, after eight hours of total input, the Polish classes continued without these constraints.

The data collected at first exposure come from two first exposure groups: one of 8 and one of 9 participants. These participants resembled our learners of Polish with regard to gender, age, language background and other sociobiographical parameters. Tasks

administered to the first exposure groups were designed to collect data at the moment of first exposure to Polish. For this reason, each participant performed only one task. Having completed the task, the participant was considered as having been "exposed" to the Polish language, and was therefore no longer eligible to perform another first exposure task.

We report on the results of two tests, namely a sentence repetition test and a sentence translation test (see Appendix). During the sentence repetition test, participants listened to 20 Polish sentences comprising between 3 and 12 words each, recorded by their Polish instructor (the 3-word sentences were eliminated in the final analysis, resulting in a total of 17 sentences or a total of 113 words). Sentences were only spoken once, and participants were asked to repeat them as best they could. During the sentence translation test, participants listened to the same set of sentences used for the sentence repetitions; however, in this case, participants were told to listen to the sentences and to write, in French, the words they thought they understood. They also heard the sentences only once.

The sentence repetition test was administered to a group of first exposure participants (referred to as Period 0) and to our group of learners after 4 hours of exposure to Polish (Period 4) and again after 8 hours (Period 8). The sentence translation test was taken by a group of first exposure participants only. We will only briefly summarize results of the sentence repetition test given that they are discussed in detail in Rast and Dommergues (2003) and Rast (2008); however, we will provide an in-depth analysis of results of the sentence translation test. A comparison of results on these two tests provides useful information about what learners "take in".

Results

The sentence repetition test

The sentence repetition test (cf. Klein, 1986) was used to determine how a learner perceives, memorizes and reproduces, in the short term, an expression in the TL. The data collected from this test allowed us to examine the relevance of certain factors to our participants' ability to correctly repeat a word, including such lexical characteristics as word length, word stress, phonemic distance (between Polish and French), lexical transparency (with respect to French) and word position. We also investigated the effect of the frequency of words in the Polish input on correct repetitions (cf. Rast & Dommergues, 2003 for details).

The data revealed certain phenomena that aid an absolute beginner's perceptual activity (a first step in the intake process) by rendering a TL item more or less salient. This can be seen in Table 1. The results show factors that affected what the participants "took in" (perceived and repeated) in certain cases but not in others.

Table 1 shows that word length (measured in number of syllables) had no significant effect on participants' ability to correctly repeat the word at all periods. In addition, no significant effect of word frequency in the input was found in the first exposure group or in the input group after 4 hours of instruction. It was only after 8 hours of instruction that we begin to observe an effect of word frequency. The remaining factors presented in Table 1 show a significant effect of the factor on correct repetitions across the three periods (word stress, phonemic distance between L1 and TL, lexical transparency between L1 and TL, and word position in the sentence). We also examined interactions between these

Table 1. The effect of certain factors on participants' ability to repeat Polish words at three different intervals

	Period 0	Period 4	Period 8
word length	–	–	–
word stress	+	+	+
phonemic distance	+	+	+
transparency	+	+	+
word position	+	+	+
word frequency	–	–	+

\+ significant effect found[1]

– no significant effect found

independent variables and found some significant relationships. In summary, our participants repeated words correctly according to the following parameters:

1. More short than long words in initial position and more long than short words in final position;
2. More phonemically familiar than unfamiliar words especially when in initial and final positions and when stressed;
3. More very transparent words than opaque or only fairly transparent words.

Saliency of input items is considered in L1 and L2 acquisition to be a phenomenon that affects whether or not the input item becomes intake (cf. Peters, 1985; Ellis, 2006a, 2006b). The sentence repetition test allowed us to examine the factors that render (or not) an item in the input salient. In other words, the results should help us to predict what factors will aid learners in transforming input into intake. Sentence repetitions may be one way to get at a learner's perceptual activity, but they admittedly require production (oral repetition). It follows that if an item is not repeated, we cannot claim that it, or some part of it, was not detected, noticed or perceived (cf. Robinson, 2003 for distinctions between these terms); however, if it is repeated, we can assume that it was perceived, (and probably noticed and detected). The sentence repetition test did not allow us to examine participants' comprehension of the test items; we can therefore say nothing about their comprehension of items based on our test results. For this reason, we pursued our investigation of what helps learners "take in" input with a sentence translation test.

The sentence translation test

The sentence translation test was used to determine how a learner comprehends, memorizes and translates, in the short term, a word in the target language. The subject group comprised 9 native French speakers with no previous knowledge of Polish. As mentioned above, participants listened to the same 20 sentences as those used for the sentence repetition test; however, in this case, participants were asked not to repeat but to write, in their L1 (French), the Polish words they thought they understood. The 17 sentences analyzed for the purposes of this study represent a set of 113 experimental words. Their translations

were judged as either "correct" or not on a word by word basis. French translations that were semantically related to the Polish test word were considered correct translations regardless of their grammatical composition. The Polish word *studiuje*, for example, means "he/she studies"; we accepted *étudier* (to study) and *étudiant* (student) as correct translations of *studiuje*. By means of global one-way ANOVAs, we examined correct translations with regard to the same factors as we did with correct repetitions; however, given that the sentence translation test was administered to first exposure subjects and not to our learners, we compare results at Period 0 only.

Word length. Figure 1 shows the distribution of correct translations of the 113 words categorized into three word-length groups: 0–1 syllable (n = 48); 2 syllables (n = 37); 3–6 syllables (n = 28). A one-way ANOVA confirmed a significant effect of word length, F (2,110) = 7.193, $p < 0.01$ on correct translations. In other words, the length of the word as measured in syllables appeared to play a role in the ability of our subjects to correctly translate a word at first exposure, 3–6 syllable words showing a higher success rate than 0–2 syllable words. This is in contrast to our findings from the sentence repetition task where word length was found to have no significant effect on correct repetitions at first exposure (Period 0) or over time (Periods 4 and 8).

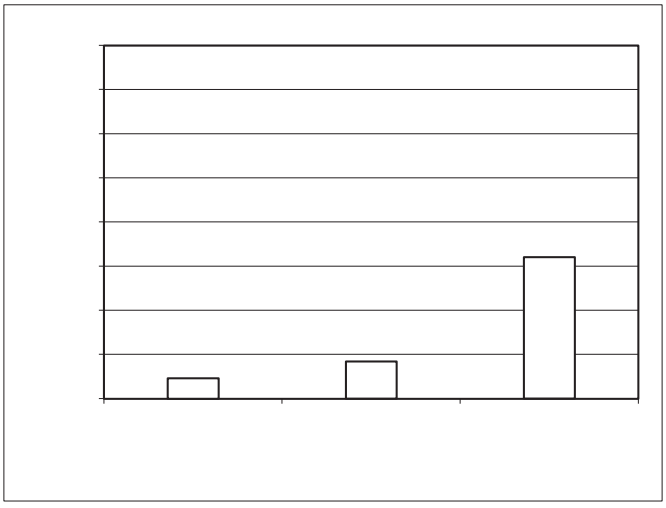


Figure 1. Correct translations relative to word length

Word stress. Correct translations were measured as a function of word stress with two stressed values, stressed (n = 58) and unstressed (n = 55). Results presented in Figure 2 show a main effect of stress on our subjects' ability to translate the Polish words, F (1,111) = 17.557, $p < 0.01$. In fact, no unstressed words were correctly translated. This corresponds to our findings in the sentence repetition test in that the difference between stressed and unstressed words was the greatest at first exposure and gradually weakened with exposure to TL input. This suggests that stress aids a learner's perception and comprehension of TL items at this initial stage of L2 acquisition.

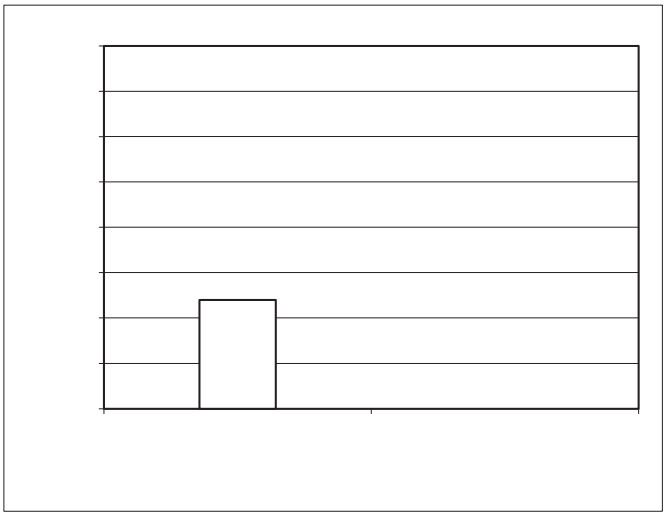


Figure 2. Correct translations relative to word stress

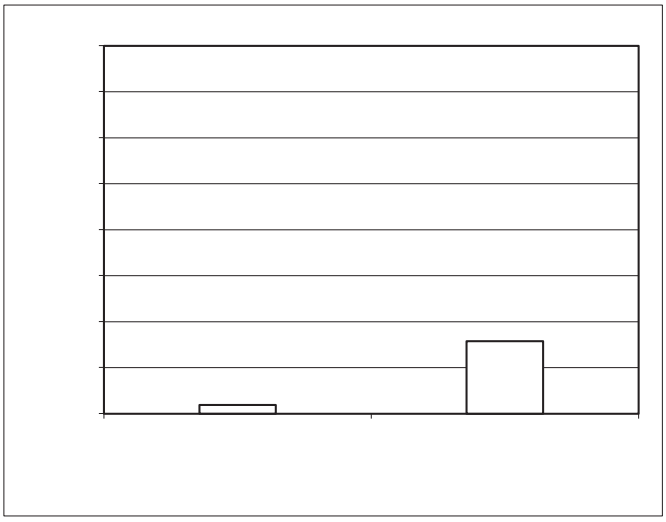


Figure 3. Correct translations relative to phonemic distance

Phonemic distance. Figure 3 shows the percentage of translations of Polish words considered as "close" (n = 89) or "distant" (n = 24) with respect to the French phonemic system. Phonemic distance was measured by means of a contrastive analysis of the French and Polish phonemic systems. A word was considered "distant" if it fulfilled one of the two following criteria: (1) contained a consonant that does not exist in the system of standard Parisian French, such as *cz* or *ci*, both approximately realized with a /tš/ in the word *nauczyciel* (teacher); (2) contained a consonant cluster that does not exist in French, such as *wł* /vw/ in *włoszką* (Italian). Words that did not meet either of the above criteria were

categorized as phonemically "close". Unlike our findings in the sentence repetition test, the results of a one-way ANOVA of sentence translations show no main effect of phonemic distance on correct translations, F (1,111) = 3.244, p = 0.0744, n.s. Whereas subjects appeared to rely on phonemic familiarity to repeat Polish words (not surprisingly), this factor provided little assistance for comprehension.

Transparency. The three values assigned to the variable "transparency" were established by means of a word translation test in which 15 French native speakers who knew no Polish were asked to translate Polish words out of context. We calculated the percentage of correct translations and then categorized the Polish words as "opaque" (rating of 0% correct translations, n = 76), "fairly transparent" (rating of 1–50%, n = 22) and "very transparent" (rating of 51–100%, n = 15). As expected, results as seen in Figure 4 show a main effect of transparency on correct translations, F (2,110) = 55.573, $p < 0.01$. A PLSD Fisher comparison confirms that the main effect was due to a significant difference between opaque and very transparent words and between fairly transparent and very transparent words ($p < 0.01$ in both cases). These results strongly suggest that transparency plays an important role in a learner's perception and comprehension of TL input at the initial stage of L2 acquisition.

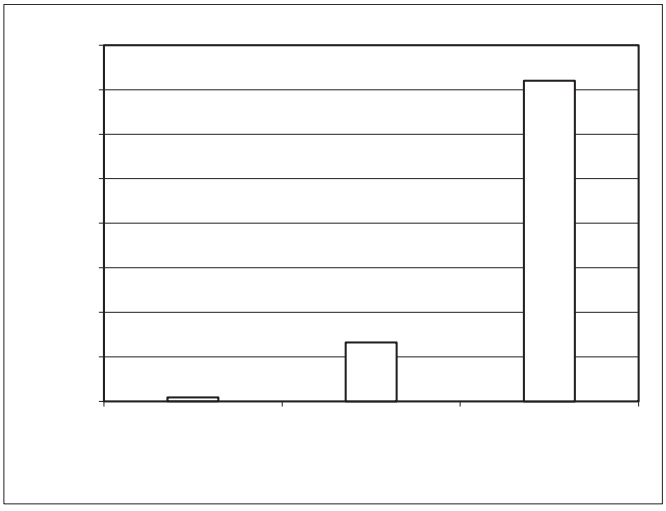


Figure 4. Correct translations relative to transparency

Word position. Three sentence positions were considered with regard to translations: sentence initial (n = 17), middle (n = 79), and final (n = 17). Figure 5 shows a main effect of position on sentence translations, F (2,110) = 17.777, $p < 0.01$. A PLSD Fisher comparison shows that this effect was due to a significant difference between initial and final and between middle and final positions ($p < 0.01$ in both cases). A further analysis revealed an interaction between position and transparency. In fact a comparison of the transparency ratings between the 17 words in initial position and the 17 words in final position showed that more words in final position than in initial position appeared in the category "very transparent". As seen in Figure 4 above, transparency had a significant effect on correct

translations and could well be the reason why words in final position were better translated than those in other positions. An interesting observation, nonetheless, is in the comparison between the sentence repetitions and translations. A summary of the results of the sentence repetitions relative to position reads as follows: "The percentage of correct repetitions was higher for words in initial position than for those in final position; however, this particular difference was not statistically significant in our data" (Rast, 2008: 154). It follows that the combination of very transparent and final position may be a more powerful predictor for comprehension/translation than for perception/repetition. This provides an excellent example of the importance of considering interacting variables in SLA research; tests need to be designed in such a way that variables can be clearly measured against each other.

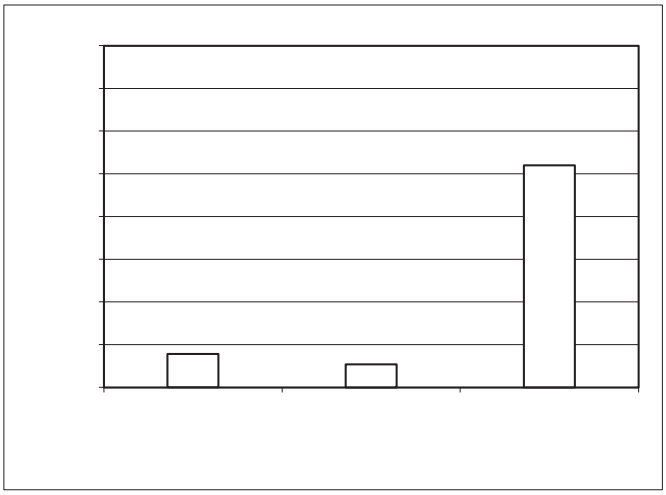


Figure 5. Correct translations relative to word position

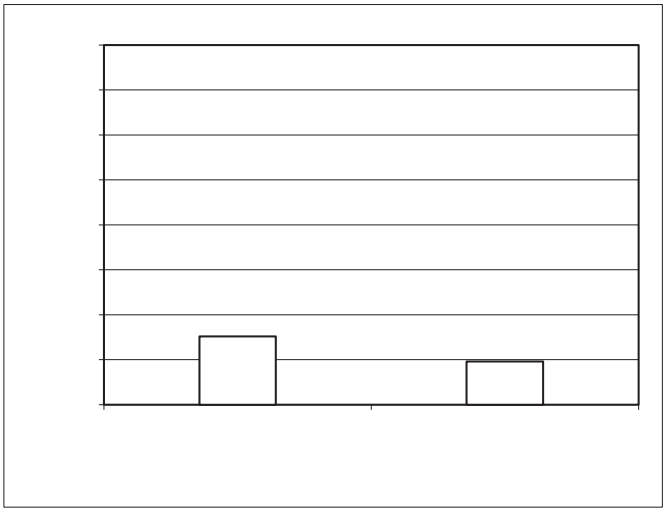


Figure 6. Correct translations relative to word frequency

Word frequency. Due to the fact that some words appeared several times in the translation test, we wanted to check the possible effect of this recurrence on our subjects' translations. We used the same criteria for frequency as discussed in Rast (2008: 149): "absent" in the input (0 tokens); "rare" (1–20 tokens); "frequent" (21–600 tokens). At Period 0, we find only two categories: absent ($n = 64$) and rare ($n = 49$), "rare" being words that appeared in previous test items, such as *i* (and). Results as seen in Figure 6 show no effect of frequency on correct translations, $F(1,111) = .801$, $p = .373$, n.s. The same was found to be true with sentence repetitions; that is, even limited tokens in the input did not help our participants repeat or translate the Polish words.

To summarize the six figures presented above, Table 2 shows the overall results of the sentence repetition and sentence translation tests at first exposure (Period 0).

Table 2. The effect of certain factors on participants' ability to repeat and translate Polish words upon first exposure to Polish (Period 0)

	Repetitions (Period 0)	**Translations (Period 0)**
word length	–	+
word stress	+	+
phonemic distance	+	–
transparency	+	+
word position	+	+
frequency	–	–

\+ significant effect found
– no significant effect found

Word length appeared to have more influence on correct translations than on correct repetitions; longer words were not repeated with more accuracy than shorter words, but they were better comprehended. Phonemic distance played an important role in sentence repetitions; TL words that contained familiar phonemes were better repeated than those that did not. In the translation exercise, however, no significant difference was found between words categorized as phonemically "close" and those categorized as "distant". These latter results may be considered less than surprising given that the translation test required no oral production; however, we might still expect to find some positive effect of phonemic familiarity on comprehension. Such was not the case in our findings here. No effect of word frequency was found at Period 0 in either data set. Word stress, transparency and position in the sentence all had a significant effect on our participants' ability to repeat and to translate the Polish words.

An additional point to make here is that we assume that perception precludes comprehension. In other words, comprehension of an input item cannot occur if perception has not occurred. This could explain why, given the same Polish sentences to repeat or translate, participants in the repetition test repeated more words than participants in the translation test translated. Of the possible 113 Polish words to be *repeated* by the 8 participants (a total of 904 possible word repetitions), 29% of these were repeated correctly.

Of the possible 113 Polish words to be *translated* by the 9 participants (a total of 1017 possible word translations), only 7% were translated correctly.

In one final analysis, we decided to take a closer look at the translations and repetitions of a selection of Polish words that appeared in the tests. Table 3 compares partial results of the sentence repetition test with partial results of the sentence translation test.

Table 3. Comparison of correct repetitions and correct translations at Period 0

Polish word **(translation)**	**% correct repetitions (n = 8)**	**% correct translations (n = 9)**
włoską (Italian)	75 (n = 6)	0
zupe (soup)	75 (n = 6)	0
piwa (beer)	87.5 (n = 7)	0
studentem (student)	75 (n = 6)	22 (n = 2)
informatykę (computer)	63 (n = 5)	67 (n = 6)
film (film)	87.5 (n = 7)	67 (n = 6)
uniwersytecie (university)	87.5 (n = 7)	100 (n = 9)

As can be seen in Table 3, the words *włoską*, *zupe* and *piwa* were repeated by the majority of subjects, but not translated. More subjects correctly repeated *studentem* than correctly translated it. *Informatykę* was repeated at approximately the same rate as it was translated. The other two words, *film* and *uniwersytecie*, were repeated *and* translated by the majority of subjects. These results show to what extent the Polish words were processed differently and/or processed to different degrees. It is the nature of this difference that interests us. In his analysis of intake, Chaudron (1985) suggests that, "In speaking of intake we are, in effect, referring not to a single event or product, but to a complex phenomenon of information processing that involves several stages [...]"(2). He evokes "preliminary intake", namely "the initial stages of perception of input" (2). Is this the type or stage of intake we have observed in our participants' processing of *włoską*, *zupe* and *piwa*? Chaudron also refers to "final intake", that is, "the series of stages by which learners fully integrate and incorporate the linguistic information in input into their developing grammars" (2). Is this the type or stage of intake we have observed in the processing of *uniwersytecie*? Given the difference in language activity that subjects perform during different tasks, we propose that the use of the term "intake" be related to the language activity involved. A theory of SLA must somehow link these different processes to different language activities.

Discussion

In order to better understand the processes described above, we need more complementary tasks using the same stimuli in which one language activity, such as perception, is tested in one task, while another language activity is tested in another. Take, for example, the sentence repetitions presented above. Participants were asked to listen to Polish sentences and repeat them. This test was designed to discover more about perception than comprehension. Modifying the task, we then asked another group of participants to

listen and *translate* the sentences. This task informed us not only about perception, but also about comprehension. Such data indeed provide useful information about the relation between perception and comprehension. To take this a step further, future studies should include collecting data by means of perception, comprehension, and production tasks while maintaining the TL lexical items that appear in each task constant. Such a methodology would provide at least partial responses to the following questions: (1) under what conditions can learners understand what they perceive; (2) under what conditions can they produce what they have at some point perceived and/or comprehended; (3) under what conditions can they produce new words and structures that have never appeared in the input?

In this article, we have used the terms "perceive" and "perception" as encompassing the processes of detecting and noticing, without commenting on the existence or level of awareness or attention. Although crucial to SLA theory, these constructs are beyond the scope of this article. Our current objective is to use our language data to examine the conversion of input into intake. Taking perception as a starting point, we may wish to say that perception constitutes one "stage" or "type" of intake. What then is needed for the next or another stage or type of intake to transpire? It is possible, for example, to perceive a word or even to store a word in long-term memory without understanding it. In their study of Dutch native speakers' first 15 minutes of exposure to Chinese, Zwitserlood et al. (2000) discovered that learners can recognize that a TL item is a word even if they do not understand it. One can also understand a TL form without being able to (re)produce it. And one can (re)produce a TL form without comprehending it, and so forth. The following is one way of contemplating the stages and types of intake, "stage" referring to a continuum as suggested by Chaudron (1985) involving preliminary and final intake during learning, and "type" referring to the language activity involved:

0. *No* intake when the item is *not* perceived, comprehended or (re)produced;
   Intake of various types:
1. The item is perceived, but it is neither comprehended nor (re)produced;
2. The item is perceived and (re)produced, but not comprehended;
3. The item is perceived and comprehended, but not (re)produced;
4. The item is perceived, comprehended, and (re)produced.

The category "(re)produced" will need to be divided into at least three subgroups: repetition, spontaneous use, and productive mastery of all form-function mappings. As mentioned earlier, the assumption here is that intake of all types can, to various degrees, lead to learning. This paper does not address the steps beyond these different types of intake; it merely sets the stage for subsequent learning that may take place.

This article has examined intake with regard to perception, comprehension and (re)production and proposes that language activity be considered in our operationalizing of intake. It follows that methodology takes on critical importance in this endeavor. What tasks will inform us as to whether or not an item has been perceived (detected or noticed and possibly segmented) or comprehended? Types 0 and 1, introduced above, could be identified by means of neuroimaging and eye-tracking studies (cf. Godfroid & Housen, 2008). Type 2 intake could be tested in simultaneous repetition/comprehension

tasks, such as the sentence repetition and translation tests used in this study. If learners are able to reproduce the TL word, we know they have perceived it. If in a subsequent task, they are unable to show comprehension of the given word, then we have a case of type 2 intake (the item is perceived and (re)produced, but not comprehended). Type 3 could be tested by means of a picture-matching task whereby the learners are asked both to repeat and select an image. If a learner selects the correct image, but is unable to repeat the word, we can assume a case of type 3 intake (the item is perceived and comprehended, but not (re)produced). Type 4 (the item is perceived, comprehended, and (re)produced) can only be tested by means of a combination of tasks (perception, comprehension, and (re)production) in which the same items appear in each task as in our sentence repetition and translation tests. Such tasks may be designed to investigate perception, comprehension and (re)production of phonetic/phonological, lexical, morpho-syntactic, semantic, and/or pragmatic information. We predict that results of such tests will confirm that a clear distinction needs to be made between "intake" in terms of the stage of the intake process and of the language activity at hand. We propose to use the term "intake *level*" or "*level* of intake", combining in essence the concepts of stage and language activity (type). In this way, we can observe intake at any given time in the learning process at the *perceptual* level, the *comprehension* level, and the *(re)production* level, the latter being "intake" that is converted for (re)production. This does not imply that intake is required for production per se, but it is required for *meaningful* production in an L2.

Conclusion

First exposure studies are unique in that they allow us to meticulously observe what learners do with the language input provided to them, in essence responding to questions about how input becomes intake. In turn, this allows us to observe what aids learners in this conversion from input to intake with regard to the various language activities of perception, comprehension, and production, such as the effect of transparency on the ability to perceive and comprehend words in an unknown language. This approach will also enable us to observe what learners do with their intake, and to contribute to the question of whether and how intake leads to learning.

Notes

* I sincerely thank Marzena Wątorek and Clive Perdue for their contribution to task development and data collection for this project, and Jean-Yves Dommergues for his assistance with data analysis. I would also like to thank an anonymous reviewer whose insightful comments helped me articulate my thoughts more effectively.

1. Statistics are provided in Rast and Dommergues (2003) and Rast (2008).

References

Carroll, S. (2001). *Input and Evidence: The raw material of second language acquisition*. John Benjamins, Amsterdam.

Carroll, S. (2004). Segmentation: Learning how to 'hear words' in the L2 speech stream. *Transactions of the Philological Society, 102*(2), 227–254.

Chaudron, C. (1985). Intake: On models and methods for discovering learners' processing of input. *Studies in Second Language Acquisition, 7*, 1–14.

Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics, 5*(4), 161–170.

DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *The Handbook of Second Language Acquisition* (313–348). Malden, MA: Blackwell.

Dimroth, C., Gullberg, M., Indefrey, P. & Roberts, L. (2006). The effects of exposure to an unknown L2+. Annual Report, Max Planck Institute for Psycholinguistics, Nijmegen.

Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. Doughty & M. Long (Eds.) *The Handbook of Second Language Acquisition* (256–310). Malden, MA: Blackwell.

Ellis, N. C. (2006a). Language acquisition as rational contingency learning. *Applied Linguistics, 27*(1), 1–24.

Ellis, N. C. (2006b). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*(2), 164–194.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New Approaches to Language Mechanisms* (257–287). Amsterdam: North-Holland.

Frauenfelder, U. & Tyler, L. (1987). The process of spoken word recognition: An introduction. *Cognition, 25*, 1–20.

Godfroid, A. & Housen, A. (2008). The role of noticing in incidental L2 vocabulary acquisition: Evidence from an eye-tracking study. *Proceedings of the 33rd International LAUD Symposium, Cognitive Approaches to Second/Foreign Language Processing: Theory and Pedagogy* (191–210). Essen: Linguistic LAUD Agency.

Hatch, E. (1983). Simplified input and second language acquisition. In R. W. Andersen (Ed.), *Pidginization and Creolization as Language Acquisition* (64–88). Rowley, MA: Newbury House.

Havranek, F. & Cesnik, H. (2001). Factors affecting the success of corrective feedback. *EUROSLA Yearbook, 1*, 99–122.

Hendriks, H. & Prodeau, M. (1999). Isn't Dutch a Mixture of English and German?: On the influence of a second language on the acquisition of a third one. Paper presented at *The Ninth Annual European Second Language Association Conference*, Lund.

Hudson Kam, C. & Newport, E. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development, 1*(2), 151–195.

Hulstijn, J. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition, 27*(2), 129–140.

Hulstijn, J. & DeKeyser, R. (Eds.) (1997). Special Issue, Testing SLA Theory in the Research Laboratory. *Studies in Second Language Acquisition, 19*(2).

Klein, W. (1986). *Second Language Acquisition*. Cambridge: Cambridge University Press.

Lyster, R. & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 19*(1), 37–66.

Mackey, A. & Gass, S. (Eds.) (2006). Special Issue, Interaction Research: Extending the Methodological Boundaries. *Studies in Second Language Acquisition, 28*(2).

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*(1–2), 71–102.

Matthey, M. (1996). *Apprentissage d'une langue et interaction verbale*. Bern: Peter Lang.

Peters, A. (1985). Language segmentation: Operating principles for the perception and analysis of language. In D. Slobin (Ed.), *The Crosslinguistic Study of Language Acquisition, Vol. II* (1029–1067). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rast, R. (2008). *Foreign Language Input: Initial Processing*. Clevedon: Multilingual Matters.

Rast, R. & Dommergues, J-Y. (2003). Towards a characterisation of saliency on first exposure to a second language. *EUROSLA Yearbook, 3*, 131–156.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 855–863.

Robinson, P. (2003). Attention and memory during SLA. In C. Doughty & M. Long (Eds.), *The Handbook of Second Language Acquisition* (631–678). Malden, MA: Blackwell.

Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.) *Implicit and Explicit Learning of Languages* (165–209). London: Academic Press.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129–158.

Sharwood Smith, M. (1996). The Garden of Eden and beyond: On second language processing. *CLCS Occasional Paper, 44*, 1–20.

VanPatten, B. (2000). Thirty years of input. In B. Swierzbin, F. Morris, M. Anderson, C. Klee & E. Tarone (Eds.), *Social and Cognitive Factors in Second Language Acquisition: Selected Proceedings of the 1999 Second Language Research Forum* (287–311). Somerville: Cascadilla Press.

Wong, W. & Simard, D. (2000). La saisie, cette grande oubliée! *Acquisition et Interaction en Langue Étrangère, 14*, 59–86.

Zwitserlood, P., Klein, W., Liang, J., Perdue, C. & Kellerman, E. (2000). The first minutes of foreign-language exposure. Unpublished manuscript, Max-Planck Institute for Psycholinguistics, Nijmegen.

Appendix

The sentence repetition test and the sentence translation test
(The same sentences were used for both tests.)

1. Piotr, mieszka w Krakowie, i studiuje informatykę.
2. Anna jest Włoszką, i mieszka we Włoszech.
3. Nauczyciel zna Marka, i mój brat zna go również.
4. Piotra mama, Marka zna również. (*Not analyzed*)
5. Piotra mama, wykładowcą jest na uniwersytecie.
6. Ewa zje lody. (*Not analyzed*)
7. Jestem w Warszawie, i znam ją bardzo dobrze.
8. Zupe Marek zje chętnie.
9. Marek, zna mojego kolegę, i mojego kolegę, zna również mój brat.
10. Anka i Marek, nie mieszkają w Krakowie, ale znają go dobrze.
11. Juan jest Hiszpanem, i mówi świetnie po polsku.
12. Lekarz, nie zna mojego kolegi.
13. Piwa napije się Piotr.
14. Lekarza, zna mój kolega.
15. Książke mojego kolegi, zna Piotr bardzo dobrze.
16. Podchodzi kelner do Jacka.
17. Mojego nauczyciela zna moja mama.
18. Marek pali papierosa. (*Not analyzed*)
19. Piotr, Marii, i Jackowi, opowiada film.
20. Marek jest studentem, i mieszka w Krakowie.

CHAPTER 7

On the stability of representations in the multilingual lexicon

Kees de Bot and Wander Lowie
University of Groningen

1. Introduction

Studies investigating the (multilingual) mental lexicon have traditionally used stable and invariant lexical representations as their starting point. This is essentially the position taken in models based on word associations and translation studies like the Revised Hierarchical Model (Kroll & Stewart, 1994), but also in current activation based models geared towards visual word recognition, like the Bilingual Interactive Activation model or BIA+ (Dijkstra & Heuven van, 2002). If it is the researchers' objective to explore the steps required for the recognition of a word, it would indeed be very inconvenient if the representation of the word changed in the process of its recognition. Yet, in view of recent insights into human cognition pointing to the intrinsic dynamics of mind and body (Van Orden, Holden, & Turvey, 2003), the assumption of invariant and static representations must be approached with some reservations. Instead, similar to other aspects of cognition, lexical representations must probably be regarded as dynamic, episodic and therefore inherently unstable.

In this paper we will address the issue of the (in)stability of lexical representations in the multilingual mind from both a theoretical and an empirical point of view. After discussing the most pertinent findings from recent insights about lexical representations and elaborating of the stability of representations, we will report on a single subject study investigating this issue. We will place this study into the context of a very early single subject study that we have replicated using modern techniques.

2. Representation and storage in the mental lexicon

Even though all researchers working on the mental lexicon will agree that word knowledge and word representations are affected by use, in fact the way lexical items are generally dealt with assumes an invariant representation. For instance, in the widely used semantic and phonological priming paradigm (Neely, 1991) it is assumed that the prime

causes a change in reaction times, while the control item is assumed to be stable and unaffected by the unrelated prime. Here we take a different perspective by including time as a factor in the representation of lexical items. Several interacting factors have an impact on representations over time. Our most important claim is that activation is about dynamically changing links in time. Representations are essentially not invariant but episodic with different types of links that can activate the next word, including meaning, linguistic context but also sensory perceptions, like smell, taste and touch. In this sense we are in line with van Gelder & Port (1995: 3) who argue in the introduction to their book *Mind as motion* that the main issue in the cognitive system is change over time:

> The cognitive system is not a discrete sequential manipulator of static representational structures: rather, it is a structure of mutually and simultaneously influencing change. Its processes do not take place in the arbitrary, discrete time of computer steps: rather, they unfold in the real time of ongoing change in the environment, the body, and the nervous system. The cognitive system does not interact with other aspects of the world by passing messages and commands: rather, it continuously coevolves with them.

Adding time as a factor in representation is part of a larger development in the study of the lexicon (de Bot, Lowie, & Verspoor, 2007; Van Geert, 2008) that seems to move from an invariant perspective to an episodic perspective (Merritt, Hirshman, Zamani, Hsu, & Berrigan, 2006). In the invariant perspective the assumption is that representations are canonical, stable and abstract, while in an episodic approach representations are rich, detailed and gradual. There is an interesting parallel with work on automatic speech recognition (ASR) (Strik, 2006). In ASR the task is to map incoming speech onto representations in the system that can be matched in order to arrive at an interpretation. Strik (2006) presents an overview of the contrast between an invariant and an episodic perspective on representations, showing many relevant oppositions between the two approaches:

Table 1. The invariant approach versus the episodic approach.

invariant approach	*episodic approach*
complex mapping	simple mapping
parsimonious representation	extensive representation
canonical representations	detailed stimulus information
Quantal	gradual
encode generalities	encode particulars
Categorical	graded
abstract units, often described by symbols	episodes, exemplars
abstractionist view	holistic view
analytic approach	analogical approach
variation is noise, a nuisance	variation contains information useful for perception
strip off variation during normalization	use variation
dissociation of form and content	mutual dependencies
linguistic information (content)	also indexical information (form)

Figure 1. Table 1 from Strik (2006)

As this overview shows, there are clear parallels between ASR models and models of human language processing. Therefore, the ASR models are also relevant for human language processing. One of the important characteristics of an episodic approach to the lexicon is that variation contains useful information, as it may reflect restructuring processes and the constant dynamics of the cognitive system. Variation can therefore not be discarded as 'noise'.

The main problem of the invariant approach is that it is not clear what the invariant parts comprise. From a computational perspective, relating the mental lexicon to Simple Recurrent Networks (SRNs), word 'types' are not explicitly represented, but are sensitive to their context and to prior use. Coleman (2002) discusses this issue with respect to phoneme representation and concludes that in speech production there are no invariant parts at all, as all sounds are influenced by the context in which they are used. This is in line with ideas expressed by Hawkins & Smith (2001): "Boundaries between categories are plastic, changing with factors such as surrounding phonetic context, talker, speaking rate and all kinds of linguistic experience such as lexicality, frequency and recency… The context sensitivity indicates that the mental representation of linguistic-phonetic categories is fundamentally relational and multidimensional" (2001:21). At the same time, the phonological representation within a fixed context like a word will be relatively stable. This apparent contradiction is best accounted for from a dynamic perspective of speech production, in which no invariant phonemes are the basis of articulation, but groups of articulatory gestures (see, for instance, Goldstein & Fowler, 2003).

3. Instability of lexical knowledge in L1 and L2

The apparent stability of lexical knowledge may also depend on subsets of the lexicon, in which a subset is to be defined from a multi-dimensional and dynamic point of view. The same lexical item can simultaneously be part of several subsets, like a register, a language variety or a language. Some of these subsets can be expected to be relatively more stable than others. The subset of lexical items associated with a language that is not someone's native language can be expected to be less stable than lexical items associated with the mother tongue, assuming that more frequent use and more frequent co-occurrence lead to less variable connections in the network. There are a number of studies that show that L2 lexical knowledge is not stable over time. Schmitt & Meara's (1997) year-long study of lexical development by Japanese learners of English showed that "small decrements often appear in repeated measures studies" and "vocabularies seem to exhibit a certain degree of flux over even relatively short periods of time. Words do not seem to be learned smoothly from one frequency level to the next, and even high-frequency words that are learned seem to be forgotten" (25) .

De Bot & Stoessel (2000) present data on two cases of the attrition of an L2 learned in childhood in an educational setting. Through a relearning paradigm some of the residual lexical knowledge could be reactivated, but the data show that the knowledge is highly unstable: words that were recognized in one session were not recognized in another one, and the other way around.

Hulsen, De Bot, and Weltens (2000) present data from Dutch migrants in New Zealand showing that over generations latencies in a naming task in Dutch become longer and standard variations increase proportionally. This suggests that retrieval becomes slower and more variable as a function of decreased exposure and use.

Montrul (2004) also shows that other parts of the language system, in particular the interfaces between syntax and other components are unstable and strongly sensitive to change.

All these observations suggest that the relative stability of representations depends on use. As we have argued in Section 2, even adult native speakers show a certain degree of instability of their lexical representation. The studies above demonstrate that instability and variation are particularly strong in language systems that are used in less than ideal language contact situations. It seems reasonable to assume that with increased use, representations will become more stable and are more easily retrieved. This is supported by data from Segalowitz (1991) who looked at the effect of level of L2 proficiency on L1 word recognition. His data show that automatic L1 word recognition appears to be unaffected by advanced levels of L2 proficiency, whereas controlled processing of words involving strategic behaviour may be.

In sum, there seems to be growing evidence for a perspective on the mental lexicon in which variation rather than stability is the rule. There may be various sources affecting stability and instability, like frequency of use, context and recency. In order to study these sources we need to establish whether there is some baseline variation that is inherent in systems of interacting variables, such as the mental lexicon in individuals.

The fundamental question we want to address is to what extent lexical representations are stable in L1 and L2 as measured by a simple processing task. It follows from the dynamic perspective we take that representations are inherently unstable and constantly changing due to internal restructuring in the lexical system. There is no point in time during which the mind is not changing. "There is simply no such thing as a static internal representation, as required by the computer metaphor of the mind." (Spivey 2007: 31) This variance is expected both between and within subjects and between experimental items. The assumption is that variability in a reaction time experiment reflects variability in representations and/or access of these representations.

4. Single subject studies on lexical processing

In an editorial for a special issue of the journal Second Language Research, Meara (1995) argues in favour of single subject studies on vocabulary acquisition. He argues that group-based studies using standard methodologies, 'rather quickly settle into a run-of-the-mill-approach' (ii). He argues for more studies in which not only single subjects are studied in their development but also for self studies in which the researchers themselves are the providers of the data. He refers to some of the groundbreaking work by Cattell from 1886 who studied his own ability to recognize words in English and German, using what we now consider to be primitive instruments to assess latencies. Cattell conducted a number of experiments to test the relation between sensory impulses and motor reactions to these

impulses. These included colour naming, letter naming and word naming. He also looked at the effects of attention by adding noise in perception tasks and at the effects of fatigue. For the latter he took no halfway measures; in one experiment, 1950 reactions had to be recorded, requiring 'the observer reacting almost continuously from early in the morning until late into the night' (537). Although no effects of fatigue were found over the day, its effect spilled over to the next day: 'In the case of C [Cattell himself] the brain substance concerned in the simple reaction seems to have been so far exhausted that his reaction-time remained abnormally long for two days' (540).

Some of the experiments reported are directly relevant for the present article since we will be using data that have been gathered partly from a single subject experiment. It is hard to make a direct comparison between the latencies found in Catell's experiments and the ones we will report here, since the experimental devices differ markedly. Still a number of findings with respect to word naming are interesting and relevant. Catell's data show that there was an overall tendency to shorter reaction times over the 11 day period. Although the developmental pattern is not smooth, this suggests a learning effect over sessions. Another remarkable finding is that both Catell and his assistant were faster in their L1 than in their L2. This finding cannot be attributed to the difference between the languages, as the assistant was a native speaker of German and Catell was a native speaker of English.

Cattell saw it as an advantage that the participants in his study (he and his assistant) were 'experienced observers', since they would know what they were doing and could adjust to the conditions.[1] The main reason why they took themselves as subjects was that they wanted to eliminate sources of variation as much as possible: 'I only give the determinations made on B and C: I have made similar determinations on other subjects of different age, sex, occupation, etc. but these can be better considered after we know the results of careful and thorough experiments on practiced observers' (233). And also 'The reaction time for practiced observers is consequently quite a constant quantity' (234). We will come back to the issue of researchers as subjects in the discussion.

More recently, Segalowitz, Watson & Segalowitz (1995) carried out a single case assessment to find out to what extent automaticity in word recognition as measured in a lexical decision task changes over time with both instructed and non-instructed words. This study is particularly relevant for our purpose, as it focuses on variability in retrieving lexical items, which we have argued to represent the relative instability of representations. The aim of their experiment was to find out to what extent decreases of latencies as a function of learning reflect changes in overall processing speed in what they call the 'blend of underlying mechanisms' in a lexical decision task. The 'blend of underlying mechanisms' reflects the blend of automatic and controlled processing components involved in visual word recognition, from letter identification to decision making. The subject was tested in four sessions over a period of 3 weeks. The assumption was that a reduction of the variability in the reaction time, proportionate to the reduction of the latencies themselves would reflect a general speeding up of the system, while a disproportionate change of the variability in relation of the latencies would reflect a change in the balance of automatic and controlled underlying processes. For this they used the coefficient of variation (CV) which is defined as the Standard Deviation divided by the Mean Reaction time (SD/RT).

The CV data showed that there were no clear patterns of development and that the subject did not become more skilled in doing the task repeatedly. The authors do report an effect of frequency, with higher frequencies showing lower CVs.

The present contribution follows this line of single case studies, in combination with between-subject data to investigate the relative stability of lexical representations. Similar to Cattell (1886) and Segalowitz et al. (1995), we will focus on changes of response time over time and similar to Segalowitz et al. (1995), we will concentrate on the change of variability in different conditions as represented by the coefficient of variance.

5. The experiments

The purpose of our experiments was to investigate the relative (in-)stability of lexical representations, focusing on variability and change. To this end, we set up two experiments involving response time measurement. The first experiment was a study involving a small number of bilingual participants that were tested in their L1 and L2 to assess inter-individual variation. The second experiment was a single subject study involving 10 repeated measures of the exact same L1 and L2 data to assess intra-individual variation.

To assess baseline variation in the mental lexicon we decided to use the simplest possible measure we could find, a word naming task (see, for instance, de Groot, Borgwaldt, Bos & Van den Eijnden, 2002). In a word naming experiment, subjects are presented with words on a computer screen they have to name as quickly as possible. The speech signal is picked up by a microphone and the response time is measured by the RT software.

In this section we will discuss the setup and the results of each of the experiments separately, followed by a general discussion. The materials and procedures were the same for the two experiments and will therefore be discussed first.

5.1 Materials

Since response times tend to be sensitive to a wide variety of characteristics of the words involved, like word length (New, Ferrand, Pallier, & Brysbaert, 2006), number of syllables (Jared & Seidenberg, 1990), number of neighbours (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), morphological complexity (Schreuder & Baayen, 1994) and word frequency (Monsell, 1991), we had to make a careful selection of items for the experiments. And since we wanted to focus on variation between conditions (L1 and L2) and on variation over time, the variation between the items had to be reduced to a minimum. We reduced the item variation through a number of steps. First we selected 300 high frequency 4–5 letter words in English and in Dutch, based on the CELEX/Cobuild lexical database (Baayen, Piepenbrock, & Gulikers, 1995). These items were tested in a pretest on six bilingual participants, all students from the English department of the University of Groningen,[2] in two sessions, one for Dutch, and one for English. The mean reaction times, SDs and CVs are presented in Table 1. Although this pretest was only used for the selection of items, the results show clear differences in both variability (CV) between the Dutch (L1) and English (L2) scores of the participants. This trend will be further explored in the actual experiment.

Table 1. English and Dutch pretest data based on 300 items (results in ms)

English	Part. 1	Part. 2	Part. 3	Part. 4	Part. 5	Part. 6
Mean	493	482	487	495	446	461
SD	80	76	59	83	131	98
CV	0.16	0.16	0.12	0.17	0.29	0.21
Dutch	**Part. 1**	**Part. 2**	**Part. 3**	**Part. 4**	**Part. 5**	**Part. 6**
Mean	458	493	449	442	537	467
SD	62	56	45	49	100	90
CV	0.14	0.11	0.10	0.11	0.19	0.19

To select the items showing the least variation, the z-scores for the individual items were calculated and the 100 items with the highest z-scores were removed from both language lists. In this way, we made sure that the remaining 200 test items showed minimal degree of variation between the items. This selection procedure was done for both the English and the Dutch dataset.

5.2 Procedures

All experiments took place in a sound proof booth. The experiment was administered as a self-paced naming task using E-prime 1.2 (E-prime, 2006) and the PST Serial Response Box (E-prime, 2006) to which a microphone had been attached. Before the experiment took place, the sensitivity of the microphone was tested and optimized. The items were presented one by one in three random blocks with 1 minute breaks between the blocks. The target was preceded by a fixation point in the middle of the screen for 1000 ms. Response times were logged and the sessions were recorded using a voice recorder for later reference.

5.3 Inter-individual variation

The next step was the assessment of inter-individual variation for this selected set of items. Four very advanced Dutch learners of English were tested in both languages. The same procedure was used as in the pretest, but now the items were presented in a fixed order to all participants. For each participant the means and SDs for the two languages were established using the same procedure as in the pretest. Before further analyses took place, obvious cases of outliers (due to early sound triggers or faulty registration) were deleted from the data. The cutoff points for the outliers in the Dutch test were 300 ms and 600 ms, which in all cases was more than 3 SDs from the mean RT in each direction. The cutoff points in the English test were 300 ms and 700 ms, which was also more than 3SDs from the mean RT in each direction. Then the correlation between subjects over items was calculated. The assumption was that when items have stable and fixed representations, correlations between subjects should be high.

5.4 Inter-individual results

The four very advanced learners of English were tested with the sets of 200 items in the two languages that remained after the pretest. The resulting means, SDs and coefficients of variance (CV) for English and Dutch are represented in Table 2.

Table 2. Means and SDs for English and Dutch

English	Part. 1	Part. 2	Part. 3	Part. 4
Mean	498	536	525	464
SD	56	71	88	70
CV	0.11	0.13	0.17	0.15

Dutch	Part. 1	Part. 2	Part. 3	Part. 4
Mean	510	497	448	387
SD	49	58	59	42
CV	0.10	0.12	0.13	0.11

These data indicate that even for these advanced learners, the latencies for L1 were shorter than those for L2. The CV is higher in L2 than in L1. In order to assess variation in latencies for items between participants, correlations have been calculated. Table 3 presents the correlations for Dutch and English. The correlations between participants appear to be very low and nonsignificant, suggesting that there are no consistent reactions to items, but rather variation between individuals.

Table 3. Correlations between participants over items

English Dutch	Part. 1	Part. 2	Part. 3	Part. 4
Part. 1		0.24	0.19	0.09
Part. 2	0.12		0.02	0.10
Part. 3	0.08	0.12		0.11
Part. 4	−0.02	0.21	0.11	

5.5 Intra-individual variation

To establish (within-subject) variation over time, one advanced learner of English (the first author) has been tested in a series of 10 sessions in five days over a period of 3 months. The first two sessions took place in early June 2007. The participant was tested early and late in the afternoon (1 pm, session 1a and 6 pm, session 1b) on one day and again on the same moments in time the next day (sessions 2a and 2b). Based on these sessions, variation within one day and between two consecutive days could be established. In order to test variation over weeks, two sessions on the same day (3a and 3b)

were administered three weeks after the first two sessions. One month later, the participant was tested again twice on the same day, this time after a week in which he nearly only used English, both professionally and privately (with visiting English speaking relatives). Three weeks later the final sessions took place. These sessions followed a week in which he refrained from using English completely, again both professionally and privately. The last two sets of data should allow us to assess the impact of frequency of use on stability of representation.

In short, to assess intra-individual variation, the same participant was tested 10 times with intervals of a few hours, a day and three weeks, and again twice on the same day after an English only and a non-English period. The same set of items in the same order was used throughout the sessions. The English test consistently preceded the Dutch version.

5.6 Intra-individual results

The second set of data aimed to assess stability of representations over time within a single subject. As indicated before, there were 10 sessions in total. First we will present the outcomes for the first 6 sessions in Tables 4 and 5. These data show longer latencies for English than for Dutch and no clear development over time. CVs are fairly similar for the two conditions.

Table 4. Means and Sd's English from 6 sessions

Session English	Mean	SD	CV
1a: day 1 1pm	528	50	0.09
1b: day 1 6 pm	522	53	0.10
2a: day 2 1 pm	513	52	0.10
2b: day 2 6 pm	479	38	0.08
3a: day 22 9 am	516	41	0.08
3b: day 22 2 pm	489	44	0.09

Table 5. Means and SDs over 6 sessions Dutch

Session Dutch	Mean	SD	CV
1a: day 1 1 pm	454	30	0.07
1b: day 1 6 pm	470	43	0.09
2a: day 2 1 pm	465	38	0.08
2b: day 2 6 pm	470	38	0.08
3a: day 22 9 am	464	36	0.08
3b: day 22 2 pm	481	41	0.09

Table 6. Correlations between for the RTs of the individual items in 6 sessions (1.1–3.2) for English (horizontally) and Dutch (vertically). Significance of correlations signalled by * ($p < 0.05$) and ** ($p < 0.01$)

English → Dutch ↓	**1.1**	**1.2**	**2.1**	**2.2**	**3.1**	**3.2**
1.1		0.45**	0.12	0.32**	0.17*	0.29**
1.2	−0.16*		0.22**	0.22**	0.22**	0.24**
2.1	0.11	0.16*		0.14	0.02	0.16*
2.2	0.10	0.09	0.14		0.19	0.13
3.1	0.03	0.16*	0.17*	0.16*		0.21**
3.2	−0.03	0.12	0.13	−0.12	0.02	

In order to test stability of items over sessions, correlations for individual items between sessions were calculated for English and Dutch. The correlations are presented in Table 6. These data show that the correlations are generally very low, apart from the one for first and second session in English. For the other pairs less than 5% of the variance is explained. Figure 2 presents a part of the reaction time data from the English experiment over six sessions. The data show that there is considerable variation between sessions and items and that therefore the low correlations cannot be explained by a restriction of range effect.

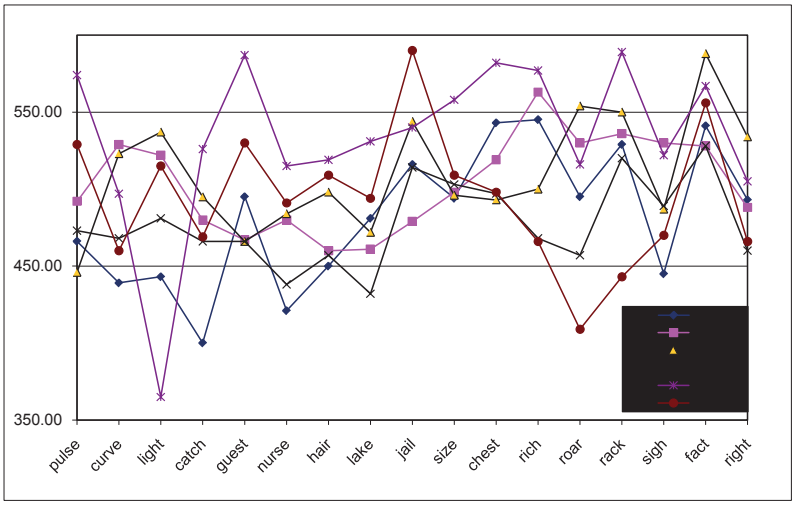


Figure 2. Excerpt from RTs for individual Dutch items in 6 sessions, represented by the different lines

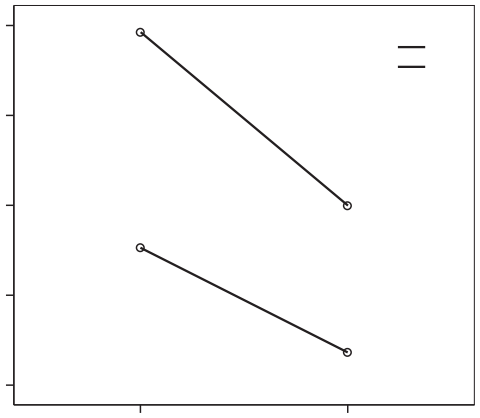


Figure 3. Mean RTs for the Dutch (1) and English (2) test in the two immersion contexts

The last 4 sessions are presented as one set, since the focus was on the same issue: to what extent are means and SDs sensitive to use or no-use of one of the languages. The sessions 4a and 4b were administered after 10 days of English only,[3] while sessions 5a and 5b were administered after a week without using English. Figure 3 presents the mean reaction times for the test sessions after the maximum-English week and the no-English week.

In Tables 7 and 8 the mean response time latencies and SDs are presented for English and Dutch. 'a' refers to the morning sessions, 'b' to the afternoon session.

Table 7. Means and SDs for 4 sessions for English: 4a and 4b after complete English exposure, 5a and 5b after "no English" period

Session English	Mean	SD	CV
4a	547	44	.08
4b	528	43	.08
5a	503	36	.07
5b	498	36	.07

Table 8. Means and SDs for 4 sessions for Dutch: 4a and 4b after complete English exposure, 5a and 5b after no English period

Session Dutch	Mean	SD	CV
4a	492	34	.07
4b	490	33	.07
5a	455	30	.07
5b	481	36	.07

A repeated measures ANOVA items analysis with Language, Context and Time as variables revealed that the main effect for language was significant, $F(1,174) = 333.8$, $p < 0.01$, with Dutch consistently faster than English. The main effect of Context also turned out to be significant, $F(1,174) = 334$, $p < 0.01$, with faster means response times after the "No English" period. The main effect of Time (morning vs afternoon) was not significant. Significant interactions were found for Context x Language, $F(1, 174) = 18.6$, $p < 0.01$, with the strongest language effect in the "Max English" condition (see Figure 3). The interaction between Time and Context also turned out to be significant, $F(1,174) = 29.1$, $p < 0.01$, with the strongest Context effect in the morning sessions (see Figure 4). Finally, a significant interaction was found between Time and Language, $F(1,174) = 39.5$, $p < 0.05$, with the strongest Language effect in the morning sessions (see Figure 5).

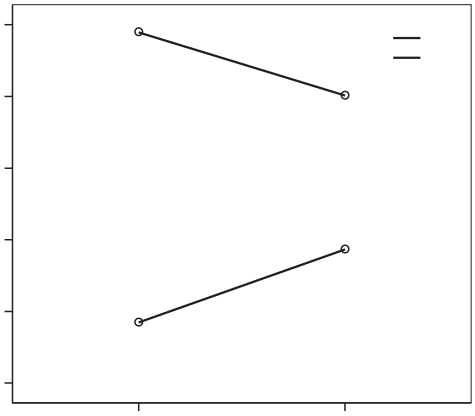


Figure 4. Mean RTs for the interaction between Context and Time. Context 1 is the "Max English" context, Context 2 is the "No English" context

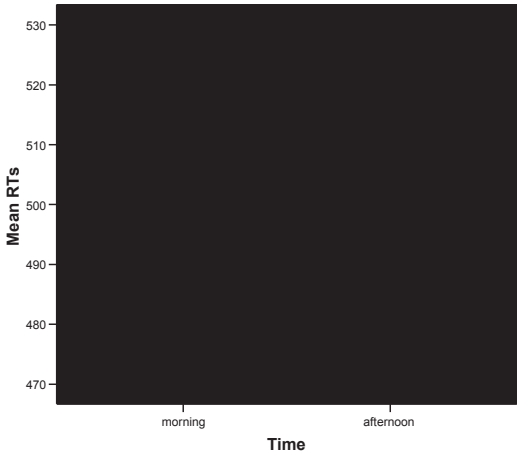


Figure 5. Interaction of RTs for Time x Language. Language 1 = Dutch, Language 2 = English

Table 9 represents the correlations for the four sessions for English and Dutch. Again we can see longer RTs and SDs for English than for Dutch, and very low correlations between sessions over items.

Table 9. Correlations for the scores on the individual items between the 4 sessions for English (horizontally) and Dutch (vertically). Significance of correlations signalled by * ($p < 0.05$) and ** ($p < 0.01$)

English **Dutch**	**4a**	**4b**	**5a**	**5b**
4a		.083	.251**	.111
4b	.071		.212**	.132
5a	.193*	.071		.152*
5b	.150*	.145*	.156*	

Finally, we looked at the correlations between mean RTs and SDs for the individual items. In line with the finding that the correlations between the different tests were not significant, pointing to variability between the items, the RTs and SDs of individual items varied strongly. Some items showed very little variability (low SDs). Also the CVs of the individual items varied strongly over the different sessions, ranging from 0.03 for the item *lach* ('laugh') to 0.14 for the item *helft* ('halve') in the Dutch data. Also in the English data the RTs and SDs varied widely and the same was found for the CV (0.13 for *dark*; 0.03 for *goods*). Some words had a relatively low SD (20 ms for *meat*) and some a very high SD (72 ms for *young*).

In the English subtest, the correlation between the SD and the Mean RT for all items was not significant ($r = 0.03$). This is illustrated in Figure 6.

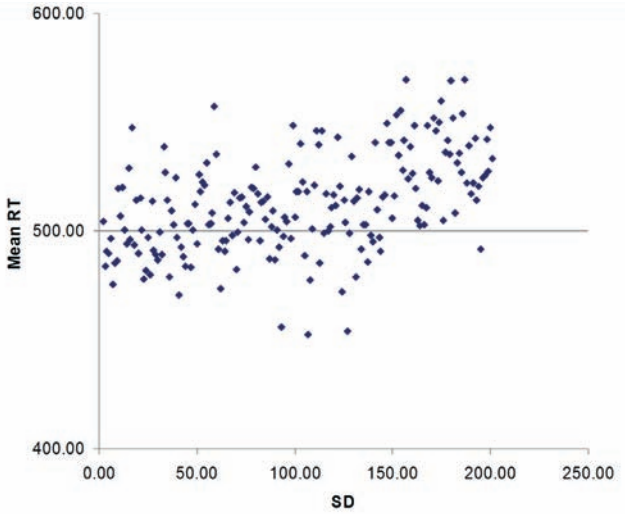


Figure 6. SD and Mean RT for all 200 items in the English subtest

For the Dutch subtest the correlation between Mean RT and SD was 0.14 (ns). This is illustrated in Figure 7.

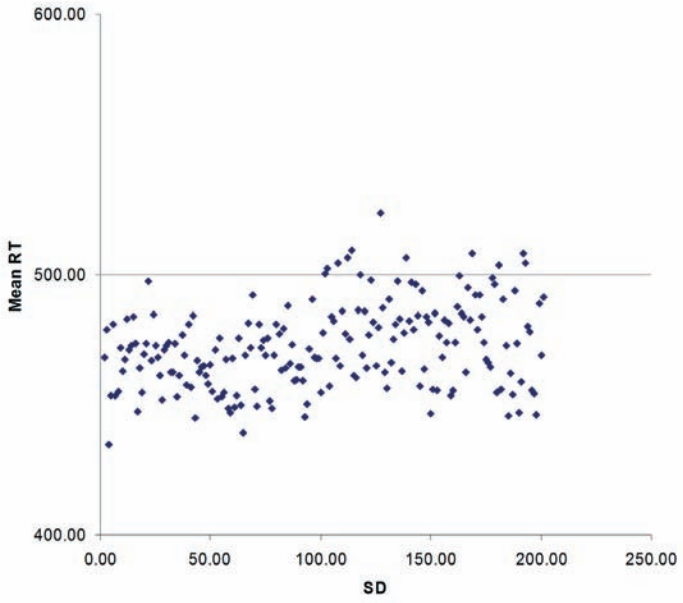


Figure 7. SD and Mean RT for all 200 items in the Dutch subtest

6. Discussion and conclusions

The results of our experiment show that response times in the L1 are significantly faster than response times in the L2. This finding is not very surprising and has been reported before (Dijkstra 2005). When the same test was administered after maximum exposure to English and after a minimum of exposure to English, the striking result is that the RTs for both the English and the Dutch subtest are faster after a minimum of exposure. The English test was even more strongly affected by the difference in context than the Dutch test. This result is hard to account for, and must probably be sought in the fact that the minimum of exposure to English coincided with a relatively quiet period with less travel and interaction.

The selection of items using strict initial criteria and subsequent pretesting was focused on keeping the variation between the items to a minimum. Nevertheless, the amount of variation found within the individual items is considerable. Generally, the standard deviations in the English test were larger than the standard deviations in the Dutch test. This observation cannot be explained by the generally lower RTs in the English subtest, as no correlation was found between mean RT and SD. The mean RTs over sessions clearly showed that the variation in response times could not be caused by item-specific

characteristics. Moreover, partial data analyses showed that this result could not be accounted for by periods of decreased concentration. The only conclusion that can logically be drawn from this is that the activation of individual lexical representations is quite variable. Qualitative analyses revealed that some individual representations (like *meat*) were relatively stable, while others (like *young*) very unstable. We also found that the instability of lexical representations was found to be strongest for the L2 English items. The variability found in the data cannot be accounted for by a change in the blend of underlying processes due to increased complexity of the task. The explanation that is not contradicted by the results of our experiment is an episodic view of the lexicon, in which activation of lexical representation varies as a function of recency of use and type of association.

From the findings presented several conclusions can be drawn, keeping in mind the limitations of the type of single subject design used here. The first is that there is considerable variation that poses problems for current processing models as presented earlier. The second is that while representations are clearly not stable, it is hard to tell how much instability the language processing system can deal with. The obvious ease of delivery in normal language production and language perception suggests that the variability and instability do not pose serious processing problems. A third and more fundamental issue is whether outcomes like the ones presented should have an impact on our thinking about within and between subject variability. If the variation in repeated measures within a subject is similar to that found between subjects this would have considerable consequences of how we should look at statistical approaches like analyses of variance, but it would be preposterous to draw such massive conclusions on the basis of a single and maybe biased subject.

Having said that, the line of research proposed here, using multiple case studies in investigating the multilingual mental lexicon, is definitely promising. The use of reaction time data on isolated items is not unproblematic from a dynamic perspective. As Spivey (2007) argues: "The fundamental weakness of some of the major experimental techniques in cognitive psychology and neuroscience is that they ignore much of the time course of processing and the gradual accumulation of partial information, focusing instead on the outcome of a cognitive process rather than the dynamic properties of that process." (53). In this article we have tried to move one step further by looking at relations between items over the whole experiment rather than isolated items

In future studies four directions of development could be taken. First, the influence of different types of variables could be explored, like word type (cognate status, abstractness). Second, the dynamic approach could be further investigated by making the task more cognitively demanding (e.g. adding finger tapping to the task) and thereby limiting the availability of resources. Third, we need to gather data that are more continuous in nature: "To properly address the temporal continuity of perception, action, and cognition, more cognitive experimentation in the future will need to not only use more continuous response measures but also use continuous (and ecologically valid) dynamic sensory stimulation." (Spivey, 2007: 63) Finally, advanced statistical analyses, like time series analyses, autoregression and Fourier analyses of reaction time data could be applied to further probe the dynamic relation between variables.

What does all of this mean for our views on bilingual language processing? It seems like we are about to witness a major change in paradigms in which the computer metaphors

referred to by Spivey and others that imply static representations and a modular approach to language processing needs to be revised. What the new models and paradigms will look like is very unclear and we are faced with a situation in which the old models no longer seems to be valid while no alternative models with equal sophistication are available. One thing that is obvious from what is happening now on the dynamic stage is that discreteness will be replaced by continuity on all levels and at all time scales and that the type of modularity that has dominated models of language processing over the last decades will vanish. It is very obvious that the new models will have a time dimension that was clearly missing in what now might be called 'old school' models. Clearly, the present paper can be no more than a modest first attempt to investigate the status of lexical representations from a dynamic perspective, but is foremost a reflection of the struggle with old and new ways of thinking that is typical of a period of paradigm changes.

Notes

1. 'The two subjects (Dr. G. O. Berger and the writer) on whom the determinations were made had already had much practice in psychological work. They were in good health and lived regularly, not even using coffee.' (Cattell 1886: 230)

2. None of these participants took part in the actual experiment.

3. Complete non-use is hard to achieve for both languages, since there will inevitably be signs like advertisements or lyrics on songs on the radio.

References

Baayen, Harald, Richard Piepenbrock, & Leon Gulikers (1995). The *CELEX Lexical Database*. [(Release E25)], http://celex.mpi.nl/. Philadelphia, PA, Linguistic Data Consortium.

Cattell, James M. (1886). The time taken up by cerebral operations. *Mind, 11*, 523–538.

———. (1886). The time taken up by cerebral operations. *Mind, 11*, 377–392.

———. (1886). The time it takes to see and name objects. *Mind, 11*, 63–65.

Coleman, James (2002). Phonetic representations in the mental lexicon. In J. Durand & B. Laks (Eds.), *Phonetics, phonology and cognition* (96–130). Oxford: Oxford University Press.

Coltheart, Max, Kathleen Rastle, Conrad Perry, Robyn Langdon, & Johannes Ziegler (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256.

De Bot, Kees, Wander Lowie, & Marjolijn Verspoor (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition, 10*, 7–21.

De Bot, Kees & Saskia Stoessel (2000). In search of yesterday's words: Reactivating a long forgotten language. *Applied Linguistics, 21.3*, 364–388.

De Groot, Annette, Susanne Borgwaldt, Mieke Bos, & Ellen van den Eijnden (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language, 47,* 91–124.

Dijkstra, Ton (2005). Bilingual visual word recognition and access. In J. Kroll & A. de Groot (Eds.), *Handbook of bilingualism. Psycholinguistic approaches* (179–201). Oxford: Oxford University Press.

Dijkstra, Ton & Walter J. B. van Heuven (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175–197.

E-prime v. 1.2 (2006). Pittsburg (PA): Psychology Software Tools Inc. http://www.pstnet.com.

Goldstein, Louis. & Carol A. Fowler (2003). Articulatory phonology: A phonology for public language use. In N. Schiller & A. Meyer (Eds.), *Phonetics and phonology in language comprehension and production* (159–208). Berlin, New York: Mouton de Gruyter.

Hawkins, Sarah & Rachel Smith (2001). Polysp: A polysystemic, phonetically-rich approach to speech understanding. *Journal of Italian Linguistics/Rivista de Linguistica, 13.1*, 99–188.

Hulsen, Madeleine, Kees de Bot, & Bert Weltens (2002). Between two worlds: Social networks, language shift and language processing in three generations of Dutch migrants in New Zealand. *International Journal of the Sociology of Language, 153*, 27–52.

Jared, Debra & Mark S. Seidenberg (1990). Naming multisyllabic words. *Journal of Experimental Psychology: Human Perception & Performance, 16,* 92–105.

Kroll, Judith & Erika Stewart (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149–174.

Meara, Paul (1995). Single-subject studies of lexical acquisition. *Second Language Research, 11.2*, i–iii.

Merritt, Paul, Elliot Hirshman, Shane Zamani, John Hsu, & Michael Berrigan (2006). Episodic representations support early semantic learning: Evidence from midazolam induced amnesia. *Brain and Cognition, 61,* 219–223.

Monsell, Stephen (1991). The nature and locus of word frequency effects in reading. In D. Besner and G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (148–197). Hillsdale, NJ: Erlbaum.

Montrul, Silvina A. (2004). Convergent outcomes in L2 acquisition and L1 loss. In M. Schmid, B. Köpke, M. Keijzer, & L. Weilemar (Eds.), *First language attrition* (259–280). Amsterdam: John Benjamins.

Neely, James (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner, & G. W. Humphreys (Eds.), *Basic processes in reading* (264–336). Hillsdale, NJ: Lawrence Erlbaum Associates.

New, Boris, Ludovic Ferrand, Christophe Pallier, & Marc Brysbaert (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review, 13,* 45–52.

Schmitt, Norbert & Paul Meara (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition, 19.1*, 17–36.

Segalowitz, Norman (1991). Does advanced skill in a second language reduce automaticity in the first language? *Language Learning, 41*, 59–83.

Segalowitz, Norman, Vivien Watson, & Sidney Segalowitz (1995). Vocabulary skill: Single-case assessment of automaticity of word recognition in a timed lexical decision task. *Second Language Research, 11.2*, 121–136.

Spivey, Michael (2007). *The continuity of mind.* Oxford: Oxford University Press.

Schreuder, Robert & R. Harald Baayen (1994). Prefix stripping re-revisited. *Journal of Memory and Language, 33*, 357–375.

Strik, Helmer (2006). How to handle pronunciation variation in ASR by storing episodes in memory? *ESR-SRV06*, 33–38.

Van Geert, Paul (2008). The Dynamic Systems approach in the study of L1 and L2 acquisition: An introduction. *Modern Language Journal, 92,* 179–199.

Van Gelder, Tim & Robert Port (1995). It's about time: An overview of the dynamical approach to cognition. In R. Port & T. van Gelder (Eds.), *Mind as motion: Exploration in the dynamics of cognition* (1–45). Cambridge, Mass: MIT press.

Van Orden, Guy C., John G. Holden, & Michael T. Turvey (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology-General, 132,* 331–350.

CHAPTER 8

Conceptual representations in the multilingual mind

A study of advanced Dutch students of English

Wander Lowie, Marjolijn Verspoor and Bregtje Seton
University of Groningen

In this contribution, we describe an experiment investigating the influence of first language Dutch on second language English conceptual representation. Native-like and Dutch-based word associations were included in a priming lexical decision task to study subtle differences between conceptual characteristics of native and non-native words. The reaction time study revealed some interesting differences between native speakers of English and advanced Dutch students of English. These results are in line with a dynamic distributed activation model of the multilingual lexicon that allows for different degrees of acquisition of a lexical item. Even very advanced learners may have acquired L2 conceptual representations that deviate from those of native speakers.

1. Introduction

What is a word and what is a concept? How is a word related to its concept(s)? Actual conceptual representation is difficult to grasp, but we will assume that a concept emerges from the physical, emotional, cognitive, pragmatic, and linguistic experience an individual has with a particular lexical item, phrase or construction, resulting in associations of varying strength. As Langacker (2000) suggests in his Usage-Based model, these associations are part of "massive networks in which structures with varying degrees of entrenchment, and representing different levels of abstraction, are linked together in relationships of categorization, composition, and symbolization" (p. 5). Associations can thus be numerous and change over time, causing them to be part of a dynamic network.

Because associations are the result of personal experience, they do not only change over time within the individual, but they also differ among different individuals within a speech community. Nonetheless, because of their interaction and therefore their mutual experience, people may negotiate and exchange commonalities in their conceptual representations. Sharifian (2008) calls this conceptualization "heterogeneously distributed cultural cognition" (p. 6), meaning that the conceptual representation of two individuals belonging to the same cultural group will partly overlap and partly differ. In this way, both conceptual similarities within a cultural group and differences between cultural groups

can be accounted for. Following this argumentation, it can be expected that different languages bring along different conceptual representations as well. Our question is to what extent the conceptual representations of languages differ and to what extent the concepts of an L1 may influence those in L2.

From a developmental point of view the question is whether L2 learners can attain L2 concepts that are identical to the concepts of native speakers. A great number of studies have investigated this question using translation studies and cross-language priming tasks. In these studies the question is if the words in the L2 mental lexicon are *conceptually mediated* or if they can only be accessed through the L1 lexicon. For instance, a major conclusion of the translation studies carried out by Judith Kroll and her associates (e.g. Kroll & Stewart, 1994) is that translation from L1 to L2 is slower and more troublesome than translation from L2 to L1. Kroll and Stewart refer to this observation as "translation asymmetry", which they claim is caused by direct (L1) versus indirect (L2) access to conceptual representations. Other studies report that "when translating words in either direction, bilinguals of various L2 fluency levels apparently access and exploit conceptual memory representations at least most of the time" (de Groot & Poot, 1997: 252). The starting point of this discussion is always that conceptual representations are language independent and that the language specific lexical semantic information has to be matched with the conceptual representations. From a usage-based perspective, though, one would argue that conceptual representations are not language specific but specific to each individual and through a bottom-up process (interaction among individuals) they become somewhat more specific to a speech community.

Taking the usage-based approach, one would expect that the concept developed in the L1 will always play some role in the L2, because no matter how much one has used the L2, one cannot completely erase all previous experience and associations, and it would be difficult to develop a conceptual representation similar to one by a native speaker of the target language.

However, how can it be ascertained to what extent an L2 concept has been acquired and to what extent it is native-like? Thus far, the overlap of conceptual representations has been expressed in the percentage of overlap between within-language and between-language word associations (Kolers, 1963; Taylor, 1976). In addition to the objections against using a translation task for word associations (see Van Hell & De Groot, 1998) this purely quantitative approach can only roughly indicate a measure of conceptual overlap, and does not give any insight in the actual conceptual representation in multilingual speakers. We will claim that even for seemingly very strongly overlapping meanings of words in two different languages, like concrete cognates, there may be subtle differences in semantic interpretation. In other words, even though both form and meaning are similar, an English *rose* is not necessarily the same as a Dutch *roos*. We will refer to converging evidence from two areas of research, word association studies from cognitive linguistics and reaction time experiments from psycholinguistic research to point to the many fine grained differences between seemingly equivalent words in different languages and to gain insight into the organization of the conceptual representations of bilinguals.

The present study is concerned with how the *culturally dependent cognition* of one language influences native speakers of this language when they are speaking their second

language. Being interested in the conceptual representation of bilingual speakers, we investigated the differences in conceptual representation between advanced learners of English (Dutch third-year university students) and native speakers of English. To come to this, we used word association tests to compare conceptual representation in both English and Dutch. These different word associations were used as primes to their targets in a priming experiment, in which a lexical decision task was used to gather reaction times to the target words. With this experiment, we wanted to put conceptual representation in a multilingual perspective and investigate if people's first language conceptual representation influences their conceptual representation in another language, even when being an advanced speaker of this language.

2. A usage-based and dynamic approach to the mental lexicon

Most, if not all current models of the multilingual mental lexicon assume an interactive activation network of lexical items. Rather than assuming separate lexical representations for words in different languages, it has now been convincingly demonstrated that the lexicon must be seen as one network, regardless of language or register. At the same time, most models assume some kind of "language tag" that is associated with each item in the lexicon. In this way, the recognition of a word in one language may inhibit lexical items from other language subsets (Dijkstra & Van Heuven, 1998; Dijkstra & Van Heuven, 2002). The mechanism that is assumed to account for lexical selection is referred to as "activation". Each lexical item has its own level of activation that increases with use and that decreases over time. As the lexicon is considered a network, the activation of one lexical item can lead to increased activation levels of lexical items that are closely associated with that item. Lexical items that are associated with one particular language can thus be assumed to interactively activate each other, allowing for faster recognition of words from that lexical subset. Subsets of lexical items are not limited to words that belong to the same language. Similar mechanisms can be assumed for any characteristic that words share, like register or conceptual overlap. In this way, a multidimensional picture of the mental lexicon emerges in which each lexical item can simultaneously be part of an infinite number of subsets. The activation metaphor allows for a flexible interpretation of the subsets. Even though in principle lexical items from a particular language subset will be activated in a language specific situation, like speech production, individual lexical items may be activated that share other characteristics and that happen to have a very high level of activation due to recency or frequency of use. This accounts for the observation that while speaking one language, lexical items from other language subsets may be predominant. As argued for the BIA (Bilingual Interactive Activation) and BIA+ model (Dijkstra & Van Heuven, 2002), another principle of lexical selection is competition. The selection of one lexical item will suppress the activation level of competing words in the BIA model: "activated language nodes send inhibitory feedback to all word nodes in the other language" (Dijkstra & Van Heuven, 2002: 176). More information on language processing and comprehension models can be found in Seilhamer's (this volume) discussion of different models. In BIA+, the role of the language node is moderated and it may even be argued, as has

been done by Jacquet and French (2002), that it is redundant. If all lexical items contain information about the language to which they belong, it is unnecessary to assume a separate language node if the starting point is a distributed rather than a modular model of lexical processing. Having said that, a language node may be required to control language specific production processes (Kroll & Dijkstra, 2002). Although the BIA+ model has shown to be able to account for a wide variety of research results, especially on visual recognition, increased explanatory power can be expected if the ideas of BIA+ are considered from the perspective of a distributed model. A distributed approach is not only helpful in accounting for language selection, but it can also account for the development of the lexicon and its dynamic nature. Furthermore, it can account for the semantic overlap between lexical items, which is left implicit in localist processing models. Below we will discuss each of these issues in some more detail.

The emerging picture of the mental lexicon is a multidimensional connectionist model with interactive activation spreading between the lexical nodes and in which lexical selection takes place through inhibition and facilitation. Approaches like the BIA model give an accurate and detailed account of the process of visual word recognition, especially in experimental settings. However, these models are not well equipped to account for the development of the lexicon. In a distributed model of the lexicon, as proposed by for instance Kawamoto (1993), McRae, Seidenberg, and De Sa (1997) or Hummel and Holyoak (2003), the dynamic nature of lexical connections, required for development and change, can be easily accounted for. Unlike a localist approach to the lexicon, a distributed model does not assume a one-to-one mapping of meaning representations and lexical items and therefore allows for partial conceptual links (or a pattern of activation) to the lexical item, in some models specified as "featural representations" (McRae, Seidenberg, & De Sa, 1997). The assumption of partial conceptual representations have two clear advantages over unitary representations. First, partial representations can easily and intuitively account for the partial acquisition of lexical items and for deviations between native speaker representations and learner representations. The partial representations will constantly change in a dynamic lexicon and can therefore also account for the instability of lexical representations (see de Bot & Lowie, this volume). Secondly, distributed representations can account for the (partial) overlap between the different lexical items. Since a distributed representation assumes partial representations at all levels of lexical representations, from the orthographic-phonological level to the conceptual level, partial overlap can also be assumed at all levels. A representation of multilingual links at several levels is represented in Figure 1.

The overlapping words (like *dress* and *skirt*) share meaning representations. For words in a connectionist framework this means that the activation pattern overlaps, leading to mutual activation of related words. The more activation is shared between conceptual representations, the more likely these items are to be associated with each other when language users are asked to give word associations. Moreover, as is argued by Van Hell and De Groot (1998), "the more the activation pattern of the stimulus resembles that of the (potential) associative response, the less time will be needed to move towards this pattern" (p. 207). This model accounts for the observation that conceptual representations related to L1 and L2 may show various degrees of overlap at different language levels. For words

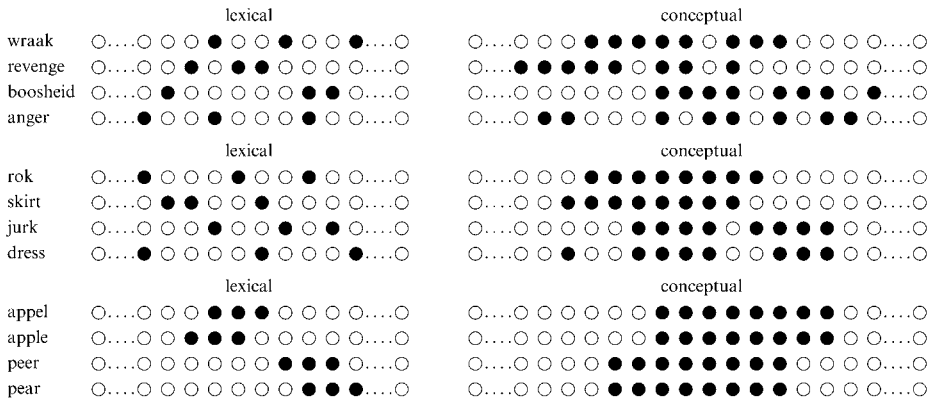


Figure 1. Distributed representation of the lexical and conceptual level of three related translation pairs in a Dutch-English bilingual lexicon. The units are interconnected within and between subsets (from: Van Hell & De Groot, 1998: 206)

that are identical between two typologically related languages, it is then assumed that they share the same semantic and conceptual representation. The degree to which words share the same representations can be seen to vary as a function of their cognate status, the concreteness and the part of speech. Concrete cognates are most likely to be *identical* between languages, while abstract non-cognates can be assumed to have different semantic representations. Van Hell and De Groot (1998) show that "in both within- and between-language association, retrieving an associate was easier to concrete than to abstract words, to cognates than to non-cognates, and to nouns than to verbs" (193). The distributed activation model can accommodate any level of overlap.

What is most important for a model of bilingual conceptual representations is that it is able to account for change. Initially, L2 learners will assume full conceptual overlap between lexical items in the L1 and in the L2. Exposure to L2 will lead to the discovery of new conceptual associations with L2 lexical items and subsequent restructuring of the links between lexical items and conceptual representations. The dynamic nature of the (multilingual) lexicon can be illustrated by considering two time slots in the development of lexical items, as illustrated in Figure 2. In this figure the dynamic lexical restructuring is displayed for two related English words, *latest* and *last*. In Dutch there is only one lexical item for the two English ones. Initially (at t_1), the Dutch learner will assume full overlap between the words *last* and *laatste*. When the learner 'discovers' the English word *latest* (at t_2) a restructuring will have to take place to create the activation pattern for this new item. This type of lexical restructuring is an ongoing process, which is found for native speakers and learners alike. Continuous lexical restructuring controlled by principles of activation make this model essentially usage based. In agreement with recent approaches of language acquisition as a dynamic system (de Bot, Lowie, & Verspoor, 2007), there is no principled distinction between use and acquisition. Second language development is only different from native speaker language use by the amount of language use and the context

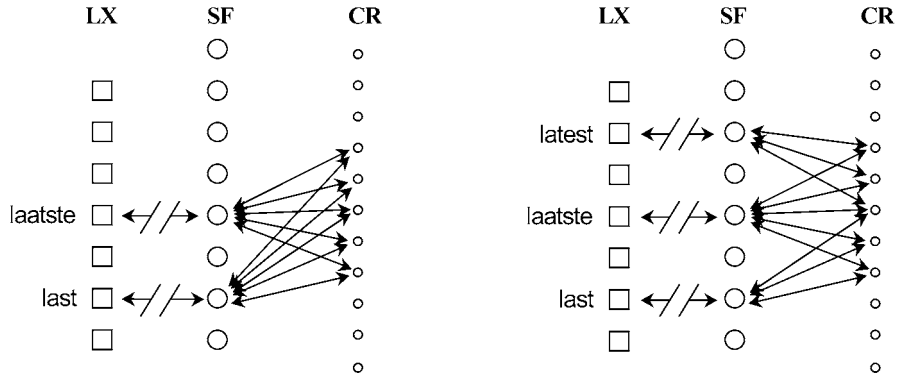


Figure 2. Two time slots in the development of a set of overlapping lexical items in a dynamic distributed activation model of the multilingual lexicon. LX = lexical representation; SF = word node; CR = conceptual representation (Lowie, 1998: 101)

in which the language is used. Applied to the conceptual representations in the multilingual mental lexicon, we can say that increasing exposure and use of L2 may eventually lead to native-like representations of L2 lexical items. However, there is a fair chance that many L2 lexical items maintain traces of L1 concepts and that due to different activation patterns L2 lexical items may deviate from those of native speakers in quite subtle ways. Therefore the next question we will address is to what extent concepts may be considered general or language specific.

3. Shared or language specific concepts

We have argued that the dynamic distributed activation model of the multilingual lexicon described above can account for the partial overlap between lexical items of different languages. Discussions on conceptual representations in the bilingual mind have concentrated on the question if L1 concepts can be considered identical to L2 concepts. Dong, Gui, and MacWhinney (2005) have argued for a converging view of conceptual representations, in which concepts contain both L1 specific information and L2 specific information. Their study contains two experiments with third-year Chinese students of English, a cross language lexical decision priming task and a word association task. In the priming task they compared four language conditions: Chinese-Chinese (CC), English-Chinese (EC), Chinese-English (CE) and English-English (EE). The experiment showed within language priming effects, but also cross-language priming effects for both EC and CE, with faster responses in the EC condition than in the CE condition. The authors interpret this as evidence that there is "a shared conceptual system for the two vocabularies in the bilingual's mind" (2005: 227) and that the conceptual links are asymmetrical. In the second experiment the participants provided closeness rankings to Chinese-English word pairs.

The results of this study show that "conceptual differences between a pair of translation equivalents tend to converge in the mind of L2 learners. The more advanced the L2 is, the greater co-effects the two languages produce on the conceptual representations of the two languages" (Dong et al., 2005: 232). At the same time, the data revealed that learners tend to "maintain their L1 conceptual system in the representation of the L1 word and to adopt the L2 conceptual system in the representation of the L2 word" (Dong et al., 2005: 233). Based on these results, the authors argue in favour of a shared distributed asymmetrical model, which is outlined in Figure 3 below. The resulting picture can be seen as a refinement of the revised hierarchical model (Kroll & Stewart, 1994). The significant improvement to this model is the recognition that shared components can coexist with language specific components. However, the representation of languages as separate units is misleading and shows traces of spatial metaphor. The findings from this study can equally well be accounted for by dynamic distributed activation model. This model has two additional advantages: it can account for representations at the level of individual lexical items and it can account for the continuously changing nature of the lexicon.

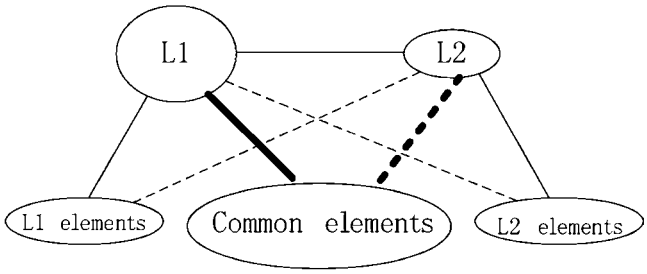


Figure 3. The shared (distributed) asymmetrical model (Dong et al., 2005)

4. Word associations and conceptual representations

Whereas Dong et al. (2005) incorporated primes and targets that were chosen according to their semantic relationships, word association studies can actually give real-time evidence on what the existing links are between lexical items. These connections eventually make up a conceptual representation. Earlier word association research in the monolingual field (e.g. Entwistle, 1966) has shown that adults often have paradigmatic associations like synonyms and antonyms, while young children give more syntagmatic responses like for example collocational associations. Another type of response is a clang association, which is a phonologically or morphologically similar word (Verspoor, 2008). This information from first-language learners forms a basis for research in second-language acquisition. However, if at all focusing on bilingual associations, former studies (e.g. Fitzpatrick, 2006; Sheng et al., 2006; Verhallen & Schoonen, 1993) have often mainly classified their investigations according to these linguistic terms, instead of qualitatively analysing per target word what the differences between these associations are.

Verspoor (2008) did look at differences in associations qualitatively. Taking a usage-based approach, Verspoor looked at the differences between Dutch learners of English and English native speakers in their associations of the word *abandon* and differences between Dutch and Vietnamese learners of English and native speakers of English in the word *career*. Clear differences were found in the associations between native speakers of English and learners of English. For example, the Dutch learners of English had a non-native-like association of *banish* with the word *abandon*, while they did not have the native association of *ending* and of *giving up*.

For her Master thesis, Van den Berg (2006) conducted a similar word association study with three groups of advanced L2 English speakers: native speakers of Dutch, German, and French. In the six stimuli words that were used for this study, differences were not only found for the three groups with the native English control group, but also for the three groups and the native Dutch, German, and French control groups (who received word association tests with the stimuli given in Dutch, German and French respectively). The concepts of *house*, *career*, *time*, *abandon*, *know*, and *drive* could therefore be concluded to be different in the L1 languages and in English.

So far, it has become clear that the conceptual representation of lexical items is dynamic and susceptible to changes. Representation differs per individual, but shows overlap between different culture and/or language groups. Moreover, word association research (see e.g. Verspoor, 2008; Van den Berg, 2006) has shown that there are differences between different languages. This would mean that the conceptual representation of L2 learners is influenced by the concepts of their L1. However, reaction time research into these influences has not been backed up by word association research. If the two were combined, it would be possible to detect the influence of L1 conceptual representation on the L2 in quantitative terms.

Thus, word associations should be able to show a difference between L1 and L2 speakers of a language, and predict some differences in conceptual information. If these differences in word associations are subsequently used for a reaction time experiment that relies on subconscious information processing of participants, different stages of the language learning process can be compared to find out whether there is an increase in L2 lexical concepts and whether there are remaining L1 concepts present in the L2. The present study will apply this in a lexical decision priming task with advanced learners of English and compare them to native speakers of English. Expected is that learners will show a slight increase in recognizing conceptual relations that are native-like, which will yield faster reaction times to these specific primes when they are more advanced. Although this increase in native-like conceptual representation in the L2 is expected, the difference between the high proficiency L2 learners in our experiment and the control data of the native speakers is still expected to be significant.

5. The experiment

Qualitative research into L2 conceptual representations as carried out by Van den Berg (2006) has tentatively shown that L2 lexical items may have many L1 related conceptual

characteristics. The deviations from L1 concepts cannot be quantitatively tested using association tasks due to individual variation. At the same time, cross-language priming studies have thus far yielded contradictory results, partly due to methodological issues (see Fitzpatrick, 2007; Pavlenko, 1999). Therefore we have conducted a two-stage study,[1] in which word associations given by native speakers and advanced L2 learners are used as primes in a monolingual lexical decision task. This method enables us to compare response times of learners to native speaker associations and vice versa. Using response time measurements we could investigate subtle conceptual differences that remain undisclosed in regular association tasks. We expected that even very advanced Dutch learners of English would still show faster reaction times to a Dutch-based association of an English word than to a native English association. Conversely, the native speakers were expected to respond faster to native associations than to the Dutch-based associations. Neutral associations were added to act as baseline RTs for both groups.

5.1 Association task

An English association task and a Dutch association task were conducted and compared to native speaker associations from the Edinburgh Word Association Thesaurus.

Van den Berg (2006) pointed out that the largest differences in associations between the two languages were to be expected with abstract words. Therefore, 100 abstract verbs and nouns were selected for a translation task carried out by 24 Dutch students of English at the University of Groningen. The 77 most frequent translations were used for the subsequent association task.

Dutch-based word associations of the 77 English (non-cognate) abstract nouns and verbs were gathered from 34 first-year and second-year students of English at the University of Groningen. Dutch associations of the same words (in translation) were gathered from 47 first-year students of Dutch. Native English associations were taken from the Edinburgh Word Association Thesaurus (EAT), which was provided by speakers of the same age group as the Dutch students in our experiment.

5.2 Priming task

Using the words from the association tasks, a priming experiment was set up with the type of association (*prime type*) as the major within-subjects factor. Additional variables were target word length, target syllable length, target frequency and word type. The additional variables were included for reference and control and were not intended to play a part in the actual analysis of the data.

15 third-year students of English at the University of Groningen and 16 native speakers of English participated in this experiment. The age range (20–40) and backgrounds of the two groups of participants were similar, although the native speakers were slightly older on average. None of the native speakers spoke any Dutch, which warrants a maximum difference in the conceptual representation investigated through the associations of the two language groups. Although the majority (12) of the native speakers were British,

some were from other English speaking countries, like Australia, New Zealand, Canada and the USA.

For the priming task 60 target word associations were selected. The neutral associations were randomly selected from the British National Corpus out of all items with a frequency between 30 and 150. An equal number of 60 pseudo words were taken from the ARC nonword database (Rastle, Harrington, & Colthaert, 2002). The items were equally distributed over three versions of the experiment, each containing all target words appearing in different conditions: 20 target words with an English prime (hereafter ENprime), 20 target words with a Dutch prime (hereafter NLprime) and 20 target words with a neutral prime. Moreover, all target associations selected were tagged for a particular association type, classified into Antonym, Synonym, Compound, Collocation and Classification.

Using *Eprime*, the items were presented to the participants in two randomized blocks of 60 stimuli, with a 30-second break in between. Before the presentation of the prime a fixation point was displayed on the center of the screen for 1000 ms. The prime was presented with an SOA of 50 ms, followed by the target item. The participants were instructed to respond to the target word as quickly as possible. No feedback was provided during the experiment.

The analysis of the resulting correct responses did not include the responses to the pseudo-items. Mean overall response times were around 500 ms. Prior to analysis, outliers were removed from the data in two steps. First all response times slower than 1200 ms were removed. We subsequently removed all responses that deviated beyond 2.5 SD of the participant's mean response time. The data were analysed in SPSS using a 2x3 repeated measurements design (learners and natives x three priming conditions).

6. Results

Q-Q plots of the results per group per prime type yielded normally distributed plots. The standard deviations of these were all between 57 and 78 ms, which made it acceptable to use a repeated measures analysis for these data. The descriptive statistics of the conditions are listed in Table 1.

Table 1. Mean reaction times, standard deviations and accuracy per group per prime type

	Learners			Natives		
	Mean RT	**SD**	**accuracy**	**Mean RT**	**SD**	**accuracy**
EN prime	508	75	98%	486	58	99%
NL prime	503	65	100%	497	68	99%
Neutral prime	518	79	96%	498	68	98%

The overall analysis of the interaction between prime type and group appeared not to be significant, and none of the interactions reached significance at $p < 0.05$.

The classifications of the ENprime and the target word show different results for the two groups of participants. The learners reacted fastest to compound relations and the

slowest to antonyms, while the natives reacted fastest to antonyms and slowest to compound relations. Moreover, the learners were faster than the natives for compound relations of ENprimes. Due to the limited number of items for each classification, these differences were not significant.

When looking at the frequency of the target words, the high frequent words yield a picture with more differences between the groups, while the low frequent words bring them closer together. A repeated measures analysis for ENprime and NLprime in the high frequency data gives a significance of the interaction between prime and group of $F(1,29) = 4{,}653$, $p = .039$. This is graphically represented in Figure 4.

The items analysis shows quite some words that have had an opposite result in that the group of learners were faster at the ENprime than at the NLprime and for the group of natives the other way around or that learners were faster at the ENprime than the natives were.

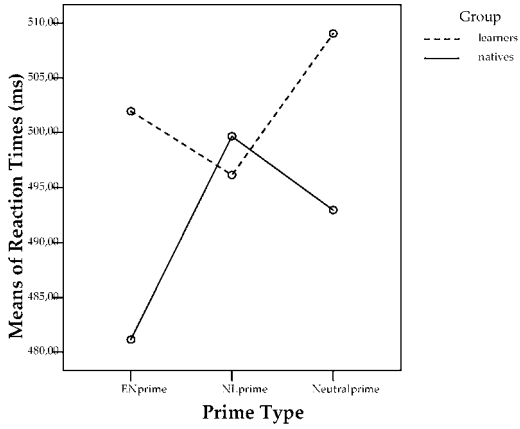


Figure 4. Means plot of prime types per group for high frequent words only

7. Discussion

The primary assumption in our study was that conceptual representations of native speakers and learners of English would show a large degree of overlap, but would not be identical. Using word association tasks we have tried to identify the differences between the conceptual representations of the learners and the native speakers. Testing the associations in a priming experiment we expected that even very advanced Dutch learners of English would still show faster response times to a Dutch-based association to an English word than to a native English association. The native speakers of English were expected to respond faster to native associations than to the Dutch-based associations.

The overall tendency found in the data, as shown in Figure 4, seems to meet these expectations. The native speakers were faster in the condition with the English primes, while the learners were faster in the condition with the Dutch primes. This effect did

not reach significance when all items were included, but it was significant for the high-frequency items. This tentatively shows that even though the conceptual representations of lexical items will largely overlap, especially in closely related languages like Dutch and English, the subtle differences that appeared from the association data were corroborated by the response time experiment. Given the inherently temporal and individual nature of lexical associations and the relative instability of the mental lexicon (compare e.g. the results found in De Bot and Lowie, this volume), especially for L2 learners, the small effect detected is certainly meaningful.

Having said that, the experiment could be improved at several points. One of the outcomes of the experiment was that there was a lot of variation between the individual items. Even though the overall tendency was in line with the effect expected, many items showed a pattern that was diametrically opposed to the trend found. In some cases this was due to irregularities in the setup of the experiment, like typing errors. In other cases the intended prime was typically British and could not have affected the non-British native speakers. The data contain several different classification types of association, but the types were not equally represented, leading to a misbalance in the experiment that could not be consistently neutralized in the analysis. Some classification types yielded reversed results due to the direction of the relation between the words. For instance, one of the associations used for *poverty* was *stricken*. In the priming experiment, it could only be tested if *stricken* primes *poverty*, while the direction of the association was reversed. Similar directionality effects occurred with *gracious* as a prime of *goodness*, *luck* as a prime for *bad*, etc. The analyses further showed that nouns yielded more differences between the groups than verbs, which is in agreement with cross-language priming effects found in the literature discussed (Van Hell & De Groot, 1998). In retrospect, it would have been better to limit the items to one word type (nouns) only. Finally, even though only relatively frequent lexical items were selected for the priming experiment, we found the strongest effect for the most frequent items. Frequency effects are well attested in the literature on lexical decision and apparently the bandwidth used was not sufficiently narrow. Moreover, since the experiment was carried out with learners, there may have been insufficient familiarity with less frequent items from the association task. This frequency issue is also discussed in Seilhamer (this volume).

8. Conclusion

In this study we have advocated a usage-based dynamic distributed activation model of the multilingual mental lexicon. Using converging evidence from crosslinguistic response time measurements and insights from cognitive linguistics, we have argued that the dynamic distributed model has the strongest explanatory power for data from these fields. Not only can it account for the relative instability of lexical representations in the mental lexicon, it can also account for the development in terms of lexical restructuring processes.

Perhaps the most valuable asset of the model is its ability to account for the partial overlap between lexical representations of learners and native speakers. The focus of this study was on the overlap of conceptual representation between languages and the extent

to which the L2 concepts of advanced learners are affected by their L1 concepts. This question was investigated in a sequence of two experiments, word association and lexical priming. The tentative outcome of this study was that even advanced learners are still affected by their L1 conceptualisation. This outcome makes sense from a usage based approach, as the entrenchment of L1 concepts cannot be erased. The partial acquisition of L2 concepts can be accounted for by the dynamic distributed activation model. The conclusions from this study will have to be tentative, as the nature of this type of experimentation is complex and the two stages had to be closely interlinked. Some methodological issues call for a refinement of the approach. Especially the categorization type of the items, the part of speech and the frequency of the targets need further attention in follow-up experiments. These issues will be resolved in further studies. For the current experiment, the conclusion seems to be justified that the experiment has tentatively demonstrated the subtly differential nature of lexical representations of advanced learners and native speakers.

Notes

1. We would like to express our gratitude to Bregtje Seton, Sabrina Eikens and Vanessa Pietersz, who carried out the first version of this study in partial completion of their MA in English or Applied Linguistics.

References

The Edinburgh Word Association Thesaurus (2007). URL [On-line]. Available: http://eat.rl.ac.uk

De Bot, Kees & Wander Lowie (this volume). On the stability of representations in the multilingual lexicon.

De Bot, Kees, Wander Lowie, & Marjolijn Verspoor (2007). A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and Cognition, 10.1*, 7–21.

Dijkstra, Ton & Walter J. B. van Heuven (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (189–225). Manwah, NJ: Lawrence Erlbaum.

Dijkstra, Ton & Walter J. B. van Heuven (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175–197.

Dong, Yanping, Shichun Gui, & Brian MacWhinney (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition, 8,* 221–238.

Entwistle, Doris R. (1966). *Word Associations of Young Children*. Baltimore: John Hopkins Press.

Fitzpatrick, Tess (2006). Habits and rabbits: Word association and the L2 lexicon. *EUROSLA Yearbook, 6,* 121–145.

Fitzpatrick, Tess (2007). Word association patterns: Unpacking the assumptions. *International Journal of Applied Linguistics, 17,* 319–331.

Hummel, John E. & Keith J. Holyoak (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110,* 220–264.

Jacquet, Maud & Robert M. French (2002). The BIA++: Extending the BIA+ to a dynamical distributed connectionist framework. *Bilingualism: Language and Cognition, 5,* 202–205.

Kawamoto, Alan H. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language, 32,* 474–516.

Kolers, Paul A. (1963). Interlingual word association. *Journal of Verbal Learning and Verbal Behaviour, 2,* 291–300.

Kroll, Judith F. & Ton Dijkstra (2002). The bilingual lexicon. In R. Kaplan (Ed.), *The Oxford Handbook of Applied Linguistics* (301–321). Oxford: Oxford University Press.

Kroll, Judith F. & Erika Stewart (1994). Category interference in translation and picture naming – Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149–174.

Langacker, Ronald W. (2000). A dynamic Usage Based Model. In M. Barlow & S. Kemmer (Eds.) *Usage Based Models of language* (1–63). Chicago: University of Chicago Press.

Lowie, Wander M. (1998). *The Acquisition of Interlanguage Morphology: A Study into the Role of Morphology in the L2 Learner's Mental Lexicon.* Ph. D. University of Groningen.

McRae, Ken, Virginia R. de Sa, & Mark S. Seidenberg (1997). On the nature and score of featural representations of word meaning. *Journal of Experimental Psychology: General, 126,* 99–130.

Pavlenko, Aneta (1999). New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition, 2,* 209–230.

Rastle, Kathleen, Jonathan M. Harrington, & Max Colthaert (2002). 358,534 nonwords: the ARC Nonword Database. *Quarterly Journal of Experimental Psychology, 55A,* 1339–1362.

Seilhamer, Mark Fifer (this volume). Concept stretching and model merging: An attempt to better account for L2 processing and acquisition of grammatical constructions.

Sharifian, Farzad (2008). Distributed, emergent cultural cognition, conceptualisation, and language. In R. M. Frank, R. Dirven, T. Zienke, & E. Bernárdez (Eds.), *Body, language, and mind (Vol. 2): Sociocultural situatedness* (109–136). Berlin/New York: Mouton de Gruyter.

Taylor, Insup (1976). Similarity between French and English words: A factor to be considered in bilingual language behaviour? *Journal of Psycholinguistic Research, 5,* 85–94.

Van den Berg, Saskia W. (2006). How much huis, Haus, or maison do you find in house? Master thesis at the University of Groningen.

Van Hell, Janet G. & Annette M. B. de Groot (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1,* 193–211.

Verspoor, Marjolijn (2008). What bilingual word associations can tell us. In F. Boers & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to Teaching vocabulary and phraseology* (259–287). Berlin: Mouton de Gruyter.

CHAPTER 9

Formulaic language in L2

Storage, retrieval and production of idioms by second language learners

Anna Cieślicka
Texas A&M International University, Laredo, TX

1. Introduction

Understanding and producing language involves not only dealing with individual words, but also with a large number of formulaic utterances, or fixed expressions, such as collocations, phrasal verbs, proverbs, idioms, slogans, common quotations, sayings, etc. The purpose of the present paper is to address the question of how second language learners process a particular variant of formulaic language, namely, idioms. Idioms are traditionally defined as multi-word phraseological units, whose meaning is not predictable from their constituent parts (see, for example, Fraser, 1970; Makkai, 1972; Weinreich, 1969). One of the issues discussed in the psycholinguistic literature devoted to idiom processing has been the degree to which literal meanings of idiom constituent words become activated in the course of retrieving the idiom's figurative meaning from the mental lexicon. The most recent psycholinguistic models seem to agree that some degree of literal activation is obligatory in the course of idiom processing. Whereas the activation of literal meanings of idiom constituents seems well-documented in the idiom comprehension literature, much less has been written about constituent lexical access during the production of idiomatic phrases, and psycholinguistic research into how second/foreign language users produce idiomatic language is even scanter. This is surprising, given the fact that direct pedagogical implications might stem from the discovery that second language learners obligatorily activate literal meanings of second/foreign (henceforth L2) idiomatic phrases. Sprenger (2003) has proposed the superlemma theory of idiom production, a hybrid model which assumes both a unitary and a compositional nature of formulaic expressions. The model has been confirmed with monolingual speakers of Dutch (Sprenger, Levelt, & Kempen, 2006).

The aim of the study described here was to test the assumptions of the superlemma theory in a foreign language learning context. Following the review of the major assumptions of the superlemma theory and experimental work into L2 idiom processing, this paper describes an experiment modeled after Experiment 3 in Sprenger, Levelt, and Kempen (2006), but conducted with Polish advanced learners of English.

2. Compositional and noncompositional views of idioms

Two opposing views on idioms have been proposed by researchers interested in psycholinguistic aspects of idiom storage and processing. The first is the noncompositional view which treats idiomatic phrases as arbitrary strings whose figurative meanings are not directly related to the literal meanings of their individual words (see, for example, Chomsky, 1980; Fraser, 1970; Katz, 1973). In psycholinguistics, the most well-known noncompositional models have been developed by Bobrow and Bell (the *idiom list* hypothesis) (1973), Swinney and Cutler (the *lexical representation* hypothesis) (1979), and Gibbs (the *direct access* model) (1980, 1985). While these models differ with regard to the proposed temporal sequence of literal and figurative meaning activation and the storage format of idioms, all noncompositional models share the assumption that idiom meanings are stipulated arbitrarily and understood by retrieving the meaning of an idiomatic phrase as a whole, rather than by processing its component parts. Because of this feature of noncompositional models, Glucksberg (1993) refers to them as *direct look-up* models.

In contrast, compositional theories propose that idioms vary with respect to their compositionality, that is, the degree to which the literal meanings of their constituent words contribute to their overall figurative interpretation. Several lines of research have convincingly shown that idiom processing cannot be exclusively reduced to the holistic retrieval of a lexicalized meaning, and that it involves an obligatory semantic and syntactic analysis of its constituent words (see, for example, Cacciari & Tabossi, 1988; Glucksberg, 1993; Tabossi & Zardon, 1995; Titone & Connine, 1994). Most of the research undertaken within the compositional framework has thus attempted to investigate the varying degree to which literal meanings of idiom constituents contribute to their overall figurative interpretation. Major compositional models of idiom storage and comprehension developed in the literature have been the *idiom decomposition* model (Gibbs & Nayak, 1989; Gibbs, Nayak, & Cutting, 1989), the *configuration* model (Cacciari & Tabossi, 1988), and the *phrase-induced polysemy* (PIP) model (Glucksberg, 1993, 2001). These models view the processing of idioms as essentially similar to literal language comprehension. Under a compositional account, individual meanings of idiom components play a significant role in constructing the idiom's overall figurative interpretation.

An important issue within the compositional framework has been the distinction between nondecomposable and decomposable idioms. Nondecomposable (or noncompositional) idioms have meanings which are arbitrary and unrelated to their compositional analysis (e.g., *kick the bucket*), whereas the figurative meanings of decomposable (compositional) idioms are highly related to the literal meanings of their constituent words (e.g., *play with fire*). Gibbs and Nayak (1989) have suggested that nondecomposable idioms might be more lexicalized and hence their meanings might be more easily retrievable from lexical storage, whereas the meanings of decomposable idioms might be computed analogically to the way in which literal language is analyzed, with each component contributing to the developing interpretation. Gibbs and Nayak hypothesized that this could slow down the processing of decomposable idioms in comparison to nondecomposable ones.

In between the strictly noncompositional and compositional views on idiom processing lie hybrid approaches, such as the *hybrid* model developed by Titone and Connine

(1999), which views individual idioms as simultaneously arbitrary and compositional. The arbitrary, noncompositional nature of idioms is reflected in their conventionality and the fact that they are highly automatized multi-word phrases directly retrievable from the mental lexicon. Idioms also behave compositionally, in that some of them are transparent and decomposable. Such an approach to idioms allows accounting for the apparent inconsistency in idiom processing studies which have shown either faster retrieval of idiomatic than literal phrases, suggesting that idioms are stored as holistic units in the mental lexicon, or literal activation of idiom component words, suggesting that idioms are stored and processed very much like literal utterances (see Titone & Connine, 1999 for a comprehensive review). Hybrid accounts of idiom comprehension thus seem to offer the best solution to the problem that any theory of idiom processing necessarily encounters, namely the simultaneously compositional and noncompositional nature of idiomatic expressions.

3. Idiom production

The first attempt to address questions concerning the storage and retrieval of idiomatic phrases during language production was Cutting and Bock's (1997) series of error-elicitation studies in which participants were briefly presented with two idioms and, after a short pause, cued to produce one or the other as quickly as possible. The dependent measures were production latencies and blending errors, that is, unconventional combinations of two idioms. In the first experiment, Cutting and Bock (1997) employed pairs of idioms with similar (*kick the bucket, meet the maker*) or different (*shoot the breeze, raise the roof*) idiomatic meanings and with the same (*chew the fat, raise the roof*) or different (*chew the fat, nip and tuck*) syntactic forms. If idioms are stored as unitary forms then their syntactic structure should have no effect on the resulting idioms blends. On the other hand, if idiomatic expressions do undergo syntactic analysis in the course of their production, then idioms with a similar structure should produce more blends than those with differing structures. The experiment showed that same-syntax idioms with similar figurative meanings were more likely to blend than different-syntax idioms with different figurative meanings, which was taken by the authors as an argument against a lexicalized view of idiom storage.

In Experiment 3 Cutting and Bock investigated the differential lexicalization status of decomposable and nondecomposable idioms hypothesized by Gibbs and Nayak (1989). They thus presented participants with pairs of idioms that are either decomposable (*hold your tongue, button your lip*) or nondecomposable (*shoot the breeze, chew the fat*). The reasoning was that if lexical representation of nondecomposable idioms is more unitary, such idioms should be less susceptible to the production of idiom blends (e.g., *shoot the fat*) in the error elicitation task than decomposable idioms, in line with the idiom decomposition model. On the other hand, decomposable idioms, with individual components mapping directly onto the idiomatic senses, should elicit a substantially bigger number of idiom blends (e.g., *hold your lip*). Analysis of participants' responses revealed that both decomposable and nondecomposable idioms elicited a comparable proportion of idiom blends, which, according to the authors, suggests that the lexical representations of both idiom types are identical, especially when it comes to the production process.

Accordingly, Cutting and Bock (1997) have proposed a hybrid model of idiom production whose architecture is compatible with models of language production proposed in the psycholinguistic literature (Dell, 1986; Levelt, 1989, MacWhinney, 2008). The hybrid model presupposes that idioms are stored as whole units at the lexical-conceptual level of the lexicon. The lexical-concept nodes representing them are connected to the syntactic component of the system in such a way that they retrieve the phrasal frames specifying the ordering of grammatical slots in idiom phrases. In addition, the lexical-conceptual node representing an idiom is connected in the lexicon with lemmas for individual words that enter into the idiomatic phrase (the term *lemma* referring to representation of a lexical item's grammatical class information plus a pointer to the word forms (see Levelt & Meyer, 2000; Roelofs, 1992). The model easily accounts for an increase in a number of blending errors as a function of structural and meaning similarity which was demonstrated in Experiment 1. Idioms with the same syntactic form share the same syntactic frames, and idioms with similar meanings activate similar conceptual representations, which results in more competition than in the case of syntactically or semantically dissimilar idioms. Since one concept can activate multiple lexical concept nodes, including those representing idioms, similar meaning or similar structure idioms, such as *meet your maker* and *kick the bucket*, will become activated simultaneously and will compete in the course of language production, which might lead to blend errors such as *meet the bucket.*

In order to test the hybrid account of idiom production, Sprenger, Levelt and Kempen (2006) investigated production of idioms in a series of studies employing reaction time paradigms. In their first experiment, Sprenger et al. (2006) tested the claim that idiom constituents are the same lemmas which get activated in the course of producing literal phrases and that idiom phrases have their unitary representations in the mental lexicon. They thus employed a cued-recall task, in which native speakers of Dutch produced either idiomatic or literal phrases they had learned earlier in response to a visually displayed prompt word. While looking at the prompt word, participants heard a prime, which was either a word identical to the noun of the phrase to be produced or a word unrelated to any of the phrase's elements. The results showed that both idiomatic and literal phrases were produced faster when primed by one of their content words. This result, according to the authors, supports the view that idioms are compositional phrases made up of the same simple lemmas that are activated in the course of literal language production. In addition, the comparison of the priming effects found for idiomatic and literal phrases revealed that idioms were primed significantly stronger than literal utterances. Sprenger et al. (2006) explain this difference in the strength of priming by postulating a separate meaning representation for idioms in the mental lexicon. Since lemmas belonging to an idiomatic phrase are subsumed under a unitary lexical entry, priming one of the lemmas belonging to this common idiomatic representation results in the spreading of activation to the remaining lemmas of the idiom entry, thus making them more available for retrieval. Since no such common representation exists for literal phrases, priming can only speed up the retrieval of the semantically related lexical item but it cannot influence the activation of the remaining components of the literal utterance. These conclusions, confirmed in the remaining experiments, unequivocally support the hybrid account of idiom representation, under which idioms are both unitary and compositional phrases.

In a post hoc analysis of the results obtained in all the three experiments, Sprenger et al. (2006) found that idiom decomposability did not influence the extent to which literal meanings of idiom elements became activated in the course of idiom production. The results showed some sensitivity of speech onset latencies to idiom decomposability in Experiments 2 and 3, in which idioms rated as highly decomposable evoked longer speech onset latencies. However, the effect was inconsistent, as they also found the reversed effect of shorter speech onset latencies for decomposable idioms in the data. This was in contrast to the results of Cutting and Bock (1997), who failed to find any effect of idiom decomposability on the percentage of error blends. In view of the inconsistencies of the effects of idiom decomposability demonstrated in their study, Sprenger et al. do not speculate on the origin of these effects.

All in all, the study conducted by Sprenger et al. (2006) corroborates the hybrid account of idiom production put forward by Cutting and Bock (1997). In an effort to make the hybrid model applicable to production and comprehension, Sprenger et al. propose an extension and modification of the original hybrid model, which they call the *superlemma* theory. A superlemma is a separate holistic representation of an idiom, introduced into the hybrid model at the lexical-syntactic processing level in order to account for the evidence that the syntactic properties of an idiom are in some way independently represented. This superlemma is linked with the individual lemmas which make up the idiomatic phrase. One of the advantages of Sprenger et al.'s model is that it posits a simplified mechanism of idiom production, suggesting that it is identical to the mechanism of processing single words. This is so because idioms are represented with their own superlemma. This superlemma (for example, *skate on thin ice*) gets activated along with other word and phrase lemmas (such as *risk* or *gamble*) in the course of language production and competes with them in the selection process, so that, in accordance with Luce's (1959) ratio proposed in the WEAVER model of language production (see Roelofs, 1997), the superlemma for a given idiom will only be selected if the activation of its node exceeds the summed activation of all the other lemmas and superlemmas in the lexical network system. Then, upon selection of the superlemma, the simple lemmas that it points to (such as *skate, thin*, and *ice*) are selected as well (Sprenger et al., 2006).

Another advantage of the superlemma theory over the hybrid model is, according to Sprenger et al. (2006), the ease with which it can account for the differing syntactic flexibility of various idiomatic expressions. Since syntactic information about idiomatic expressions is stored in the superlemma, all the constraints operating on a particular idiom, as well as the actual grammatical relations holding between its constituents, are coded at the superlemma level. Such a solution is much simpler than the phrasal frames with open slots proposed in the hybrid model, which cannot straightforwardly account for syntactic idiosyncrasies of idioms.

4. Idiom processing in L2

The abundance of studies concerning idiomatic expressions might imply that all areas of idiom storage, comprehension, and production have been exhaustively examined. There

is, however, an aspect of idiom research which, regrettably, has been ignored in scholarly investigations- the processing of idiomatic expressions by L2 learners. The scarcity of studies of the processing of idioms and other forms of formulaic language by L2 learners is even more intriguing if one considers the widely held view that idioms constitute one of the most problematic areas of foreign language learning (see, for example, Alexander, 1987; Boers, 2000; Cacciari, 1993; Charteris-Black, 2002; Fernando, 1996; Howarth, 1998; Irujo, 1986, 1993; Kovecses & Szabo, 1996; Littlemore, 2001; Lattey, 1986; Moon, 1997; Yorio, 1989; Zughoul, 1991). In line with this view, it has been shown that figurative competence in L2 lags behind native language figurative competence (Danesi, 1992) or may even be of a qualitatively different nature, as the study into the comprehension of L1 and L2 conversational implicatures carried out by Bromberek-Dyzman and Ewert (this volume) seems to imply.

Few idiom scholars have attempted to present a theoretical account of L2 idiom learning. Gibbs (1995) has suggested that second language learners learn idioms in a rote manner, by establishing arbitrary links between idiom forms and their figurative meanings. An essentially similar view of L2 idioms has been expressed by Matlock and Heredia (2002), who assume that non-experienced (beginner) second language learners must establish direct connections between literal and nonliteral meanings of figurative expressions. When processing an L2 idiom, beginner L2 learners first attempt to make sense of it by translating it literally into L1. Only then is the figurative meaning accessed. With increasing L2 proficiency though, foreign language learners become capable of processing figurative expressions directly, in the same manner as native monolingual users. Liontas (2002) has argued that comprehension of L2 idioms requires a special processing mode, a claim reminiscent of Bobrow and Bell's (1973) idiom list hypothesis. According to Liontas, L2 learners process literal and idiomatic senses separately, deriving two alternative interpretations of an idiomatic phrase. He further claims that L2 idiom understanding comprises an initial 'prediction phrase' and a 'confirmation or replacement, reconstructive phase'. In the prediction phase, in the absence of context, an L2 learner relies on the literal analysis of idiom constituents, constructing a number of hypotheses and predictions. In the second, confirmation or replacement, reconstructive phase, "interpretation of idioms is restricted to its own context through the gradual elimination of possible interpretations" (2002: 182). Analysis of the information available in the input leads either to the confirmation of the earlier hypotheses or their replacement in light of the contextual constraints. L2 learners' inferences will depend on their semantic, syntactic, pragmatic, and cultural knowledge. Regardless of these factors, accessing literal meanings of idiom constituent words is an obligatory step in idiom processing, even if the idiom is highly familiar. It emerges then from Liontas' proposal that literal meanings of idioms enjoy a special status in the course of processing idioms by L2 learners.

A similar claim, concerning the special status of literal meanings in the course of L2 figurative processing has been put forward by Kecskes (2000), who has suggested that, due to the lack of metaphorical competence in L2, foreign language learners are more likely to rely on literal meanings of figurative utterances and on their L1 conceptual system when producing and comprehending formulaic expressions. A special status of literal meaning analysis in the course of L2 idiom processing is also apparent in Abel's (2003) *Model of*

Dual Idiom Representation. In a decomposability rating study conducted with German speakers of English, she showed that L2 learners tend to rate nondecomposable idioms as decomposable, even if this is not correct. Participants in her study were shown to assume that literal meanings of individual constituents of nondecomposable idiomatic phrases actively contribute to the idioms' overall figurative interpretation.

The strategy of reliance on literal meanings of idiom constituents also underlies Cieślicka's (2006) *Literal Salience Resonant* (LSR) model of L2 idiom comprehension, support for which has been found in a series of on-line studies employing the cross-modal priming paradigm (Cieślicka, 2006, 2007). The LSR model of L2 idioms assumes primacy of literal over figurative meanings of idiom constituents, assigning literal meaning a higher salience status in idiom comprehension (see Giora's *graded salience* hypothesis, 1997, 1999, 2003 for a discussion of the notion of salience in language processing). Literal salience in L2 idiom processing reflects the way in which a lexical representation of an idiom entry is dynamically constructed in the course of figurative language acquisition by an L2 learner. In line with the widely accepted view of the bilingual lexicon as being a multi-layered hierarchical structure (see, for example, De Groot, 2002; Heredia, 1997; Kroll & De Groot, 1997; Kroll & Stewart, 1994; Kroll & Tokowicz, 2001), the LSR model assumes that the mental lexicon of an L2 learner consists of a language-independent conceptual level of meaning representation and language-specific lexical, or word form level, both of which are interconnected via lexical and conceptual links. Since the construction of meaning representations in the L2 lexical network is a dynamic process (for a discussion of the dynamic and unstable nature of representations in the mental lexicon, see de Bot and Lowie, this volume), creating an idiom entry will involve establishing associative links between lexical nodes for idiom constituents, so that the idiomatic configuration can be recognized as a phrase when processed. Gradually, those lexical-level connections between idiom constituents become strengthened and the whole figurative sequence becomes automatized in the lexical network. Automatization of the idiomatic sequence can only occur with sufficient exposure to a given idiomatic form in the L2 input, accompanied by extensive oral practice opportunities.

In addition to creating strong links between idiom lemmas at the lexical level, building an idiom entry entails developing a conceptual level representation of the idiom's figurative meaning. This conceptual representation is likely to differ from that of the idiom's native language translation equivalent and hence will pose additional difficulties for the L2 learner (see Lowie, Verspoor and Seton, this volume, for a discussion of the development of conceptual representations in the L2 mental lexicon). The LSR model further assumes that the newly created conceptual links for L2 figurative meanings will be much weaker than conceptual links holding between idiom constituents and their literal meaning representations. This is so because figurative expressions in L2 are most often learned by L2 learners who are already familiar with literal meanings of L2 words making them up. Since conceptual links are likely to be much stronger for literal than figurative meanings of idioms, literal meanings of idiom components will be the most strongly and fastest activated in the course of L2 idiom comprehension.

It thus appears from the review of on-line research into idiom processing in L2 that literal meanings might be activated to a considerable degree in the course of processing

idiomatic phrases by second language learners. The superlemma theory proposed by Sprenger et al. (2006), with its emphasis on literal meaning activation, should thus adequately capture the literal-meaning priority in L2 idiom processing. The study reported in the remainder of this paper was designed to test this assumption with regard to idiom production by second language learners.

5. The study

The aim of the study was to test the superlemma theory in the L2 learning context, more specifically, to explore the degree of activation of literal meanings of L2 idiom constituents in the course of language production. The study employed a variant of the completion and naming task used by Sprenger et al. (2006) in their third experiment with native speakers of Dutch. This task measures the so-called 'preparation effect' that arises when a language user is preparing to produce the last word of an idiomatic expression and has to switch the task and read out loud another word instead of completing the idiomatic phrase. If literal constituents of idioms become activated in the course of language production, then the language processing mechanism, which is preparing to produce the last idiom component, should co-activate other lemmas, either semantically or phonologically related with the idiom's lemma, and thus make such related words easier to name than unrelated ones. To provide an example, in the course of uttering the idiomatic expression *Jack pulled my leg*, activation spreads over the network of lexical units, co-activating those lexical items which are semantically related to the verb *pull* and to the noun *leg*. If, instead of having to complete the idiom with the last word, the language user has to name a word that is semantically related to the missing lemma (e.g., *foot*), then the time needed to read the word out loud should be much shorter than if the word was unrelated to the idiom (e.g., *peas*). Measuring speech onset latencies to words related semantically and phonologically to the last words of idiom phrases and comparing them to latencies obtained when the words are unrelated to the idiom constituents allows researchers to assess the degree of activation of literal meanings of idioms' lemmas during their production.

5.1 Participants

The participants were 30 Polish learners of English (mean age 23.4; 25 women and 5 men). All of them were students at the School of English, Adam Mickiewicz University, Poznań, enrolled in the 2-year M.A. program and fluent in English (average length of exposure to English 14.5 years). Their participation in the experiment was voluntary.

5.2 Materials

The materials included sixteen English idiomatic expressions, all of which were finite phrases in the past tense form, of the following syntactic structure: Subject [VP [V([PP

Prep) [NP art/pro/adj N]]]] (*Jack kicked the bucket, Jack sat on the fence*). They consisted of the same proper name, *Jack*, followed by the verb in the past tense, and either the determiner/ adjective plus noun (*Jack kicked the bucket*) or the preposition, determiner (optional) and noun (*Jack sat on the fence, Jack played with fire*). Out of the sixteen idioms, the majority (12) had the structure: Noun + Verb + Det./Adj. + Noun, and the remaining four had the structure: Noun + Verb + Prep. + (Det.) + Noun.

All the stimulus idioms were rated as familiar in a pilot study conducted with a group of 50 students, whose level of proficiency in English was comparable to that of the subjects taking part in the experiment. During the pilot study, the subjects were asked to rate each idiom according to its familiarity and decomposability on a 1–7 Likert scale from 'unknown'/'nondecomposable' (1) to 'well-known'/'decomposable' (7). Only the idioms whose average familiarity ratings exceeded 4.0 were selected for the experiment. Instructions for the decomposability rating required that subjects assess the extent to which the literal action or mental state denoted by the idiom is related to or similar to the figurative meaning of the idiom, with 1 indicating that the literal meaning is dissimilar from the idiomatic meaning and 7 indicating that the two meanings are very closely related. Idioms which received a decomposability rating above 3.5 were classified as decomposable, whereas those which received a rating of 3.5 or below were categorized as nondecomposable. Out of the sixteen idioms used in the study, only four were rated as decomposable. Even though the study described here was not intended to test idioms differing in decomposability, especially in light of the fact that the number of decomposable and nondecomposable idioms employed in the experiment was unequal, a follow-up analysis was run on items varying along the dimension of decomposability.

For each idiom, two words were prepared, one of which was related semantically and the other phonologically to the last word in the idiom. Semantically related primes were obtained from a word association task conducted with the same group of participants who took part in the familiarity and decomposability norming pilot study. The most frequent responses were selected as final stimuli. For example, the semantically related word for *fire* in the idiom *play with fire* was *heat*. In turn, phonologically related targets were prepared by the experimenter in such a way that they had the same leads (C(C)V_) as the idiom final word and the same number of syllables and word stress. For example, the phonologically related word for *bucket* in the idiom *kick the bucket* was *budget* (see Appendix 1 for a complete list of materials). Inclusion of phonologically related targets was motivated by the rationale provided by Sprenger et al. (2006), under which the phonological effect might function as an indicator of the experimental paradigm's sensitivity to priming in general. This is so because the phonological priming effect is independent from idiomaticity of the stimulus materials per se and reflects the preparation of the word's phonological form. Hence, if obtained, the phonological priming effect would testify to the sensitivity of the experimental task to priming, even in the absence of semantic priming. In turn, demonstrating the presence of semantic priming would provide support for the claim that the building blocks of idioms are the same lemmas which become activated when they are part of literal utterances.

5.3 Procedure

Participants were tested individually in a soundproof room. At the beginning of the session each person received a paper-and-pencil cloze task, whose purpose was to check his/her familiarity with the idioms used in the experiment. Each of the sixteen idioms was entered on a separate line, always with the last word (henceforth referred to as 'coda') missing (e.g. *Jack kicked the* _____). The participants' task was to write down the coda for each idiom in the space provided so as the experimenter could assess which idioms the participants did not know. Data pertaining to idioms unknown to each subject (5.5% of all the data points) were excluded from further analysis. The participants were subsequently seated in front of the computer screen and silently read instructions displayed on the screen for the on-line task. The instructions said that they would see incomplete idioms displayed one after another in the center of the screen with their coda (last word) missing. These would be the same idioms which they had just seen in the paper-and-pencil cloze task. Participants were to silently read each incomplete idiom displayed on the computer screen and to produce, i.e., say out loud, the missing coda when the question mark appeared below the idiom. The instructions emphasized the need to say the coda out loud into the microphone as quickly as possible and not to cough or make other noises that might trigger the voice key. The instructions further said that, on a number of occasions, instead of a question mark, a word in capital letters would appear below the idiom fragment. In such cases, participants were supposed to switch task and name the capitalized word (henceforth a 'target') out loud, instead of completing the idiom with its missing coda. Although the instructions de-emphasized the naming task, the proportion of completion and naming trials (i.e., test items) was identical. Before the experiment started, participants went through a practice block consisting of 32 trials to familiarize them with the nature of the experiment.

Each idiom was presented 16 times, half of which were completion trials, in which the participant was supposed to say the missing coda, and the other half were naming trials, in which case the task was to name the target that appeared in place of the question mark below the idiom. Naming trials included targets which were semantically or phonologically related to the last idiom word, or unrelated. Unrelated targets were taken from the pool of semantically and phonologically related targets which were next paired with different, unrelated idioms in order to provide the baseline data for these targets in the unprimed condition. Each idiom was thus presented 8 times in the naming task, twice with semantically related targets, twice with phonologically related targets, and four times with unrelated targets, two of which were semantic targets for other idioms functioning here as semantically-unrelated targets, and the remaining two were phonologically related to other idioms and so functioning here as phonologically-unrelated. All together then, there were 256 trials (i.e. sixteen idioms presented 16 times, half of which constituted completion trials and half naming trials). The experiment was divided into two blocks, each consisting of 128 trials. The order of trial presentation was randomized for each subject and the length of the rest pause between the two blocks was controlled by each participant.

The experiment made use of the E-Prime Software (Version 1.1 Psych. Tools Inc.) and the Serial Response Box (PST Deluxe model), which collected responses via a microphone connected to it. The procedure consisted of the following steps: (1) At the beginning of each trial, a fixation cross appeared in the center of the computer screen for 1000 ms. (2) The fixation cross was next replaced with an idiom fragment, which was displayed in white lowercase letters against a navy blue background. In the completion trials, a red question mark appeared below the idiom fragment. In the naming trials, a target in red capital letters appeared in the same place in which the question mark appeared in the completion trials. (3) The idiom fragment with either the question mark or the target word disappeared as soon as the voice key was triggered by the participant's response. (4) If a response was not provided within 1200 ms, the idiom fragment disappeared and the response was automatically counted as a timeout error. Response latency was thus measured from the moment the idiom fragment with the question mark or the target appeared on the screen up to the triggering of the voice key by the participant's response. (5) After the participant made a response or after 1200 ms when no response was provided, a fixation signal was displayed again to indicate the beginning of the subsequent trial. The whole session was tape-recorded, and the experimenter made notes on the errors made by the subjects during the session.

Following the rationale provided by Sprenger et al. (2006), the display time was deliberately kept short, so as to avoid strategic behavior by the participants (e.g., preparing to utter a coda in anticipation of a completion trial). However, the original procedure was slightly modified in such a way that either the question mark or the target were displayed immediately with the idiom fragment, i.e., no delay was employed. This contrasts to the procedure used by Sprenger et al. in Experiment 3, where a question mark for the completion trials or a target for the naming trials were displayed after the idiom fragment disappeared, at Stimulus Onset Asynchronies (SOAs) of 100, 200, 300, and 400 ms. By varying the delay between the presentation of the idiom and the subsequent task, Sprenger et al. wanted to explore how patterns of semantic and phonological priming will vary over time. Seeing an incomplete idiom fragment without a further clue about the upcoming task in the form of either the question mark or the target, participants in Sprenger et al.'s study were less likely to engage in strategic behavior (e.g., waiting for either the question mark or the target before starting to prepare a response). In turn, presenting an incomplete idiom fragment simultaneously with the question mark or the target, as was done in the current study, might indeed encourage strategic behavior. The reasoning is as follows. Upon seeing an idiom fragment accompanied by a question mark, participants would follow the instructions and read the incomplete idiom silently so as to be able to complete it with its missing coda. On the other hand, when seeing an idiom fragment with a target in place of the question mark they would know that the task was now naming the word, so they might prepare the naming response without engaging in careful processing of the idiom fragment. If semantic priming is still obtained for semantically related targets in the naming trials then such results would indicate an obligatory analysis of literal meanings of idiom constituents even in conditions when no deliberate effort is made to carefully process an idiomatic phrase.

5.4 Research questions

Previous research into second language idiom comprehension (see Section 4) has demonstrated that literal meanings of idiom constituents enjoy priority in the course of understanding idioms by L2 learners. This experiment was thus intended to show whether this literal priority in figurative processing holds true also for L2 idiom production. In view of the conflicting results demonstrated in the error production experiment by Cutting and Bock (1997) and Sprenger et al.'s (2006) data, the secondary goal of the experiment was to find out whether the activation of literal meanings of idiom components varies as a function of idiom decomposability. The following two research questions have been formulated:

1. Will literal meanings of idiom lemmas become activated in the course of their processing by L2 learners?
2. Will there be any difference in the patterns of semantic and phonological priming between decomposable and nondecomposable idioms?

5.5 Results

Data pertaining to those idioms which the participants marked as unknown before the experiment were removed from further analysis. Tape recordings of the sessions were analyzed, along with the notes made during the experiment, in order to identify false triggers of voice key and erroneous or missing responses. Data from three students had to be altogether removed because the percentage of errors they made in the completion trials exceeded 20%. The remaining error data were next entered into the statistical program (SPSS 13.0) and analyzed in a series of planned comparisons described below.

In order to reduce variability, data obtained for each subject in the naming and completion trials were screened to eliminate outliers. Reaction times exceeding two times the standard deviation from the subject means (per condition in the naming and in the completion trials) counted as outliers and were excluded from further analysis (3.5%). The remaining reaction time data were exported into the SPSS 13.0 for Windows.

5.5.1 *Error data*

Mean error percentages per condition are shown in Table 1.

Participants made by far the most errors in the completion task (3.38% in the subject and 12.14% in the item analysis) and the least errors in the semantically related vs. unrelated (1.16% vs. 1.39% in the subject analysis) condition in the naming task. Planned comparisons between Phon and PhonUnr and Seman and SemanUnr conditions revealed no significant differences, as shown in Table 2.

5.5.2 *Priming for related targets*

The mean reaction times obtained for each priming condition in the naming and completion tasks are provided in Table 3.

Table 1. Mean error percentages, standard deviation and standard error means for completion and naming tasks. Phon stands for items phonologically related to the idiom lemma, Seman for items related semantically, and PhonUnr/SemanUnr for unrelated targets taken from the pool of phonological and semantic primes. S and I stand for the values obtained in the subject and item analysis, respectively

	Naming task				
	Phon	**PhonUnr**	**Seman**	**SemanUnr**	**Completion task**
Mean Error Percentages	1.51% (S) 2.91% (I)	1.39% (S) 2.36% (I)	1.16% (S) 4.64% (I)	1.39% (S) 4.07% (I)	3.38 % (S) 12.14% (I)
Standard Deviation	3.81 (S) 1.46 (I)	3.28 (S) 1.20 (I)	3.14 (S) 2.39 (I)	3.28 (S) 3.04 (I)	2.65 (S) 7.79 (I)
Standard Error Mean	0.73 (S) 0.55 (I)	0.63 (S) 0.36 (I)	0.61 (S) 1.20 (I)	0.63 (S) 1.36 (I)	0.51 (S) 1.84 (I)

Table 2. t-statistics for mean error percentages obtained in the experiment.
S and I stand for the values obtained in the subject and item analysis, respectively

	Mean	**SD**	**SE**	**t**	**Sig. (2-tailed)**
Pair 1	−0.12 (S)	2.37 (S)	0.46 (S)	−0.25 (S)	0.80 (S)
PhonUnr-Phon	−0.33 (I)	0.52 (I)	0.21 (I)	−1.58 (I)	0.18 (I)
Pair 2	0.23 (S)	1.48 (S)	0.29 (S)	0.81 (S)	0.43 (S)
SemanUnr-Seman	0.33 (I)	0.58 (I)	0.33 (I)	1.00 (I)	0.42 (I)

Table 3. Mean RTs (in milliseconds), standard deviation and standard error means for completion and naming tasks. S and I stand for the values obtained in the subject and item analysis, respectively

	Naming task				**Completion task**
	Phon	**PhonUnr**	**Seman**	**SemanUnr**	
Mean RT	776 (S) 775 (I)	791 (S) 930 (I)	764 (S) 768 (I)	781 (S) 800 (I)	1086 (S) 972 (I)
Standard Deviation	108 (S) 56 (I)	110 (S) 331 (I)	107 (S) 36 (I)	100 (S) 93 (I)	120 (S) 85 (I)
Standard Error Mean	21 (S) 14 (I)	21 (S) 28 (I)	21 (S) 9 (I)	19 (S) 23 (I)	23 (S) 21 (I)

Completion trials took longer than naming trials did, and targets related phonologically and semantically to the idiom lemmas were named more quickly than unrelated targets in both the subject and item analyses. Planned comparisons of the mean reaction times obtained in the related and unrelated target conditions revealed significant effects for both phonologically and semantically related primes (see Table 4).

Table 4. t-statistics for the mean RTs obtained in the reading task for the semantically/phonologically related and unrelated conditions. S and I stand for the values obtained in the subject and item analysis, respectively.

	Mean	**SD**	**SE**	**t**	**df**	**Sig. (1-tailed)**
Pair 1						
PhonUnr-Phon	14.5 (S)	47.01 (S)	9.04 (S)	1.60 (S)	26 (S)	0.05 (S)
	155 (I)	332 (I)	59 (I)	1.08 (I)	15 (I)	0.14 (I)
Pair 2						
SemanUnr-Seman	16.5 (S)	105.15 (S)	9.65 (S)	1.71 (S)	26 (S)	0.004 (S)
	31 (I)	95.54 (I)	23.88 (I)	1.30 (I)	15 (I)	0.10 (I)

The data for the subject analysis indicate that both semantically and phonologically related targets were named significantly faster than their control, unrelated targets ($p < 0.01$; $p < 0.1$), whereas the difference in the item analysis approached statistical significance for both semantically ($p = 0.10$) and phonologically ($p = 0.14$) related targets. The analysis thus shows that the preparation of nouns which are parts of idiomatic expressions can speed up the naming of words that are semantically or phonologically related to those nouns. The obtained priming effect testifies to the validity of the superlemma theory and the claim that lexical concept nodes of the idiom lemmas become activated in the course of idiom production in L2. Figure 1 summarizes speech onset latencies for the completion and naming tasks in the related and unrelated conditions.

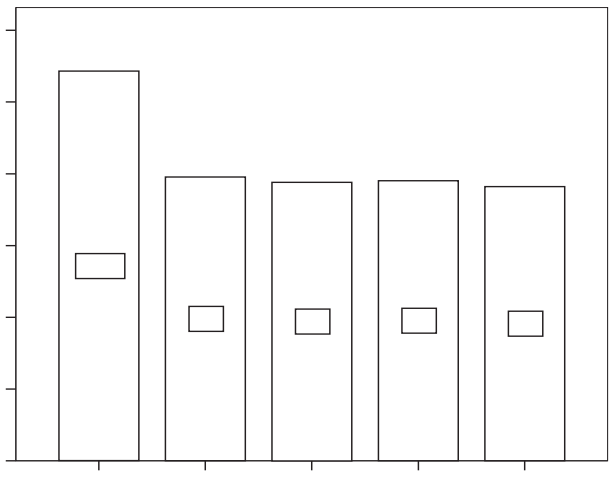


Figure 1. Speech onset latencies for completion and naming tasks obtained in Related and Unrelated conditions

5.5.3 *Decomposable vs. nondecomposable idioms*

To find out whether the results were influenced by idiom decomposability, a 2 x 5 repeated measures ANOVA was conducted on the obtained RT data, with Decomposability (Decomposable vs. Nondecomposable) and Type of Task (Completion vs. Naming Semantic vs. Naming Semantic Unrelated vs. Naming Phonological vs. Naming Phonological Unrelated) as within-subject factors. Table 5 summarizes mean RTs obtained for each Task Type as a function of idiom decomposability.

Table 5. Mean RTs (in milliseconds), standard deviation and standard error means for completion and naming tasks as a function of idiom decomposability. Decomp and Nondecom stand for decomposable and nondecomposable idioms. The figures in parentheses next to mean RTs are priming effects obtained by subtracting mean RTs elicited for related targets from RTs elicited for their control/unrelated targets.

Task/Prime	Mean RT		SD		SE Mean	
Type	Decomp	Nondecom	Decomp	Nondecom	Decomp	Nondecom
Phon	747 (54)	784 (−2)	128	108	25	21
PhonUnr	801	782	108	119	21	23
Seman	788 (−18)	755 (24)	144	108	28	21
SemanUnr	770	779	112	107	22	21
Completion	1126	1083	135	155	26	30

A significant main effect was obtained only for Task Type (F (4, 104) = 165.72; $p < .0001$). ANOVA further revealed a significant interaction between Decomposability and Task Type: (F (4, 104) = 3.50; $p < 0.05$). Phonologically related targets were produced significantly faster (747 ms) than their controls (801 ms) for decomposable (t (26) = −2.67; $p < 0.05$), but not for nondecomposable idioms: (t (26) = 0.13; $p = 0.90$). On the other hand, semantically related targets took significantly less time (755 ms) to produce than their controls (779 ms) in the case of nondecomposable idioms: (t (26) = −2.12; $p < 0.05$), but they failed to do so in the case of decomposable ones: (t (26) = 0.85; $p = 0.41$), where they actually took 18 ms longer to name than their controls. While these results do seem to show a differential sensitivity of decomposable and nondecomposable idioms to semantically and phonologically related targets, they are difficult to interpret. Whereas semantic primes appeared more efficient for nondecomposable idioms, phonological ones turned out to be relevant for decomposable idiomatic expressions. This dissociation may be the spurious effect of the uneven number of decomposable (only 4) vs. nondecomposable (12) idioms employed in the study. The fact that both idiom types primed either a word related semantically or phonologically to their last idiom lemma is, however, a strong argument in favor of the view that idioms do undergo a full literal analysis in the course of their production by second language learners, irrespective of their semantic analyzability.

6. Conclusions

The study reported here provides evidence in favor of the superlemma theory of idiom production proposed by Sprenger et al. (2006). Similarly to the findings reported by Sprenger et al. (2006) for native language speakers, the presence of a phonological priming effect in the data reported here confirmed the sensitivity of the completion/naming task to measure the preparation effect arising in the course of processing idioms by L2 learners. The effect arises when the production latency of a word that is phonologically related to the missing idiom coda is significantly affected, i.e., when this word is named faster than the word unrelated in any way to the idiom coda. In turn, the presence of a significant semantic priming effect can be seen as evidence for the activation of literal meanings of idiom lemmas during their on-line processing. Naming latencies for targets related semantically and phonologically to the idiom coda were significantly shorter than the latencies obtained for unrelated targets. This result supports the view that literal analysis of an idiomatic string is obligatory in L2 idiom production, especially in light of the fact that the design of the experiment might allow participants to ignore the idiom and focus on the naming task instead. As mentioned earlier, while subjects in Experiment 3 by Sprenger et al. saw a question mark or a target to name after the unfinished idiom disappeared from the screen, the design of the study reported here was modified in that the display of the unfinished idiom was simultaneous with the appearance of the question mark (suggesting the need to complete the idiom with its coda) or the target (to be named). Upon seeing the idiom accompanied by a target, instead of a question mark, the subjects could thus focus on the naming task, without attempting to hold the idiom in their working memory for later completion. The fact that even under such circumstances a substantial priming effect was found for related targets implies that a literal analysis of idioms did take place and that it is an automatic aspect of L2 idiom production. Those results are in line with the special status assigned to literal meanings of formulaic phrases in the course of L2 processing and reported in the L2 idiom comprehension literature reviewed earlier in the paper.

As regards the influence of decomposability on idiom production, no definitive conclusions can be drawn from the study, especially since the statistical validity of the results reported here is doubtful, on account of the uneven proportion of both idiom types employed in the experiment and the small number for the nondecomposable ones. The study found significant priming for phonologically, but not for semantically related targets displayed with decomposable idioms, and significant priming for semantically, but not for phonologically related targets displayed with nondecomposable idioms. This result is quite difficult to interpret, but it definitely undermines the claim that nondecomposable idioms are completely lexicalized in storage and only ever processed as single words. Were nondecomposable idioms to be stored and retrieved as only ever as single words, no semantic analysis of their constituent lemmas would have been demonstrated in the naming latencies. Since, however, this dimension of idiom variability was not properly controlled in the experiment, it seems that further research is needed to find out how exactly idiom decomposability affects their processing by L2 learners in the course of language production.

References

Abel, Beate (2003). English idioms in the first language and second language lexicon: A dual representation approach. *Second Language Research*, *19,* 329–358.

Alexander, Richard J. (1987). Problems in understanding and teaching idiomaticity in English. *Anglistik und Englischunterricht, 32,* 105–122.

Bobrow, Samuel A. & Susan M. Bell (1973). On catching on to idiomatic expressions. *Memory and Cognition, 1*(3), 342–346.

Boers, Frank (2000). Metaphor awareness and vocabulary retention. *Applied Linguistics, 21* (4), 553–571.

Cacciari, Cristina (1993). The place of idioms in literal and metaphorical world. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (27–56). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cacciari, Cristina & Sam Glucksberg (1990). Understanding idiomatic expressions: The contribution of word meanings. In G. B. Simpson (Ed.), *Understanding word and sentence* (217–240). Amsterdam: Elsevier.

Cacciari, Cristina & Patrizia Tabossi (1988). The comprehension of idioms. *Journal of Memory and Language, 27,* 668–683.

Charteris-Black, Jonathan (2002). Second language figurative proficiency: A comparative study of Malay and English. *Applied Linguistics, 23* (1), 104–133.

Boers, Frank (2000). Metaphor awareness and vocabulary retention. *Applied Linguistics, 21* (4), 553–571.

Chen, Hsuan-Chih & Yuen-Sum Leung (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15* (2), 316–325.

Chen, Hsuan-Chih & Man-Lai Ng (1989). Semantic facilitation and translation priming effects in Chinese-English bilinguals. *Memory and Cognition, 17* (4), 454–462.

Chomsky, Noam (1980). *Rules and representations.* New York: Columbia University Press.

Cieślicka, Anna (2006). Literal salience in on-line processing of idiomatic expressions by L2 speakers. *Second Language Research, 22* (2), 115–144.

Cieślicka, Anna (2007). Language experience and fixed expressions: Differences in the salience status of literal and figurative meanings of L1 and L2 idioms. In M. Nenonen & S. Niemi (Eds.), *Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes, Joensuu, Finland, May 19–20, 2006* (55–70). Joensuu: Joensuu University Press.

Cronk, Brian C. & Wendy A. Schweigert (1992). The comprehension of idioms: The effects of familiarity, literalness, and usage. *Applied Psycholinguistics, 13,* 131–146.

Cutting, J. Cooper & Kathryn Bock (1997). That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory and Cognition, 25* (1), 57–71.

Danesi, Marceli (1992). Metaphorical competence in second language acquisition and second language teaching: The neglected dimension. In J. E. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics* (489–500). Washington, DC: Georgetown University Press.

DeGroot, Annette (2002). Lexical representation and processing in the L2 user. In V. Cook (Ed.), *Portraits of the L2 user* (29–64). Clevedon: Multilingual Matters.

Deignan, Alice, Danuta Gabryś, & Agnieszka Solska (1997). Teaching English metaphors using cross-linguistic awareness-raising activities. *English Language Teaching Journal, 51* (4), 352–360.

Dell, Gary (1986). A spreading activation theory of retrieval in language production. *Psychological Review, 93,* 283–321.

Fernando, Chitra (1996). *Idioms and idiomaticity.* Oxford: Oxford University Press.

Fraser, Bruce (1970). Idioms within a transformational grammar. *Foundations of Language, 6* (1), 22–42.

Gibbs, Raymond W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory and Cognition, 8* (2), 149–156.

Gibbs, Raymond W. (1985). On the process of understanding idioms. *Journal of Psycholinguistic Research, 14* (5), 465–472.

Gibbs, Raymond W. (1986). What makes some speech acts conventional? *Journal of Memory and Language, 25,* 181–196.

Gibbs, Raymond W. (1995). Idiomaticity and human cognition. In M. Everaert, E.-J. van der Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms: Structural and psychological perspectives* (97–116). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gibbs, Raymond W. & Nandini P. Nayak (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology, 21,* 100–138.

Gibbs, Raymond W., Nandini P. Nayak, & Cooper Cutting (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language, 28,* 576–593.

Giora, Rachel (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics 8* (3), 183–206.

Giora, Rachel (1999). On the priority of salient meanings: Studies of literal and figurative language. *Journal of Pragmatics 31* (7), 919–929.

Giora, Rachel (2002). Literal vs. figurative language: Different or equal? *Journal of Pragmatics 34* (4), 487–506.

Giora, Rachel (2003). *On our mind: Salience, context, and figurative language.* Oxford: Oxford University Press.

Glucksberg, Sam (1993). Idiom meanings and allusional content. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (3–26). Hillsdale, NJ: Lawrence Erlbaum Associates.

Glucksberg, Sam (2001). *Understanding figurative language: From metaphors to idioms.* Oxford: Oxford University Press.

Grüneberg, Michael M., Peter E. Morris, & Robert N. Sykes (Eds.) (1988). *Practical aspects of memory: Current research and issues. Clinical and educational implications.* London: Wiley and Sons.

Heredia, R. Roberto (1997). Bilingual memory and hierarchical models: A case for language dominance. *Current Directions in Psychological Science, 6,* 34–39.

Howarth, Peter (1998). Phraseology and second language proficiency. *Applied Linguistics, 19* (1), 24–44.

Irujo, Suzanne (1986). Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *TESOL Quarterly, 20* (2), 287–304.

Irujo, Suzanne (1993). Steering clear: Avoidance in the production of idioms. *International Review of Applied Linguistics, 21,* 205–219.

Johnson, Janice & Teresa Rosano (1993). Relation of cognitive style to metaphor interpretation and second language proficiency. *Applied Psycholinguistics, 14,* 159–175.

Katz, J. Jerrold (1973). Compositionality, idiomaticity and lexical substitution. In S. R. Anderson & P. Kiparsky (Eds.), *A Festschrift for Morris Halle* (357–376). New York: Holt, Reinhart and Winston.

Kecskes, Istvan (2000). A cognitive-pragmatic approach to situation-bound utterances. *Journal of Pragmatics, 32*, 605–625.

Kövesces, Zoltan & Piter Szabo (1996). Idioms: A view from cognitive semantics. *Applied Linguistics, 17* (3), 326–355.

Kroll, Judith F. (1993). Accessing conceptual representations for words in a second language. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (53–82). Amsterdam: John Benjamins.

Kroll, Judith F. & Janet Curley (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. In M. M. Grüneberg, P. E. Morris, & R. S. Sykes (Eds.), *Practical aspects of memory: Current research and issues. Clinical and educational implications* (389–395). London: Wiley and Sons.

Kroll, Judith F. & Annette M.B. De Groot (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A. M. B. De Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (169–200). Mahwah, NJ: Lawrence Erlbaum Associates.

Kroll, Judith F. Alexandra & Sholl (1992). Lexical and conceptual memory in fluent and nonfluent bilinguals. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (191–203). Amsterdam: North-Holland.

Kroll, Judith F. & Erica Stewart (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149–174.

Kroll, Judith F. & Natasha Tokowicz (2001). The development of conceptual representation for words in a second language. In J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (49–71). Oxford: Blackwell Publishers.

Lattey, Elsa (1986). Pragmatic classification of idioms as an aid for the language learner. *International Review of Applied Linguistics, 24* (3), 217–233.

Lazar, George (1996). Using figurative language to expand students' vocabulary. *English Language Teaching Journal, 50,* 43–51.

Levelt, Willem J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, Willem J. M. & A. S. Meyer (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology, 12* (4), 433–452.

Liontas, John (2002). *Context and idiom understanding in second languages. EUROSLA Yearbook* (155–185). Amsterdam: John Benjamins.

Littlemore, Jeannette (2001). Metaphoric competence: A language learning strength of students with a holistic cognitive style? *TESOL Quarterly, 35* (3), 459–491.

Low, Graham D. (1988). On teaching metaphor. *Applied Linguistics, 9,* 125–147.

Luce, R. D. (1959). *Individual choice behavior.* New York: John Wiley.

Makkai, Adam (1972). *Idiom structure in English.* The Hague: Mouton.

Matlock, Teene & Roberto R. Heredia (2002). Understanding phrasal verbs in monolinguals and bilinguals. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (251–274). Amsterdam: Elsevier.

MacWhinney, Brian (2008). A unified model. In N. C. Ellis & P. Robinson (Eds.), *Handbook of cognitive linguistics and second language acquisition* (341–371). New York: Routledge.

Moon, Rosamund (1997). Vocabulary connections: Multi-word items in English. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (40–63). Cambridge: Cambridge University Press.

Potter, Mary C. Kwok-Faj, So, Barbara Von Eckardt, & Laurie B. Feldman (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23,* 23–38.

Roelofs, Ardi (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42,* 107–142.

Roelofs, Ardi (1997). The WEAVER model of word-form encoding in speech production. *Cognition, 64,* 249–284.

Sprenger, A. Simone (2003). *Fixed expressions and the production of idioms.* Ph.D. dissertation. [MPI Series in Psycholinguistics]. Ponsen and Looijen BV, Wageningen.

Sprenger, A. Simone, Willem J. M. Levelt, & Gerard Kempen (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language, 54,* 161–184.

Swinney, David A. & Anne Cutler (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior, 18,* 523–534.

Tabossi, Patrizia & Francesco Zardon (1995). The activation of idiomatic meaning. In M. Everaert, E.-J. van der Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms: Structural and psychological perspectives* (273–282). Hillsdale, NJ: Lawrence Erlbaum Associates.

Titone, Debra A. & Cynthia M. Connine (1994). Comprehension of idiomatic expressions: Effects of predictability, and literality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20* (3), 671–679.

Titone, Debra A. & Cynthia M. Connine (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics, 31* (12), 1655–1674.

Weinreich, Uriel (1969). Problems in the analysis of idioms. In J. Puhvel (Ed.), *Substance and structure of language* (23–81). Los Angeles: University of California Press.

Yorio, Carlos A. (1989). Idiomaticity as an indicator of second language proficiency. In K. Hyltenstam & L. K. Obler (Eds.), *Bilingualism across lifespan: Aspects of acquisition, maturity, and loss* (55–72). Cambridge: Cambridge University Press.

Zughoul, Muhammad R. (1991). Lexical choice: Towards writing problematic word lists. *International Review of Applied Linguistics, 29* (1), 45–6.

Appendix 1

A list of idioms and targets employed in the experiment. Idioms marked with an asterisk were rated as decomposable in the norming study.

IDIOM	Semantic Related	Phonological Related	Semantic Unrelated	Phonological Unrelated
Jack kicked the bucket.	WATER	BUDGET	HEAT	LIME
Jack played with fire. *	HEAT	FINE	PEAS	BOOK
Jack popped the question.	ANSWER	QUENCH	WATER	FINE
Jack pulled my leg.	FOOT	LEND	ANSWER	QUENCH
Jack walked on air.	WIND	AIM	AXE	BUDGET
Jack spilled the beans.	PEAS	BEAMS	FOOT	AIM
Jack sat on the fence.	GATE	FELL	WIND	LEND
Jack got cold feet.	HANDS	FEEL	ROPE	ISLE
Jack knocked on wood.	FOREST	WOMB	GATE	LUST
Jack read between the lines.*	ROPE	LIME	SNOW	FELL
Jack hit the road.	PATH	ROAST	HANDS	BEAMS
Jack buried the hatchet.	AXE	HASHISH	PATH	FEEL
Jack bent the rules.*	LAWS	ROOST	TUNE	MUSE
Jack beat around the bush.	TREE	BOOK	FORTUNE	ROAST
Jack pushed his luck.*	FORTUNE	LUST	TREE	ROOST
Jack faced the music.	TUNE	MUSE	LAWS	HASHISH

CHAPTER 10

A procedure for testing the Noticing Hypothesis in the context of vocabulary acquisition

Aline Godfroid*, Alex Housen* and Frank Boers*,**
*Vrije Universiteit Brussel (VUB) / **Erasmus University College Brussels

1. Introduction

1.1 The external perspective: Noticing and focus on form

Much research in applied linguistics directly or indirectly deals with the question of how learners achieve both fluency and accuracy in a second or foreign language (henceforth L2). Communicative methods of L2 teaching and meaning-focused modes of instruction have been shown to generate fluent speakers, whose production of some linguistic features, however, never reaches target-like levels (Doughty & Williams, 1998). This observation, among others, has led to the inclusion of a Focus on Form (FonF) component in recent strands of task-based learning, i.e. moments in a sequence of meaningful communicative activities where the students' attention is drawn to linguistic form[1] (Long, 1991; Doughty & Williams, 1998; Ellis, Basturkmen & Loewen, 2002).

FonF can either be proactive or reactive (Doughty & Williams, 1998). Episodes of proactive FonF are pre-established by the teacher in the lesson plan ('planned'), whereas reactive FonF occurs when a learner need arises. FonF interventions intend to make learners notice particular linguistic forms (e.g. word stress, an (irregular) verb form, the use of a plural verb after a singular noun, etc.).

The present chapter, however, takes the learner's perspective and considers episodes when the learner engages in noticing, independently of a teacher's intervention. More particularly, it will explore learners' noticing of new words in the absence of any external intervention other than the provision of comprehensible reading input.

When learners shift their attention to form 'by themselves', i.e. in the absence of a teacher's intervention, they engage in what we may term 'spontaneous' or 'learner-initiated noticing'. This is arguably the same type of cognitive processing as when FonF sequences succeed in drawing students' attention to the desired linguistic form (which is not to be taken for granted, cf. Jones, 1992, and Slimani, 1991). Thus, noticing as induced by a teacher in FonF and 'learner-initiated noticing' are complementary in the learning process: the more learners occasionally shift their attention to form by themselves, the less

they will require a teacher bringing formal linguistic items to their attention. The teacher could then spend the freed up time on further communicative activities instead.

The principal question we address in this chapter is how learner-initiated noticing can be measured. Before we do so, however, we need to discuss some of the defining characteristics of the construct of 'noticing'.

1.2 The internal perspective: Noticing in the literature on attention

Although SLA researchers have investigated the concepts of noticing, attention and awareness at least since the 1980s (e.g. Hulstijn & Hulstijn, 1984; Schmidt & Frota, 1986), Schmidt (1990) was arguably the first to offer an explicit and detailed account of noticing and its relationship to similar constructs such as attention and awareness. Schmidt claimed that subliminal learning (i.e. learning without awareness) is impossible. Rather, learners need to consciously notice language forms in order to take them in.

The concept of noticing has been very influential in SLA research since. It has been cited as a rationale for studies into the effectiveness of textual enhancement techniques (e.g. Alanen, 1995; Izumi, 2002; Jourdenais, Ota, Stauffer, Boyson, & Doughty, 1995), the use of focus on form (cf. Section 1.1) and for task-based L2 teaching (e.g. Ellis, 2003). As our research is also set within a noticing framework, it is appropriate to point out some conceptual and terminological issues that may encumber the operationalisation of the noticing construct.

Schmidt (1993) defines noticing as "the necessary and sufficient condition for the conversion of input to intake" (p. 209). Intake is an intermediary step between input and uptake, where we can conceive of *input* as all stimuli, perceptual and linguistic, that the learner is exposed to, and of *uptake* as the stimuli that are eventually committed to long-term memory.[2] Note that this characterisation does not specify exactly what *intake* is, other than that it is the product of noticing and an intermediary step in the acquisition process.

From Schmidt's (1990) description of different kinds of consciousness, it becomes apparent that it is consciousness in the sense of awareness which he understands to be involved in noticing. Importantly, consciousness (understood as awareness) is a gradual phenomenon, in which Schmidt distinguishes three levels: consciousness at the level of perception < consciousness at the level of noticing (focal awareness) < consciousness at the level of understanding. It follows that what is noticed is what one is focally aware of. Since awareness and attention are isomorphic in Schmidt's view, (Schmidt, 1993, 1994, 1995, 2001), this amounts to saying that we notice what is in the focus of our attention. Intake consists of those stimuli that have received our focal attention.[3]

Robinson (2003) adheres to basically the same view (see Figure 1). Figure 1 suggests that the allocation of focal attention is crucial for a stimulus (linguistic form) to enter working memory (WM). Only from there will the stimulus be susceptible to further processing and, eventually, storage in long-term memory (LTM). However, in addition to the connection between WM and LTM, Robinson (2003) posits the existence of a further link from stimuli in short-term memory (STM) to representations in LTM. Importantly, the status of stimuli in STM and WM is qualitatively different and, arguably, so is their impact on LTM.

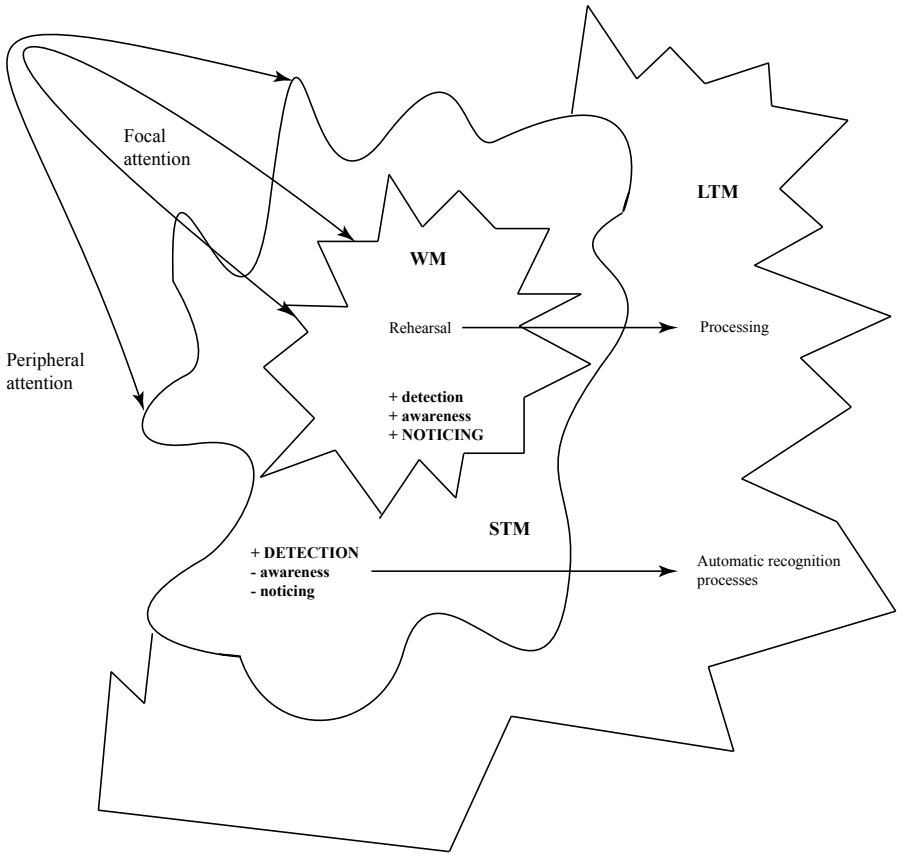


Figure 1. The difference between noticing and detection hinges on the amount of attention paid to the linguistic input, as well as the presence or absence of rehearsal in working memory (simplified from Robinson, 2003: 655)

The controversy about the impact of different stimuli on LTM dates back to Tomlin & Villa (1994) (contra Schmidt, 1990, 1993). Tomlin & Villa (1994) claim that a learner does not need to be aware of (i.e. does not need to consciously notice) stimuli in order to process them further. In their view, *detection* or "the cognitive registration of stimuli" (Tomlin & Villa, 1994: 192) is sufficient for learning. Hence, the distinguishing factor between noticed stimuli and detected stimuli (in Figure 1 situated in WM and STM, respectively) is that the learner is consciously aware of the former, but not of the latter. The learner's awareness of the noticed stimuli is the result of focal attention.

Most researchers will ascribe to the psychological reality of subliminal perception: people also register stimuli they do not attend to. The question, however, is whether mere *perception* can also lead to *learning*. Tomlin & Villa (1994) say it can, whereas Schmidt (1990 and elsewhere) and Robinson (1995, 2003) argue it cannot.

Within the field of SLA, Leow and associates conducted a very informative series of studies into the relationship between awareness and learning. In his 2000 study, Leow

assessed learners' awareness of irregular stem changes in Spanish *ir* verbs in the preterit (e.g. *preferir* becomes *prefirió*). Beginning students of Spanish were asked to solve a crossword puzzle by filling out the preterit endings of (mostly) stem-changing verbs. For example, for 'she preferred' students had to add *–ió* to the stem *prefir-*. Note that the irregular vowel change is already given, thus creating a context in which some learners will notice the irregularity and others will not. During the task, subjects were required to "speak their thoughts aloud". On the basis of the think-aloud protocols and their responses to a post-exposure questionnaire, the subjects were classified as either aware or unaware. Analyses revealed no significant differences between pre-test and post-test scores for the 'unaware' group, whereas the 'aware' group performed significantly better on the post-tests. These findings support Schmidt's (1990 and elsewhere) and Robinson's (1995, 2003) claim that conscious noticing plays a crucial role in subsequent processing of L2 data.

Leow (2000) is a replication of Leow (1997) with an important methodological improvement: it presents the stimuli in the crossword puzzle in such a way that it produces both aware and unaware learners (see above). A different, but related question is whether different levels of awareness will result in different amounts of learning. This question is addressed in Leow (1997), for Spanish irregular stem-changing verbs, and in Rosa & O'Neill (1999) and Rosa & Leow (2004), for Spanish contrary-to-fact conditionals. They used the think-aloud protocols that their participants produced during the experimental (problem-solving) task to classify them as 'aware at the level of noticing' (low-level awareness) or 'aware at the level of understanding' (high-level awareness).[4] Participants who demonstrated a high level of awareness outperformed participants who were only aware at the level of noticing. In summary, then, these four studies provide evidence that awareness at the level of noticing is facilitative of and may be necessary for L2 learning, and that higher levels of awareness lead to higher learning gains.

In contrast to the above, Williams (2005) suggests that particular form-meaning connections can be learnt *without* awareness. Williams (2005) conducted two experiments in which he acquainted college students with four novel (fictitious) articles: *gi, ro, ul* and *ne*. The "rules" behind the choice of the articles were (i) the distance to the subject of the sentence (*gi* and *ro* for 'near' and *ul* and *ne* for 'far') and (ii) noun animacy: animates take *gi* and *ul* whereas inanimates take *ro* and *ne*. The participants, however, were given instruction only about the first rule. They engaged in a sentence comprehension task in which the emphasis lay on working out the distance relationship implied by the novel word and the overall meaning of the sentence. The [+ animate] and [- animate] rule was also instantiated by the exercise items, but was not addressed by the task instruction. Next, the students were given post-tests designed to assess their appropriate article choice in compliance with the 'animacy' rule. Both stages of the post-test contained both previously seen items and new, so-called generalisation items. The latter allow assessing the learners' ability to generalise across items and apply their knowledge productively. Students were classified as either aware or unaware of the relevance of the animacy feature on the basis of their responses to post-task interviews. Crucially, students who did *not* demonstrate any awareness of the animacy feature nevertheless selected articles in compliance with the animacy rule at significantly above-chance levels, and did so also concerning the generalisation items. These results are suggestive of some form of learning without awareness: they

attest to the unaware learning of the (unnoticed) connection between a (noticed) form and an (unnoticed) meaning *feature* inherent in the (noticed) meaning of a word. They may therefore lend support to Tomlin & Villa's (1994) position that *detection* is sufficient for some forms of learning. These results certainly add evidence to Nick Ellis' implicit tallying hypothesis which states that "once a stimulus representation is firmly in existence, that stimulus (…) need never be noticed again; yet as long as it is attended to for use in the processing of future input for meaning, its strength will be incremented and its associations will be tallied and implicitly cataloged" (N. Ellis, 2002: 174).

At any rate, also in cognitive psychology a number of studies corroborate Schmidt's Noticing Hypothesis (Ericsson & Simon, 1993; Underwood, 1976, 1982; Baars, 1988, 1997, 2002). Baars (2002) reviews the role of consciousness in cognitive processing and synthesises the dominant view as follows: "There appears to be no robust evidence so far for long-term learning of unconscious input" (p. 50). Although we find automatic recognition and/or categorisation processes triggered by familiar stimuli (e.g. Nosofsky, 1992), these only appear to lead to the activation of existing memory representations, not to learning. In this view, the connection between STM and LTM in Figure 1 represents the activation of encodings already present in LTM. Other researchers, such as Kellogg and Dare (1989), claim that unattended input can actually be briefly encoded. However, they emphasise that this "does not imply that unattended encoding has any practical value (...) [since] the degree of elaboration resulting from unattended encoding appears to be too limited to have any substantive influence on human cognition or behaviour" (p. 412).

In brief, while the issue of whether there can be any learning on the basis of unconscious detection alone has not yet been settled, many researchers now agree that unconscious learning – if it exists at all – is negligible.

2. Rationale for the present study

In the introductory section we showed that the cognitive process of noticing is one of the theoretical foundations of FonF. Doughty (2001) summarises this view as follows: "The factor that distinguishes FonF from the other pedagogical approaches is the requirement that FonF involves learners' *briefly and perhaps simultaneously attending to form, meaning and use during one cognitive event*" (p. 211). Robinson (2003) also emphasises the importance of attention. He writes that "noticing involves that subset of detected information that receives focal attention, enters short-term working memory and is rehearsed" (p. 655). In his view, the resource-directing dimensions of a task will determine to what aspects of language code attention is directed. Robinson conceives of noticing within a task-based setting, which entails a focus on meaning and use. The object of noticing is then some aspect of linguistic code (or 'form'). Thus, both Doughty (2001) and Robinson (2003) put forward that noticing, or focal attention to form, cannot occur in isolation from meaning or use.

We now turn to the thorny issue of how the event of noticing can be attested or measured. So far, researchers have used a variety of techniques to gauge noticing. Some of these are *offline*, i.e. the data is elicited *after* the task. Examples include questionnaires

(e.g. Robinson, 1997a, b), stimulated recall protocols (Mackey, Philp, Egi, Fujii & Tatsumi, 2002; Mackey, 2006), interviews (e.g. Williams, 2005) and diary entries (Schmidt & Frota, 1986). In contrast, *online* techniques collect the data *while* learners are doing the task. Think-aloud protocols are a fairly common example of the latter kind (e.g. Alanen, 1995; Jourdenais et al., 1995; Leow, 1997, 2000; Rosa & O'Neill, 1999; Rosa & Leow, 2004). Self-reports in the form of note-taking (Hanaoka, 2007; Izumi, 2002) are another example. The temporal resolution of online measures is of course much higher and in view of the fleeting nature of noticing, many researchers now agree that online measures are the more reliable option. Alternatively, a combination of online and offline measures may give a more complete picture. This is the approach Leow and colleagues opted for, using both (online) think-aloud protocols and (offline) questionnaires.

Although online techniques such as think-aloud protocols provide a window into mental processes, they carry a risk of 'reactivity', i.e. concurrent verbalisation of cognitive processes may influence the very cognitive processes one is aiming to describe. For example, a think-aloud task may interfere with the main task and negatively affect performance (negative reactivity), or it may stimulate the subject to perform more accurately or systematically (positive reactivity). Whether concurrent verbalisation will be reactive depends on a complex interplay of variables, one of which is the degree of detail of verbalisation. Whereas *non-metalinguistic* verbalisations require subjects to verbalise their thoughts per se, *metalinguistic* verbalisations involve subjects giving additional information such as reasons and justifications for what they are thinking (Bowles & Leow, 2005, drawing on the typology in Ericsson & Simon, 1993). According to the review by Ericsson & Simon (1993), only metalinguistic think-alouds are often reactive for accuracy (but see Sanz, Lin, Lado, Wood Bowden, & Stafford, 2009, for contrary evidence).

So far, five published studies have investigated the potential reactivity of think-aloud protocols in SLA (Bowles, 2008; Bowles & Leow, 2005; Leow & Morgan-Short, 2004; Sachs & Polio, 2007; Sanz et al., 2009). Leow & Morgan-Short (2004) and Bowles & Leow (2005) have found that concurrent verbalisation during an L2 reading task is not reactive for accuracy in post-test performance, although it does increase time on task (Bowles & Leow, 2005). Bowles (2008) also states that metalinguistic verbalisations are reactive for the time spent on a problem-solving task. In addition, she has observed a negative effect on the written production of old (i.e. previously seen) items in the metalinguistic think-aloud group. Sachs & Polio (2007) have found that thinking aloud while processing the feedback (reformulations) on an L2 writing task negatively affects the accuracy of the subsequent revisions. Finally, Sanz et al. (2009) have found that thinking aloud enhances post-test performance following a moderately explicit lesson about the assignment of semantic functions in Latin (experiment 2). Sanz et al. (2009) are innovative in that they are the first to analyse reaction time data ('latency') in addition to accuracy as a dependent measure.[5] We think that the inclusion of reaction time data in future investigations of the reactivity of think-aloud protocols will provide a more precise picture of the potential effects of concurrent verbalisations on cognitive processing. In the meantime, the empirical evidence summarised above leads us to agree with Sanz et al. (2009) and Sachs & Polio (2007) that "think-aloud protocols should be employed and interpreted with care" (Sachs & Polio, 2007: 68).

Given the limitations of offline techniques and the risk of reactivity inherent in think-aloud protocols, we decided to explore an alternative online technique, namely eye tracking. Eye tracking, the online registration of someone's eye movements, does not involve a secondary task and therefore is not susceptible to reactivity. Nonetheless, the use of eye-tracking technology to measure noticing is not immune to methodological concern either (see final section, conclusions and perspectives).[6]

Before we turn to a description of how eye tracking can be embedded in a procedure to measure noticing, we need to explain why we will be focussing on vocabulary rather than grammar as a target for noticing (and possibly uptake).

To date, there have been several empirical studies about noticing. Table 1 provides an overview of studies that have investigated the noticing of *predetermined linguistic features.*[7]

Table 1. Overview of studies that have investigated the intake or noticing of a particular, predetermined linguistic feature

Study	Linguistic feature	Linguistic field
Alaanen, 1995	semi-artificial Finnish locative suffixes	morphology
Jourdenais et al., 1995	Spanish preterit and imperfect verb forms	morphology, semantics
Robinson, 1997a	English dative alternation	syntax
Robinson, 1997b	English pseudo-clefts of location + optional subject-verb inversion after fronted adverbial of movement or location*	syntax
Leow, 1997, 2000	Spanish irregular ir verbs in the preterit	morphology
Rosa & O'Neill, 1999; Rosa & Leow, 2004	Spanish contrary to fact conditionals	morphology, syntax
Mackey et al., 2002	English questions (recasts of non-targetlike instances)	syntax

* The examples Robinson (1997b) uses include the following: *Where Alice stands is on the right not on the left* (pseudo-cleft of location) and *Off the horse Amy fell* or *Off the horse fell Amy* (optional inversion after fronted adverbial of movement).

We can see from Table 1 that all these studies have looked at syntactic or morphological features. On the other hand, at least two studies that looked into *what linguistic items* learners tend to notice by themselves have found that learners largely focus on words (Williams, 2001; Hanaoka, 2007). This suggests that, in incidental contexts, learners are perhaps more inclined to allocate part of their attention to (new) words (which they perceive as meaning-bearing) than to 'formal' features (appreciation of which may seem less vital for comprehension and which are less perceptually salient). This makes vocabulary an attractive target in a study like ours, whose aim is to assess the connection between noticing and learning and, if the Noticing Hypothesis as put forward by Schmidt and others is valid, the association between awareness and learning.

For a characterisation of *incidental learning* in general, we largely follow Hulstijn (2003), who draws on Schmidt (1994). Incidental learning amounts to "learning without the intent to learn" (Schmidt, 1994: 16). More specifically, incidental learning can also be the "learning of one thing (e.g. grammar) when the learner's primary objective is to do something else (e.g. communicate)" (ibid). For learning to be incidental, the primary

objective can never disappear completely out of consciousness (i.e. it will always remain in the background) *even when attention is shifted to another object to be learnt* (Doughty, 2001). If the learner does lose sight of her primary goal, while focusing on this new object, then she will no longer be learning it in an incidental fashion. If we think of incidental learning as this attentional gradient shifting temporarily from the primary learning goal to a secondary goal, then it links up easily with noticing. Noticing occurs when the focus of attention shifts to secondary aspect(s) of the activity, while the primary aspects still remain in the 'fringe' of consciousness (Mangan, 1993, cited in Baars, 1997). This event then signifies the initial step in the incidental learning process.

Presenting a survey of the vast body of research on L2 incidental *vocabulary* acquisition is beyond the scope of this chapter (but see, for example, Wesche & Paribakht, 1999; Nation, 2001; Hulstijn, 2003; Laufer, 2005 and Schmitt, 2008, for reviews). However, one thing most experts in the field now agree on is that incidental uptake of vocabulary through extensive reading alone is much slower a process than was once believed, for example, by Krashen (1985, 1989). One of the reasons for this slow learning rate is that incidental vocabulary acquisition is an incremental process (Nation, 2001): the learner needs to encounter a new word repeatedly (estimates suggest 10 times or more) in a relatively short span of time in order for the item to start leaving a stable trace in the learner's memory. This requires much more reading than most learners are willing to engage in (e.g. Laufer, 2005), even more so at an intermediate-to-advanced level, where potentially useful additions to the learners' repertoires are generally of low frequency (e.g. Boers & Lindstromberg, 2009, chapter 3).

A second potential reason, which goes to the heart of our study, is that learners *may not take notice* of the unknown words they come across in a text, either because they feel comprehension of those words does not impede their overall text comprehension or because they overestimate their word knowledge (i.e. they wrongly think they know the word; Laufer, 2005).

In short, vocabulary acquisition research informs us that we should not be surprised to find low learning rates for words to which learners have been exposed only once or a couple of times. On the other hand, given the notion that learning is an incremental, cumulative process, we cannot deny the possibility that some memory trace, however feeble, *is* created on a first encounter with a word. According to the Noticing Hypothesis, however, this possibility hinges on the learner taking notice of the word in the first place. In order to detect such feeble memory traces, a sensitive instrument capable of measuring (very) partial word knowledge is needed.

3. A procedure for gauging noticing and its association with uptake: A pilot study

In what follows we propose and evaluate a procedure for measuring noticing and its association with the (gradual) uptake of new words through silent reading. The procedure consists of three stages: (1) eye tracking as a means for measuring the noticing of new words; (2) post-tests to (a) measure the memory traces (if any) created by encounters with unknown words and (b) check participants' text comprehension (to ascertain that

they engaged in reading with meaning-focus); and (3) stimulated recall protocols to gain insight into participants' subjective experience of the mental activities (if any) triggered by encounters with unknown words.

3.1 Participants

Nine native speakers of Dutch who learnt English at school as a foreign language volunteered to participate in this pilot study. Four participants were second-year English language majors (undergraduates). The other five participants were PhD students in various branches of the humanities, but *not* in linguistics. Given their habitual reading of English academic texts, we estimated the non-language majors' receptive knowledge of English would be on a par with the undergraduate language majors'.

3.2 Apparatus

Eye movements were recorded with the Eyelink II, a head-mounted eye tracker manufactured by SR Research. Text stimuli were presented in font Courier New, size 18, on a computer monitor at 68 cm distance from the participants' eyes. At this distance, 4.0 character positions equalled 1° of visual angle. The use of corneal reflections in combination with pupil tracking ('pupil + CR mode') allowed for more stable tracking of eye positions. In addition, participants were asked to place their heads on a chin rest to minimise head movements.

3.3 Target forms

As indicated in Section 2, the principal aim of the present study was to gauge the amount of learner-initiated noticing of new lexical items that took place as learners read short texts with meaning-focus. We collected data online and offline, by means of eye tracking and post-tests, respectively. We opted to use pseudo words as target forms (which, by definition, were completely unknown to the participants). Since unknown words have a (subjective) frequency of zero, the influence of frequency on fixation behaviour was controlled for. Also, vocabulary post-test performance could not be facilitated by some previously established memory traces. As a result, increased fixation times or correct responses in the post-test could unambiguously be interpreted as corollaries of noticing events.[8] In order to safeguard the ecological validity of the experiment, it was crucial that our pseudo words were plausible words. Participants should *not* debunk the pseudo words as such. We therefore made sure they were of high phonotactic probability.

Although eye movements and fixations are a widely accepted source of information about ongoing cognitive processing, they are also influenced by other, confounding variables. Apart from familiarity with the target words (which we ruled out by using pseudo words), confounds include word length, predictability, part of speech, concreteness and frequency in spoken language. Thus, a reader's viewing behaviour is determined by a complex interplay of factors. In order to control for all of the confounding variables, we used a Latin square design, in which each stimulus serves as its own control. A Latin

square is a special ANOVA design. One of its distinguishing features is the equal number of treatments and groups, which gives the design its 'square' form. We have worked with a 4x4 design. Every participant was assigned randomly to one of the four groups, which stand for a particular combination of stimuli in particular conditions (see Table 2).

Table 2. A schematic representation of a 4x4 Latin square design

		PARAGRAPHS			
		1&5	2&6	3&7	4&8
GROUP	I	A	C	B	D
	II	D	B	C	A
	III	C	D	A	B
	IV	B	A	D	C

The Latin letters A–D represent treatments or conditions. Thus, every group got to read all the stimuli (§§1–8), but in a different combination of conditions. The randomisation of conditions over groups was such that at the end, each group received each of the conditions and every stimulus was presented exactly once in every condition.

Let us now turn to the different conditions. A target form could take any one of four appearances:

– *Condition I*: the control condition
The target word is an existing word, known by the subjects, e.g.:
(1) One demographical group stands out as being most likely to push *boundaries* and break rules.

– *Condition II:* pseudo word
The target word is a pseudo word, and thus unknown to the subjects, e.g.:
(2) One demographical group stands out as being most likely to push *paniplines* and break rules.

– *Condition III:* pseudo word + existing word
The target region is a pseudo word followed by a clarifying, existing word, e.g.:
(3) One demographical group stands out as being most likely to push *paniplines or boundaries* and break rules.

– *Condition IV:* existing word + pseudo word
The target region is a pseudo word preceded by a clarifying, existing word, e.g.:
(4) One demographical group stands out as being most likely to push *boundaries or paniplines* and break rules.

Note that only conditions II to IV contained an unknown word that according to our working hypothesis would be susceptible to being noticed. Nevertheless, condition I was important, as it enabled us to compare reading times for an (unnoticed) known word with processing measures for a (noticed or unnoticed) unknown word *that occurred in exactly the same context.*

Thanks to this feature, confounding variables such as predictability, part of speech and concreteness were controlled for. In addition, we selected pseudo words such that they had the same number of letters and syllables as the existing word they substituted. (Consider the above example of *boundaries – paniplines*.) Therefore, marked differences in viewing behaviour could not be attributed to differences in word length or number of syllables either. In brief, given our tight design, significantly higher reading times for pseudo words would have to be interpreted as behavioural evidence of noticing.

3.4 Reading task

Eight target forms (pseudo words) were embedded in short paragraphs, between nine and thirteen lines in length. For an example of such a paragraph, see Appendix 1. The paragraphs with the target forms were supplemented with eight 'dummies', i.e. regular paragraphs made up of existing words only.

The participants were kept ignorant about the true aim of the experiment. We told them our study was meant to compare how people read a text on a computer screen with how they read a text on paper. Supposedly, they belonged to the group asked to read on a screen. We emphasised that for the validity of our experiment, it was very important that they read as naturally as possible, at their own pace, "the way they would read a novel in their spare time."

After we set up the cameras, the subjects read the sixteen paragraphs in a random order. Every paragraph took up three screens.[9] Subjects could move from one screen to the next by clicking any button on a pad. Appendix 2 contains an example of what a paragraph on screen looked like.

3.5 Post-tests: Description and rationale

After the reading task, the participants received two unannounced post-tests, one on vocabulary and one on content. Vocabulary was tested first to minimise temporal decay of the memory trace. However, for reasons of practicality we will present the post-tests in reverse order in this paper.

3.5.1 *Content post-test*

The content post-test ascertained that the participants had actually read with focus on meaning. It also offered the possibility of gauging possible interactions between reading comprehension and attention to form (reflected in the noticing behaviour vis-à-vis the pseudo words). The post-test consisted of sixteen true/false statements, one for each paragraph the participants read. The following statement referred to the paragraph reproduced in Appendix 2:

(5) On average, men tend to break the speed limit more often than women.
true / false
(Participants had to press 'f' for 'false'.)

3.5.2 *Vocabulary post-test*

The vocabulary-retention post-test needed attuning in order to avoid undesirable floor and ceiling effects. The fact that these floor and ceiling effects occurred, however, is interesting in its own right, which is why we will describe the steps it took us to find a suitable format.

The first format we tried was a 'multiple-choice recognition test'. In view of the repeated encounters that are generally needed to learn a word (cf. the discussion on incidental L2 vocabulary acquisition in Section 2), and the fact that our participants would have met the target words (i.e. the pseudo words) only once, we set out with the idea that our vocabulary post-test would have to be easy and gauge receptive knowledge only. Hence we created a test in which participants were given ten seconds to select the target word from a list of four possibilities. As the example below shows, the possibilities consisted of two pairs of similar pseudo words, controlled for length and number of syllables:

(6) One demographical group stands out as being most likely to push .....
and break rules.
1. dilactives
2. dilectives
3. paniplines
4. poniplines

Here the subject should give answer '3', *paniplines*.

As we will see in the results section further below, this test format produced a ceiling effect. The test was too easy; it needed rethinking.

The second test format we tried was a 'cued production test'. This was a gap-fill exercise in which the first three or four letters of the target word were given.[10] For example:

(7) One demographical group stands out as being most likely to push pani.....
and break rules.
(Participants should enter *plines*.)

The results we found were quite the opposite of the previous ones: the participants who took this test all scored at floor (for further details, see the results section).

Bearing in mind the ceiling and floor effects observed in the above tests, we returned to the multiple-choice format but made it more challenging. We extended our list of distractors to seventeen (instead of the initial three) and used a fifteen-second time limit. The new format looked as follows:

1. quirkology	7. recafic	13. amiblions
2. cobrion	8. erision	14. canimat
3. wricety	9. scrandivist	15. broadcaster
4. jurdistemy	10. perchants	16. hypergenity
5. evidoses	11. paniplines	17. dilactives
6. levider	12. offspring	18. flarrisation

(8) One demographical group stands out as being most likely to push .....
and break rules.

The participants who took this test scored between 3/8 and 6/8, thus providing the kind of spread that could be useful to investigate potential associations between future cohorts of participants' noticing behaviour (i.e. the eye-tracking data) and their post-test scores. For now, the question remains why the initial multiple-choice test and the cued reproduction test were too easy and too hard, respectively, and what this tells us about the nature of the noticing process.

3.6 Stimulated recall protocol

At the end of each experimental session, the researcher ran through the vocabulary test items again together with the participant and noted down her spontaneous comments. These turned out quite informative. For example, one participant could tell for most of the items whether he remembered reading them and how he had dealt with them. He looked over the new words that were accompanied by a synonym (conditions III and IV) as he "did not really need them to understand the text", but tried to derive the meaning of the pseudo words in isolation (condition II) from the context. This sort of qualitative information may help us gain insight into participants' mental elaboration processes that they might remember the encounters with the unknown words had brought about. We feel that the stimulated recall protocol will enrich the otherwise quantitative data and have decided to include it as an integral part of our procedure.

Stimulated recall protocols provide learners with some sort of support to facilitate the recall of the cognitive processes (or strategies) which they experienced (or applied) at the time of the treatment (Gass & Mackey, 2000). In the present study, the recall data were collected by exposing learners once more to their performance on the vocabulary post-test items rather than to their actual interaction with the target words as they appeared in the texts during the reading sessions. As such this study departs from previous studies, which typically provide support that is directly related to the actual treatment or the experimental task. Support based on the actual treatment in this study could have taken the form of showing the learners film recordings (in slow motion) of their eye movements while they were reading the paragraphs. However, we suspect that this procedure would have led learners to make many new inferences that occurred to them while viewing the film recordings rather than to reflect the cognitive processes they truly had while performing the experimental reading task (the more so since our subjects had been administered a vocabulary post-test shortly before the stimulated recall session, which oriented their attention exclusively to novel words). Hence, the stimulated recall procedure applied in our study is motivated by a concern for veridicality, that is, to ensure that the account of the learners' thought processes about the novel words is as accurate as possible. At the same time we acknowledge that the time gap between the treatment and the stimulated recall constitutes an important limitation inherent to this method (see Section 2), and one which needs to be taken into account when interpreting the protocols.

3.7 Summary of the experimental procedure

Figure 2 is a graphical representation of the different components in our procedure. In the next section we will present results for the quantitative parts of the experiment.

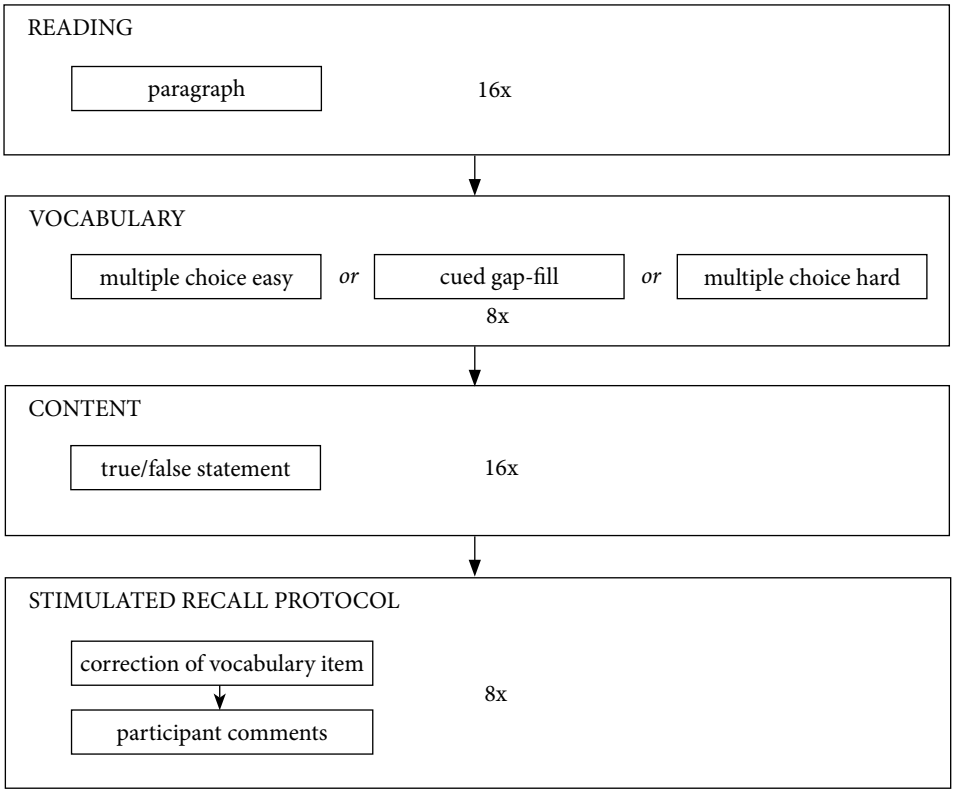


Figure 2. Flow chart illustrating the experimental procedure

4. Results

4.1 Eye measures

The suitability of eye tracking as an instrument for measuring noticing during reading is confirmed. We found plenty of instances where one can tell by sight that the target word received notably more attention than the other words on the screen. Consider the following example:

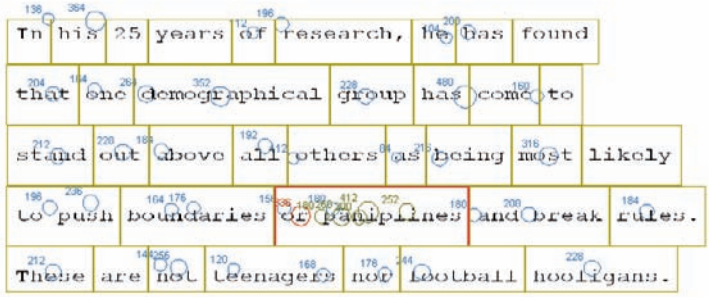


Figure 3. A clear-cut example of a noticing event

The viewing behaviour vis-à-vis the target region, *or paniplines*, can be assessed through various measures. When words are the unit of analysis, *first fixation duration* and *first run dwell time* (or *gaze duration*) are the two most frequently used (Rayner, 1998). First fixation duration is the duration of the first fixation on a word, irrespective of whether the word was fixated one or multiple times. First run dwell time (or gaze duration) is the sum of all fixations made on a word before the eyes leave that word. In Figure 3, first fixation duration is 336 msec and first run dwell time 1640 msec.[11]

However, the eyes also leave the target region (after the 252 msec fixation) and *regress* to it later on, after the 184 msec fixation made on *rules*. Longer regressions like this one occur because the reader has problems understanding the text (Rayner, 1998). Rayner adds: "In such cases, good readers are very accurate in sending their eyes to that part of text that caused them difficulty (Frazier & Rayner, 1982; Kennedy, 1983; Kennedy & Murray, 1987a, 1987b; Murray & Kennedy, 1988), whereas poor readers engage in more backtracking through the text (Murray & Kennedy, 1988)" (Rayner, 1998: 375). The very high first run dwell time for *paniplines* and the fact that this is the very word the eyes regress to later on support the claim that integration of the meaning of the new word *paniplines* into the sentence is what caused the reader to make a regression. The reader was thus cognitively engaging with the target word, which we interpret as a noticing event. Additional evidence is found in overall *dwell time*, i.e. the sum of all fixations on a word, including regressions back to the word. In Figure 3, the dwell time for the target region is 1976 msec. By way of comparison, the average eye fixation when reading an English word lasts between 200 and 250 msec (Rayner, 1998).

4.2 Content scores

Table 3 presents the results for the 16 true / false statements, juxtaposed with the scores on the vocabulary post-test.

Table 3. The content post-test scores

Subject ID	Group	Vocabulary post-test		Content post-test
		Format	Score (max. 8)	Score (max. 16)
1	1	multiple choice easy	5	14
2	2		6	16
3	2		6	16
4	4	cued gap-fill	0	15
5	4		0	12
6	3		2	11
7	3	multiple choice hard	5	11
8	2		6	15
9	1		3	9

The mean content score is 13.22 (s.d. = 2.54), which indicates that overall reading comprehension was very good. Participants took their task seriously and seem to have read the paragraphs with a general focus on meaning, as they had been instructed to.

One of the methodological improvements we intend to implement in a larger study is to include 'I don't know' as a possible response to the true / false statements. There will then no longer be 50% chance of getting the answer right, which may actually yield somewhat lower scores.[12]

4.3 Vocabulary scores

We now shift our attention to the results of the vocabulary post-test, grouped per test format (Table 4):

Table 4. The vocabulary post-test scores

Subject ID	Group	Vocabulary post-test		Content post-test
		Format	Score (max. 8)	Score (max. 16)
1	1	multiple choice easy	5	14
2	2		6	16
3	2		6	16
4	4	cued gap-fill	0	15
5	4		0	12
6	3		2	11
7	3	multiple choice hard	5	11
8	2		6	15
9	1		3	9

Three of our participants were given the ('easy') multiple-choice test with three distractors. We obtained a ceiling effect (mean score 5.67), irrespective of subjects' viewing behaviour vis-à-vis the target words concerned. In other words, these three participants recognised the targets no matter whether the eye-tracking data indicated they had taken special notice

of them during reading. This suggests that mere perception of the target words sufficed to create a memory trace that enabled participants to discriminate them from distractors shortly after exposure. One might argue that an average of 5.67 out of 8 is high but not at ceiling. This argument would hold, were it not that all three participants made the exact same mistakes: they picked 'e*p*idoses' instead of 'e*v*idoses' and 'cani*c*at' instead of 'cani*m*at'. Hence the target items failed to discriminate between the participants. In addition, recall that this test used two pairs of similar pseudo words (e.g. amiblions – amiglions – epidoses – evidoses). When participants answered incorrectly, they *always* picked the distractor that was highly similar to the target word (e.g. 'epidoses' instead of 'evidoses'). The memory trace created may not have been sufficiently refined to recall the exact word form, including segmental information, but it nevertheless enabled the participants to recognise the target word at least by approximation.

As we described in Section 3.5.2, the results of the initial multiple-choice test led us to create a more difficult test, which measured (cued) productive knowledge. Another three participants were given this test and they all scored at floor (mean score 0.67), despite the fact that the eye-tracking data for some target words suggested that they had indeed taken notice of them. It seems clear that a more stable memory trace than that created by a single encounter is needed to enable participants to recollect (part of) words for (semi-) productive purposes. The combined results of both tests suggest that a single exposure can lead to some 'intake' but not yet to 'uptake', which is fully in accordance with the notion that vocabulary acquisition is an incremental process (Nation, 2001; Laufer, 2005).

Finally, our ('harder') multiple-choice test with 17 distractors yielded an average result of 4.66, with scores ranging between 3/8 and 6/8. Unlike the two previously tried formats, this test suitably discriminated between the participants, and therefore looks like a promising format to help compute correlations between noticing behaviour and subsequent recollection of target words.

5. Conclusions and perspectives

When learners spontaneously shift their attention to linguistic form, they extract potentially useful information from the L2 input, which will eventually help them produce more accurate and varied stretches of the target language. This shift in attention is known as *noticing*. In this chapter we have described a procedure for measuring learner-initiated noticing of words during meaning-focused reading and its association with (incremental) uptake of vocabulary. The procedure combines an online measure, viz. eye tracking, and offline measures, viz. vocabulary and content post-tests, complemented by a stimulated recall protocol.

The advantage of using eye tracking as an online measure is that, contrary to concurrent verbalisations, it does not carry the risk of interfering with participants' cognitive processing. The advantage of the particular vocabulary post-test format we have in the end opted for as the offline measure is that it seems sensitive enough to pick up even weak memory traces while still being challenging enough to discriminate between different participants' performances. However, neither eye-tracking data nor post-test scores can inform

us directly about the type and depth of cognitive processing that ensues from a noticing event. It is quite possible that it is the type of *mental elaboration that follows noticing* which determines the strength of the memory trace created and, thus, determines eventual learning gains. It is in this connection that stimulated recall protocols can yield useful insights, even though we must bear in mind that their outcome is subject to memory constraints.

With regard to eye tracking as an instrument for measuring noticing, Figures 3 to 8 illustrate that learners' eyes sometimes do dwell on words that are new to them and that learners do interrupt their natural course of reading to regress to such words. It is generally agreed that "how long [readers] fixate on a word is related to the ease or difficulty with processing that word" (Rayner, 1998: 389). Furthermore, "eye movements are intimately related to the moment-to-moment cognitive processing activities of readers" (Rayner, 1998: 389–390).

A larger-scale application of the proposed procedure will enable us first to compute correlations between vocabulary post-test scores and the amount of noticing activity displayed by participants. Secondly, the Latin square design will enable us to test for quantitative differences in noticing behaviour triggered by different conditions. The following comparisons seem of particular interest:

- control condition I ↔ experimental conditions II, III and IV ("Do advanced learners of English spontaneously notice new words when reading an English text?")
- [– synonym] condition II ↔ [+ synonym] conditions III and IV ("Does a clarifying context dampen the extent to which learners notice new words?")
- [succeeding synonym] condition III ↔ [preceding synonym] condition IV ("Does the position of the clarifying co-text have a differential effect on the extent to which learners notice new words?")

Thirdly, the procedure can be used to examine potential differences among individuals in their inclination (or ability) to give notice to 'form' during meaning-oriented language processing. If marked individual differences were attested, then this would raise many new research questions, in particular about what particular characteristics might correlate with one's inclination for noticing.

Finally, in connection with the debate about awareness and learning that we started this chapter with, we were surprised to find that the first vocabulary post-test we tried, i.e. the multiple-choice test with three distractors, generated such high scores. The participants recognised the target words even though they had encountered them only once during meaning-focused reading. Crucially, they sometimes recognised these words in the post-test even though the eye-tracking data showed no marked increases in viewing activity for the target area. The question that remains to be answered, then, is whether analyses on larger data sets will reveal a critical boundary in terms of viewing activity below which input can be said to be 'perceived' and above which it is 'noticed'. The current data allow identifying clear cases of heightened attentional processing (see e.g. Figure 3). Thus, the upper end of the attentional gradient can be gauged fairly straightforwardly. The challenge for future studies is to tease apart the lower end of the attentional gradient and see how it relates to learning. In our view, this concern also holds for investigations of attention, awareness and learning that make use of think-aloud protocols (Leow, 1997, 2000; Rosa & O'Neill, 1999; Rosa & Leow, 2004). After all, learners do not verbalise everything that they

notice (Jourdenais, 2001) or, inversely, "lack of self-report or recall does not necessarily imply lack of awareness" (Rosa & O'Neill, 1999: 170, referring to Allport, 1988). Think-aloud protocols are undoubtedly a useful tool to capture clear, 'outspoken' instances of awareness (at the level of noticing and at the level of understanding), but less pronounced instances are likely to slip through the net that think-aloud procedures cast. We hope that this eye-tracking study has pointed to some fascinating but as yet underexplored aspects of the cognitive mechanisms on which the Noticing Hypothesis is built.

Acknowledgements

We would like to thank Peter Robinson and an anonymous reviewer for their helpful comments on an earlier draft of this chapter. We are also indebted to Denis Drieghe, Marc Brysbaert, Leah Roberts, Maria Ysebaert and Jiye Shen from SR Research for their valuable support and insightful advice on the experimental design and the data analysis. We would also like to thank Eric Soetens and David Henderickx for sharing their eye-tracking lab. Last but not least, we are grateful to all the volunteers who kindly gave of their time to participate in the experiment.

This research was partially sponsored by a Research Grant of the Research Foundation (FWO) to Alex Housen.

Notes

1. We use the term 'linguistic form' to refer to any aspect of linguistic code, ranging from graphemes, phonemes and morphemes over words and phrases to syntax and pragmatics. As is also apparent in this introductory section, the term 'form' is often used in opposition to 'meaning' (e.g. meaning-focused instruction versus form-focused instruction). In reality, however, the form-meaning dichotomy is oversimplistic, because conveying meaning obviously always requires some linguistic substance (e.g. words) and, conversely, all so-called formal elements also carry meaning, albeit to varying degrees (e.g. third person singular –s). An important question, then, is whether learners will also take in formal aspects of language when they primarily engage in processing its content or meaning.

2. We use 'uptake' as a synonym for 'what is learnt' or 'what is encoded in long-term memory'. This extends the behavioural definition of uptake found in interaction research, where uptake refers to learners' subsequent use of a recast (e.g. Lyster & Ranta, 1997). Although immediate uptake of a recast is produced in the next conversational turn already, it requires the learner to engage in maintenance rehearsal while awaiting her turn. This makes it likely that the recast will be embedded in long-term memory to some degree (also see Figure 1), hence our equating of uptake and learning. Loewen (2005) found that students' successful uptake was the best predictor of their post-test scores, which supports the idea that successful short-term uptake is associated with longer-term retention.

3. Robinson conceives of intake as "the cognitive registration (however temporary) of the input" (personal communication, 17th March 2008). Leow (1993) defines intake as "stored linguistic data that has been attended to by the L2 learner and may be used for immediate recognition" (p. 337–338). He operationalises intake as the participant's ability to recognise the target form immediately after the treatment. However, post-exposure recognition implicates that the target form has been stored in memory at least for the duration of the treatment, which does not necessarily follow from the allocation of focal attention (i.e.

noticing and its outcome, intake). One of our post-tests also consists of a multiple-choice recognition test, but in our view it already measures a weak form of receptive vocabulary knowledge (i.e. uptake), which presumes more elaborate cognitive processing than what is required for intake.

4. Leow (1997) uses a different terminology, but also distinguishes between a low and a high level of awareness.

5. In their first experiment, the more explicit treatment, thinking aloud was shown to slow down performance on a grammaticality judgment post-test.

6. To our knowledge, this research project is the first attempt to measure noticing by means of eye-tracking technology. Accordingly, one of the challenges is to gain a clear understanding of the strengths and weaknesses of this new approach. Given that we are only presenting the findings of an exploratory study we are not yet in a position to make more than tentative claims in this regard. Moreover, the principal aim of our research project was to investigate the occurrence of learner-initiated noticing and its connection with learning rather than to establish the methodological superiority of eye tracking over think-aloud protocols.

7. These studies should be distinguished from a complementary, also highly interesting strand of research that looks at noticing behaviour in general. The latter investigates the noticing of linguistic items or feedback as it arises spontaneously during a task or meaning-focused interaction (e.g. Williams, 2001; Loewen, 2004, 2005; Hanaoka, 2007). In such cases, the researcher does not know in advance what linguistic items the learners will take notice of.

8. We are aware that some formats of the vocabulary post-test (to be discussed later on in the text) do not exclude the possibility of guessing. The point we wished to make, however, is that none of the subjects would have an advantage due to some partial prior knowledge of the word.

9. Paragraphs were spread over different screens because every line of text is followed by two hard returns. This is a commonly used technique to avoid that possible drifts in gaze position would result in erroneous measurements. In case the recorded eye movements systematically float a little above or beneath a particular line (i.e. 'drift'), we can still reconstruct their actual position. If the lines of text followed each other immediately, we might wrongly think that the subject is looking at the line above or below the one that she is really fixating.

10. The length of the cue depended on the length of the target word. For up to nine-letter words, three letters were given. Longer target words were cued by four letters.

11. In order to be able to compute these measures, one must know the sequence of the fixations. This information is readily available in the data analysis programme, but cannot be derived from the static image above. The subject's eyes moved as follows: $[[336\ \text{msec}]_{\text{first fixation}} \rightarrow 200\ \text{msec} \rightarrow 260\ \text{msec} \rightarrow 180\ \text{msec} \rightarrow 412\ \text{msec} \rightarrow 252\ \text{msec}]_{\text{first run}} \rightarrow [208\ \text{msec} \rightarrow 180\ \text{msec} \rightarrow 184\ \text{msec}]_{\text{outside target region}} \rightarrow [180\ \text{msec} \rightarrow 156\ \text{msec}]_{\text{second run}}$.

12. We would like to thank Ronald Leow (personal communication, 31st March 2008) and the anonymous reviewer for this recommendation.

References

Alanen, Riikka. 1995. Input enhancement and rule presentation in second language acquisition. In Richard W. Schmidt (ed.) *Attention and awareness in foreign language learning.* Honolulu, HI: University of Hawai'i, pp. 259–302.

Allport, Alan. 1988. What concept of consciousness? In Anthony J. Marcel & Edoardo Bisiach (eds.) *Consciousness in contemporary science.* London: Clarendon Press, pp. 159–182.

Baars, Bernard J. 1988. *A cognitive theory of consciousness.* New York: Cambridge University Press.

Baars, Bernard J. 1997. In the theatre of consciousness. Global workspace theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies* 4(4): 292–309.

Baars, Bernard J. 2002. The conscious access hypothesis: origins and recent evidence. *TRENDS in Cognitive Sciences* 5(1): 47–52.

Boers, Frank and Seth Lindstromberg. 2009. *Optimizing a lexical approach to instructed second language acquisition*. Basingstoke, UK: Palgrave Macmillan.

Bowles, Melissa A. 2008. Task type and reactivity of verbal reports in SLA. A first look at a L2 task other than reading. *Studies in Second Language Acquistion* 30: 359–387.

Bowles, Melissa A. and Ronald P. Leow. 2005. Reactivity and type of verbal report in SLA research methodology. Expanding the scope of investigation. *Studies in Second Language Acquisition* 27: 415–440.

Doughty, Catherine J. 2001. Cognitive underpinnings of focus on form. In Peter Robinson (ed.) *Cognition and second language instruction.* Cambridge: Cambridge University Press, pp. 206–257.

Doughty, Catherine J. and Jessica Williams, eds. 1998. *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.

Ellis, Nick. 2002. Frequency effects in language processing. A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24: 143–188.

Ellis, Rod. 2003. *Task-based language learning and teaching.* Oxford: Oxford University Press.

Ellis, Rod, Helen Basturkmen, and Shawn Loewen. 2002. Doing focus on form. *System* 30: 419–435.

Ericsson, K. Anders and Herbert A. Simon. 1993. *Protocol analysis: verbal reports as data* (revised edition). Cambridge, MA: Bradford books/MIT Press.

Frazier, Lyn and Keith Rayner. 1982. Making and correcting errors during sentence comprehension: eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology.* 14: 178–210.

Gass, Susan and Alison Mackey. 2000. *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hanaoka, Osamu. 2007. Output, noticing, and learning: an investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research* 11(4): 459–479.

Hulstijn, Jan H. 2003. Incidental and intentional learning. In Catherine J. Doughty and Michael H. Long (eds.) *The handbook of second language acquisition.* New York: Blackwell Publishing, pp. 349–381.

Hulstijn, Jan H. and Wouter Hulstijn. 1984. Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning* 34(1): 23–43.

Izumi, Shinichi. 2002. Output, input enhancement, and the Noticing Hypothesis. An experimental study on ESL relativization. *Studies in Second Language Acquisition* 24: 541–577.

Jones, Francis R. 1992. A language-teaching machine: Input, uptake and output in the communicative classroom. *System* 20(2): 133–150.

Jourdenais, Renée. 2001. Cognition, instruction and protocol analysis. In Peter Robinson (ed.) *Cognition and second language instruction.* New York: Cambridge University Press, pp. 354–375.

Jourdenais, Renée, Mitsuhiko Ota, Stephanie Stauffer, Beverly Boyson, and Catherine J. Doughty. 1995. Does textual enhancement promote noticing? A think-aloud analysis. In Richard W. Schmidt (ed.) *Attention and awareness in foreign language learning.* Honolulu, HI: University of Hawai'i, pp. 183–216.

Kellogg, Ronald T. and Ruth. S. Dare. 1989. Explicit memory for unattended information. *Bulletin of the psychonomic society* 27: 409–412.

Kennedy, Alan. 1983. On looking into space. In Keith Rayner (ed.) *Eye movements in reading: perceptual and language processes.* New York: Academic Press, pp. 237–251.

Kennedy, Alan and Wayne S. Murray. 1987a. The components of reading time: eye movement patterns of good and poor readers. In J. Kevin O'Regan and Ariane Lévy-Schoen (eds.) *Eye movements: from physiology to cognition.* Amsterdam: North Holland, pp. 509–520.

Kennedy, Alan and Wayne S. Murray. 1987b. Spatial coordinates and reading: Comments on Monk. *Quarterly Journal of Experimental Psychology* 39A: 649–656.

Krashen, Stephen D. 1985. *The Input Hypothesis: Issues and implications*. London: Longman.

Krashen, Stephen D. 1989. We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis. *Modern Language Journal* 73: 440–464.

Laufer, Batia. 2005. Focus on Form in second language vocabulary acquisition. In Susan Foster-Cohen (ed.) *EUROSLA Yearbook 5*, Amsterdam / Philadelphia: John Benjamins, pp. 223–250.

Leow, Ronald P. 1993. To simplify or not to simplify: A look at intake. *Studies in Second Language Acquisition* 15: 333–355.

Leow, Ronald P. 1997. Attention, awareness, and foreign language behavior. *Language Learning* 47: 467–505.

Leow, Ronald P. 2000. A study of the role of awareness in foreign language behavior. Aware versus unaware learners. *Studies in Second Language Acquisition* 22: 557–584.

Leow, Ronald P. and Kara Morgan-Short. 2004. To think aloud or not to think aloud. The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition* 26: 35–57.

Loewen, Shawn. 2004. Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning* 54(1): 153–188.

Loewen, Shawn. 2005. Incidental focus on form and second language learning. *Studies in Second Language Acquisition* 27: 361–386.

Long, Michael H. 1991. Focus on form: A design feature in language teaching methodology. In Kees de Bot, Ralph B. Ginsberg, and Claire Kramsch (eds.) *Foreign Language Research in Cross-Cultural Perspective*. Amsterdam / Philadelphia: John Benjamins, pp. 39–52.

Lyster, Roy and Leila Ranta. 1997. Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition* 19: 37–66.

Mackey, Alison. 2006. Feedback, noticing and instructed second language learning. *Applied Linguistics* 27(3): 405–430.

Mackey, Alison, Jenefer Philp, Takako Egi, Akiko Fujii and Tomoaki Tatsumi. 2002. Individual differences in working memory, noticing of interactional feedback and L2 development. In Peter Robinson (ed.) *Individual differences and instructed language learning.* Philadelphia: John Benjamins, pp. 181–209.

Mangan, Bruce. 1993. Taking phenomenology seriously: the 'fringe' and its implications for cognitive research. *Consciousness and cognition* 2(2): 89–108.

Murray, Wayne S. and Alan Kennedy. 1988. Spatial coding in the processing of anaphor by good and poor readers: Evidence from eye movement analyses. *Quarterly Journal of Experimental Psychology* 40A: 693–718.

Nation, Paul. 2001. *Learning vocabulary in another language.* Cambridge: Cambridge University Press.

Nosofsky, Robert M. 1992. Exemplar-based approach to relating categorization, identification, and recognition. In Gregory Ashby (ed.) *Multidimensional models of perception and cognition.* Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 363–395.

Rayner, Keith. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124/3: 372–422.

Robinson, Peter. 1995. Attention, memory and the "noticing" hypothesis. *Language Learning* 45/2: 283–331.

Robinson, Peter. 1997a. Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition* 19: 223–247.

Robinson, Peter. 1997b. Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning* 47(1): 45–99.

Robinson, Peter. 2003. Attention and memory during SLA. In Catherine J. Doughty and Michael H. Long (eds.) *The handbook of second language acquisition.* New York: Blackwell Publishing, pp. 631–678.

Rosa, Elena and Ronald P. Leow. 2004. Awareness, different learning conditions, and L2 development. *Applied psycholinguistics* 25(2): 269–292.

Rosa, Elena and Michael D. O'Neill. 1999. Explicitness, intake, and the issue of awareness. Another piece to the puzzle. *Studies in Second Language Acquisition* 21: 511–556.

Sachs, Rebecca and Charlene Polio. 2007. Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition* 29: 67–100.
Sanz, Cristina, Hui-Ju Lin, Beatriz Lado, Harriet Wood Bowden, and Catherine A. Stafford. 2009. Concurrent verbalizations, pedagogical conditions, and reactivity: two CALL studies. *Language Learning* 59(1): 33–71.
Schmidt, Richard W. 1990. The role of consciousness in second language learning. *Applied linguistics* 11(2): 129–158.
Schmidt, Richard W. 1993. Awareness and second language acquisition. *Annual review of applied linguistics* 13: 206–226.
Schmidt, Richard W. 1994. Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA review* 11: 11–26.
Schmidt, Richard W. 1995. Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In Richard W. Schmidt (ed.) *Attention and awareness in foreign language learning.* Honolulu, HI: University of Hawai'i, pp. 1–63.
Schmidt, Richard W. 2001. Attention. In Peter Robinson (ed.) *Cognition and second language instruction*. Cambridge: Cambridge University Press, pp. 3–32.
Schmidt, Richard W. and Sylvia Frota. 1986. Developing basic conversational ability in a second language: a case study of an adult learner of Portuguese. In Richard R. Day (ed.) *Talking to learn: conversation in second language acquisition.* Rowley, MA: Newbury House, pp. 237–326.
Schmitt, Norbert. 2008. Instructed second language vocabulary learning. *Language Teaching Research* 12: 329–363.
Slimani, Assia. 1991. Evaluation of classroom interaction. In J. Charles Alderson and Alan Beretta (eds.) *Evaluating second language education.* Cambridge: Cambridge University Press, pp. 197–220.
Tomlin, Russell S. and Victor Villa. 1994. Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition* 16(2): 183–203.
Underwood, Geoffrey. 1976. *Attention and memory.* Oxford: Pergamon Press.
Underwood, Geoffrey. 1982. Attention and awareness in cognitive and motor skills. In Geoffrey Underwood (ed.) *Aspects of consciousness. Vol. 3: Awareness and self-awareness.* London: Academic Press.
Wesche, Marjorie and T. Sima Paribakht, eds. 1999. Incidental L2 vocabulary acquisition: theory, current research, and instructional implications. Special issue of *Studies in second language acquisition* 21(2).
Williams, Jessica. 2001. Learner-generated attention to form. *Language Learning* 51(s1): 303–346.
Williams, John. 2005. Learning without awareness. *Studies in Second Language Acquisition* 27: 269–304.
Wiseman, Richard. 2007. *A quirky look at our quirky species.* Retrieved August 21, 2008 from http://www.newscientist.com/channel/being-human/mg19426031.700-a-quirky-look-at-our-quirky-species.html.

Appendix

1. Example of a paragraph presented in the four conditions
   The target forms have been italicised for ease of reference. In the actual experiment they are printed in regular style like the other words.

   Condition I: the control condition

   > John Trinkaus of the City University of New York studies ordinary people going about their everyday lives. One of his specialities is the study of minor acts of dishonesty and antisocial behaviour. In his 25 years of research, he has found that one demographical group has come to stand out above all others as being most likely to push *boundaries* and

break rules. These are not teenagers nor football hooligans. They are women van drivers. Among other things, Trinkaus found that as many as 96 per cent of them break the speed limit, compared with 86 per cent of male offenders.

Condition II: pseudo word

John Trinkaus of the City University of New York studies ordinary people going about their everyday lives. One of his specialities is the study of minor acts of dishonesty and antisocial behaviour. In his 25 years of research, he has found that one demographical group has come to stand out above all others as being most likely to push *paniplines* and break rules. These are not teenagers nor football hooligans. They are women van drivers. Among other things, Trinkaus found that as many as 96 per cent of them break the speed limit, compared with 86 per cent of male offenders.

Condition III: pseudo word + existing word

John Trinkaus of the City University of New York studies ordinary people going about their everyday lives. One of his specialities is the study of minor acts of dishonesty and antisocial behaviour. In his 25 years of research, he has found that one demographical group has come to stand out above all others as being most likely to push *paniplines or boundaries* and break rules. These are not teenagers nor football hooligans. They are women van drivers. Among other things, Trinkaus found that as many as 96 per cent of them break the speed limit, compared with 86 per cent of male offenders.

Condition IV: existing word + pseudo word

John Trinkaus of the City University of New York studies ordinary people going about their everyday lives. One of his specialities is the study of minor acts of dishonesty and antisocial behaviour. In his 25 years of research, he has found that one demographical group has come to stand out above all others as being most likely to push *boundaries or paniplines* and break rules. These are not teenagers nor football hooligans. They are women van drivers. Among other things, Trinkaus found that as many as 96 per cent of them break the speed limit, compared with 86 per cent of male offenders.

(This paragraph is an adapted extract from an article in *New Scientist* magazine, 9 May 2007, by Richard Wiseman.)

2. Example of a screen display

The next three pages show an example of how the paragraphs are displayed on screen. Participants can move from one screen to the next by pressing a button on a response pad. They cannot return to an earlier screen. The target form, here *paniplines* (condition II), always appears somewhere in the middle of the second screen. (Paragraph adapted from Wiseman, 2007.)

John Trinkaus of the City University of New York studies ordinary people going about their everyday lives. One of his specialities is the study of minor acts of dishonesty and antisocial behaviour.

In his 25 years of research, he has found that one demographical group has come to stand out above all others as being most likely to push paniplines and break rules. These are not teenagers nor football hooligans.

They are women van drivers. Among other things, Trinkaus found that as many as 96 per cent of them break the speed limit, compared with 86 per cent of male offenders.

3. Examples of increased visual attention to the target word ('noticing') and the corresponding control condition ('no noticing')

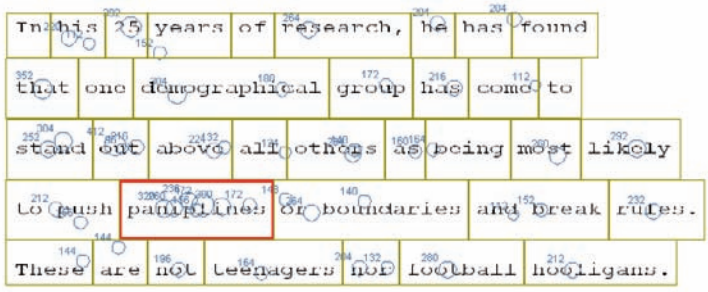


Figure 4. Increased visual attention to 'paniplines', interpreted as a noticing event (condition III)

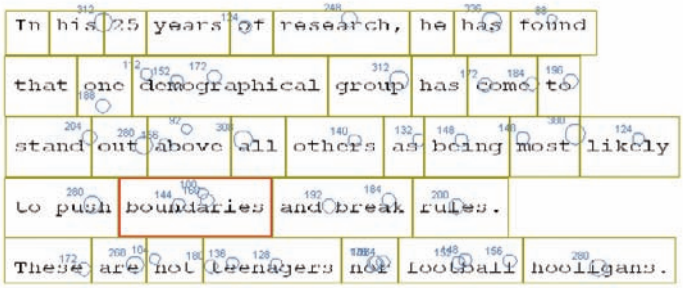


Figure 5. The control condition for Figures 3 and 4. 'Boundaries' receives far less attention than 'paniplines'. It is a familiar, unnoticed word

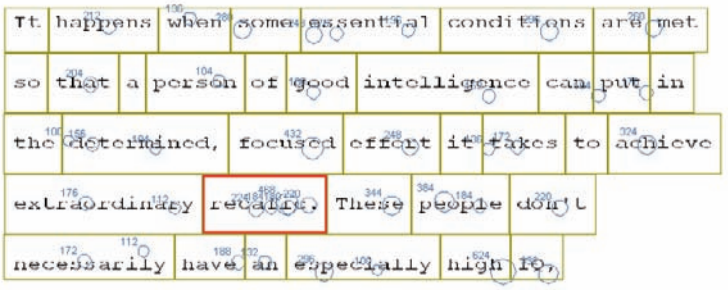


Figure 6. Increased visual attention to 'recafic', interpreted as a noticing event (condition II)

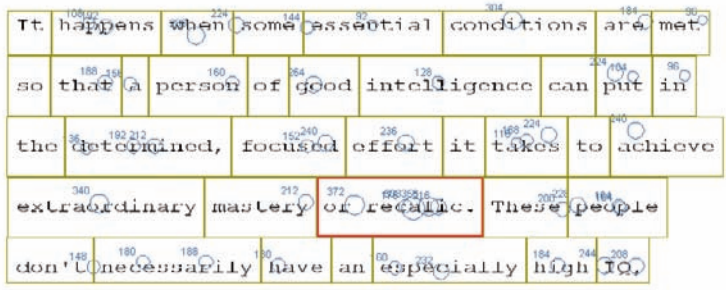


Figure 7. Increased visual attention to 'recafic', interpreted as a noticing event (condition IV)

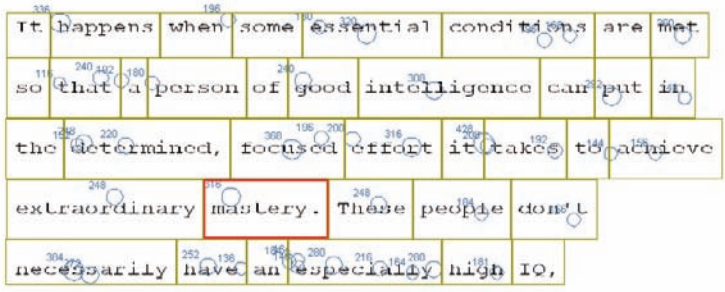


Figure 8. The control condition for Figures 6 and 7. 'Mastery' receives far less attention than 'recafic'. It is a familiar, unnoticed word

CHAPTER 11

Construal and the use of impersonalisation strategies in English and Spanish in an FLL context

Juana I. Marín Arrese
Universidad Complutense de Madrid

1. Introduction

Languages make use of different strategies for the mystification of agency, by backgrounding of the role of agency and/or obfuscation of the identity of the agent, which represent a continuum in the degree of defocusing of the agent in various constructions. The term 'impersonalisation strategies' is used to refer to the use of a variety of linguistic resources which allow for varying degrees of defocusing of the agent or cause, ranging from implicit reference to a specific agent, to the evocation of the type of agent involved, and finally to some abstract and schematic notion of causation (Marín Arrese, 2002).

Strategies of impersonalisation in English include the use of *be*-passives, resultatives, *get*-passives, unmarked intransitives (coding intrinsically and non-intrinsically spontaneous events, Kemmer, 1993), existentials, non-finites (*-ed*, infinitive and *-ing*) clauses, nominalizations, impersonal pronouns (*one, you, we, they…*), and miscellaneous lexical devices (metaphor, metonymy …). Spanish exhibits a more complex system of codification: agentless periphrastic *ser*-passive, resultative *estar*, *se*-passive, *se*-impersonal, inchoative or anticausative *se* (non-intrinsically spontaneous events), unmarked intransitive (intrinsically spontaneous events), modalised impersonal expressions (*hay que, urge*), existentials, non-finite clauses, nominalisations, impersonal use of pronouns (*uno, nosotros* …), and the use of metaphor, metonymy and other lexical devices. (Nedjalkov and Jaxontov, 1988; Shibatani, 1988; Gómez Torrego, 1992; Maldonado, 1999; Marín Arrese, 2002, 2003, *inter alia*).

This paper focuses on passives and other situation types related to the middle domain. In a comprehensive study on middle semantics,[1] Kemmer (1993: 7) defines the notion of situation types as "the sets of situational or semantic/pragmatic contexts that are systematically associated with a particular form of expression". Kemmer (1993: 142) establishes a distinction between middle situations, such as reflexive and reciprocal events, body action and cognition middles, which "all involve events occurring in the mind and/or body of human or at least animate entities", and related situation types such as spontaneous events, passive middle and impersonal middle, which "prototypically involve purely affected, often inanimate, entities".

The paper presents results of a case study on the use of strategies of impersonalisation by Spanish university students in translation tasks into English and Spanish of sentences involving defocusing of agency/cause. It is hypothesized that variation in the constructions used by Spanish students in the translations derives from differences in the construal of the event designated (Langacker, 2000). Other factors involved, it will be argued, are the following: degree of explicitness in the designation of the subject element in both languages; grammatical vs. pragmatic motivation for word order; and avoidance strategies involving the use of the periphrastic passive. That is, they derive from cognitive and pragmatic factors, and by the constraints imposed by the codification system in each language.

The purpose of this study is twofold: (i) the identification of qualitative and quantitative differences between the two languages in the use of impersonalisation strategies in the translation tasks; (ii) the explanation of the variation in the translation equivalents, due to differences in construal, and discourse-pragmatic differences.

This paper is organized as follows. Sections 1 and 2 describe impersonalisation strategies in English and Spanish. Section 3 explores the characterizing features of the passive, impersonal and spontaneous systems in the two languages. The case study is described in 4 and the results and discussion are presented in Section 5. The final section is devoted to the concluding remarks.

2. Impersonalisation strategies in English

2.1 Constructions with *Be*

Be-constructions are found in the semantic space of stative relational expressions, in the subjective resultative construction (*One foot was badly swollen*, A18.R2),[2] in resultatives of translational motion (*Now that the girls were gone with their young attendants*, A4.R1.), in objective resultatives or stative passives (*Even if America's burned - even if it's destroyed*, A14. R15-16), and in agented and agentless passives (Nedjalkov and Jaxontov, 1988; Langacker, 1991a; Déchaine, 1995; Marín Arrese, 1999). The periphrastic passive construction is perhaps the most frequent grammatical strategy of impersonalisation in English. In the vast majority of cases, the passive is agentless; according to Svartvik (1966) omission of the by-phrase ranges from 76% to 80%. Similarly, Marín Arrese (1993) found that agentless passives represented a percentage of 81.90% in contrast with 18.09% agented passives in the corpus studied.

In discourse, we find different degrees of defocusing and/or degrees of recoverability of the identity of the agent. In some cases, the underlying agent is recoverable from the preceding or following co-text. In others, it may be inferred on the basis of shared knowledge, or shared event or context models, which allows us to predict the type of agent characteristically involved in the event, though the identity of the agent is not recoverable to the point where unique reference can be established. Finally, there are cases where mystification of the identity of the agent is absolute (Marín Arrese, 1993, 1997). The following examples exemplify the scale of defocusing of agency.

(1) a. Kleopatra was swept by a red surge of anger. (A4.11)
b. He did not take the hand of Monsignor Campanati, since that was not proffered. (A6.3) (i.e. by Monsignor Campanati)
c. The core of the house was a rugged old fort going back to the civil wars: later the thatched house had been built beside it. (A4.14) (i.e. by the builders)
d. …, and somewhere a kettle was being filled. (A3.2) (i.e. by somebody)

From a Cognitive Grammar analysis, the passive construction involves three meaningful 'grammatical' elements: *be*, *-ed* and *by* (Langacker, 1991a, b). The passive variant of *be*, be2 or BEp, is a schematic process, which "imposes its processual profile on the composite expression", be+V-ed (Langacker, 1991a: 138). The meaning attributed to the passive variant of the past-participial morpheme, PERF3, is that of a shift in trajector/landmark alignment: the primary landmark of V becomes the passive trajector. Langacker (1991b: 203) notes that PERF3 enhances the salience of the "'terminal participant', i.e. one that lies downstream from another with respect to the flow of energy (or some analog thereof), and it is prominent by virtue of being made the relational figure (trajector)". In addition, PERF3 shares with the other variants of [PERF][3] the emphasis on the terminal stage of an event, the construal of the event as completed. Finally, the passive *by* "profiles a stative relation whose trajector is a schematic passive participle and whose landmark is the unspecified actor (i.e. the trajector of the verb stem on which the participle is formed)" (Langacker, 1991b: 337).

The passive schema evokes a process where the theme participant, the new trajector, is construed as the most prominent participant, and where the emphasis is on the terminal stage of the event. The agent is a non-focal participant, and may thus be left unspecified. In this respect, as Langacker (1991b: 333–336) observes, "the markedness of a passive or passive-like construction does not derive from profiling (…). Rather, it resides in the fact that the participant otherwise expected to be the subject is bypassed in favor of a less qualified candidate".

Langacker (2006: 127) also notes that the choice of a particular participant for trajector, and thus for focal prominence, "enhances the salience of those facets of the overall relationship that it anchors". Thus, although the active construction and its alternative alignment in the passive are equivalent in profiling, they differ in that the active highlights "the agent's activity", whereas the passive highlights "the core process which the theme thereby undergoes".

2.2 Constructions with *Get*

The verb *get* was originally associated with the meaning of 'obtain or attain' (c. 1200), which later extends to the meaning of 'procure for oneself' (c. 1300), and can be found with a dative or reflexive pronoun. This notion is linked to the association of responsibility of the theme participant. The verb *get* was also found to designate change effected in the position or state of the object (c. 1350), often co-occurring with a reflexive. The meaning later extended to 'cause or induce an action to be performed', as in the causative reflexive

(c. 1500) (OED, 1992). Responsibility in inducing a change-of-state is undoubtedly the basis for the extension to the *get*-passive.

Get-constructions are found in the causative, reciprocal/reflexive, inchoative, middle and passive domains. *Get* in causative (*We got him to sign the papers*) and causative-locative constructions (*He got her into the house*) (Givon and Yang, 1994) invokes the expression of causation falling somewhere between direct (*make*) and mediated (*cause*) cause (Cameron, 1990). There is also a correlation between passive and causative in so-called 'indirect passive constructions' (*He got his radio confiscated by the police*) (Shibatani, 1985). We also find reflexive-causative or 'complex reflexive passives' (*I got (myself) appointed to the governing board*) coded with *get*, as well as some instances of reciprocal situations and some middle situation types (body action middles) (Marín Arrese, 1993, 1999; Givon and Yang, 1994; Collins, 1996).

The aspectual features of 'mutation', 'change' or 'inception' are most obvious with *get* as a copula (Cameron, 1990), and with predicates designating spontaneous events (*The crustacean got fossilized*). This feature is also found in the category identified by Collins (1996) as 'psychological *get*-passives' (*get fed up with*), which is best subsumed under cognition and emotion middle situation types (Kemmer, 1993).

Get-passive events involve an external controlling agent, which is typically either non-identifiable or non-specific. In the *get*-passive, the theme participant may be viewed as holding some responsibility for the event, since it very often functions as catalyser (Barber, 1975), though it is not the primary 'initiator' of the event. It has been observed in the literature that the passive with *get* implies initiative and/or partial responsibility or causal involvement of the theme participant in the event.[4] This attribution of responsibility for the event favours the choice of referents with high inherent topicality as subjects (Marín Arrese, 1993, 1995, 1997). In this respect, it is interesting to note that in co-occurrence with deontic modals, necessity is attributed to the subject in the *get*-passive, thus enhancing the subject responsibility reading, whereas in the *be*-passive necessity is attributed to the agent (Lakoff, 1971: 156).

(2) a. Radicals must get arrested to prove their machismo.
b. Radicals must be arrested if we are to keep the Commies from overrunning the U.S.

The original emphasis on the initiative of the subject has extended to cases of 'generic', theoretical responsibility, especially in events of an adversative nature. The corresponding construction with the reflexive bears a more pointed suggestion of responsibility (Hatcher, 1949).

(3) a. He got killed in a car accident.
b. You'll get yourself killed one of these days.

Langacker (2000: 312–314) observes that the evolution of *get* in the passive domain involves "progressive attenuation of the degree of control exercised by its subject". Attenuation in subject control (from volitional participation to non-volitional responsibility) may further extend to cases where "the experiencer is no longer the subject per se, but rather an individual associated with the subject, the possessor" (c) and to cases where "the locus of the experience need not be overtly specified or clearly delimited" (d).

(4) a. Sue got (herself) appointed to the governing board.
b. Ralph got fired again.
c. All my books got stolen.
d. Another bank got robbed last night.

Conversely, the role of the agent participant in the *get*-passive appears to be subordinated, that is, "the agent does not completely dominate the situation" (Hatcher, 1949: 436). This probably accounts for the fact that the expression of an agent phrase is rare, particularly if it refers to a human (volitional) entity, though human agents with a low degree of individuation are found (Marín Arrese, 1999, 2003). Agents, when expressed, tend to be inanimate and non-identifiable.

(5) a. */?The woman got run over by John.
b. */?The window got broken by John.
c. The woman got run over by a drunken driver.
d. The woman got run over by a car.

2.3 Spontaneous situations

Languages without middle systems, like English, use 'heavy marking' in direct reciprocals and reflexives (*He saw himself in the mirror*), and unmarked intransitives ('zero marking') for naturally reciprocal events (*They embraced*) and for middle situation types (Kemmer, 1993). English uses the same verb form, so-called reversible predicates (labile alternation,[5] Haspelmath, 1993), to designate transitive (lexical causatives) and non-intrinsically spontaneous events (inchoative or anticausative construction).

(6) a. Mary/The key/The wind *opened* the door.
b. The door *opened*.

In spontaneous situations, intentional actions are presented as 'events' occurring spontaneously. Spontaneous events typically involve change of state, both physical and psychological, and change of position or location, associated with the following (Kemmer, 1993: 269–270) (my examples in Spanish):

i. Animate beings: physiological processes (grow, *crecer*); non-translational motion (spin, *girar*).
ii. Inanimate beings: change of position/location (sink, *hundirse*); change of shape (shorten, *acortarse*); physico-chemical change (melt, *derretirse*); partial disruption of an object's integrity (break, *romperse*); global disruption of object's integrity (disintegrate, *desintegrarse*); existential change (dissolve, *disolverse*); object-specific (container: fill, *llenarse*) (port: open, *abrirse*); property of activities (finish, *terminarse*); inanimate state (split, *dividirse/partirse*).
iii. Inanimate state/activity impinging on human sensory organs: visual (shine, *brillar*); auditory (ring, *sonar*), tactile (scratch, *picar*).
iv. Property of inanimate entity filtered through cognitive interpretative mechanisms: auditory (sound, *sonar*), visual (appear, *aparecer*).

Most of these event types, as Kemmer (1993:142) notes, involve *intrinsically spontaneous events*, that is, those viewed as "occurring without direct initiation by an Agent" or any external cause. There are some cases, however, which we might term *non-intrinsically spontaneous events*, where the change-of-state event is construed as involving some human agent or at least evokes some external schematic cause.

Non-intrinsically spontaneous events, coded with the inchoative construction, involve an effector as theme participant. The event is construed as taking place spontaneously in that the theme is attributed a certain degree of control over the event. At the same time, they evoke the presence of 'external energy input' as some covert force (agent, instrument, natural force or circumstance), however abstract or schematic it may be (Langacker, 1991b). As Langacker (1991b: 333–36) notes, in non-intrinsically spontaneous events, "only the change of state is profiled, the implicit reference to an agent is non-salient and may be absent altogether".

(7) a. Their withdrawal *had* already *begun*. (AEg04)
b. The Stock Exchange *will reopen* this morning and a gold auction will go ahead. (BEt08)

The central type class of predicate found in the inchoative construction is a causative verb involving change of state (physical and psychological) and change of position. However, causative verbs such as *build, clean, cut, paint, wash,* etc., also involving a change of state, are not compatible with the inchoative construction, since the type of action they describe, which includes some specification of the means or manner of performance, requires a human volitional agent. The absence of agent-oriented meaning components appears to be crucial in the construal of the event as spontaneous, thus licensing the inchoative construction. The *get*-passive makes available an inchoative reading in the case of events involving agent-oriented components. The parallel between the two constructions lies in the fact that both focus on the change of state.

(8) a. A local architect *built* the new library./*The new library built./The new library got built.
b. The baker *cut* the bread./*The bread cut./ The bread *got* cut.
c. She finally read the book/*The book read./The book *got* read.

In non-intrinsically spontaneous events, the conception of the theme as exerting a certain degree of control can be reinforced by the presence of an adverbial modifier *by itself* (i.e., 'without outside help'). On the other hand, intrinsically spontaneous events, which are construed as inherently autonomous, are incompatible with the expression of the notion 'without outside help', since the event is necessarily internally initiated and controlled.

(9) a. The door *opened* by itself.
b. The boat *sank* by itself.
c. *The roses *blossomed* by themselves.

In intrinsically spontaneous events, there is no implicit reference to any agent or external energy input; they are construed as taking place spontaneously and internally induced, without the direct initiation of an external cause. As Langacker (2000: 84) notes, the events are construed as autonomous in that "we are able to conceptualize them independently

of any agent or initiative force that might be responsible for them", and thus the theme element is wholly responsible for the event. We find a whole range of unaccusative predicates[6] (Perlmutter, 1978) coding this situation type (Marín Arrese, 2002).

(10) a. By mid-1999 a new stratagem *emerged* when Frei had a long telephone conversation with Blair in which the Chilean sought help in getting Pinochet released back to Chile on humanitarian grounds. (BEg03)
b. The Home Secretary intervened on the night before the resignation after Mr Mandelson had cast doubt on whether the call *had happened*. (CEt02)

Intransitive events are also designated by means of unergative predicates[7] (Perlmutter, 1978), where the inanimate entity coded as subject is construed as endowed with 'volition' and able to act autonomously (Marín Arrese, 2002).

(11) a. The FSA admits it has no idea how much MRM *enters* the country. (BEt01)
b. The e-mail and many other papers *have gone* before Sir Anthony Hammond, the QC asked by Mr Blair to investigate the affair on the day Mr Mandelson resigned for a second time. (CEt02)

The designation of intrinsically spontaneous events disallows the use of causative transitive predicates (b). Where causation is referred to, it typically involves natural forces, which are not perceptually salient causes. The expression of external cause requires the use of the grammatical causative construction (c), and thus the causative view is marked. It has been argued that the impossibility of expressing the causative event by means of a lexical causative derives from the presence of "internal control" exerted by the entity engaged in the event (Smith, 1970: 170). Sometimes the use of *get* alternates with the intransitive (d), allowing for the evocation of some form of external force.

(12) a. The rosebush *blossomed*.
b. *Spring *blossomed* the rosebush.
c. Spring *made* the rosebush blossom.
d. The crustacean (*got*) fossilized.

2.4 Other strategies of impersonalisation

Strategies of avoidance of individual specification of the agent in impersonal situations include the use of indefinite pronouns (*one, some, …*) and of personal pronouns with generic or uncertain reference (*they, you, we, …*), thus allowing for "reduced participant identifiability" (Langacker, 2000: 40).

(13) …– but written in late 1938 after Munich when by now, *you* would have thought the world was tumultuous in its desire to act. (ABst: 221)

Impersonal or generic and vague uses are also attained by the use of indefinite pronouns, *everybody, nobody*, *anybody*, or by the use of 'inclusive *we*' in reference to "an incompletely defined collectivity that includes the speaker and one or more others, without specifying who they are" (Kitagawa and Lehrer, 1990: 745).

(14) *We* all know that there are many other influences on our children's development beyond school, and *we* all know I think that education must look at the whole picture.
(GBsp: 296)

Existentials involve a non-causal, non-eventive construal. The event is not construed as an interaction between participants but as a location or abstract setting (Langacker, 1991b). As Langacker (2000: 41) notes, "the archetypal basis for transitivity is an energetic interaction, and an interaction only holds between participants".

(15) After the accident *there was* a news blackout regarding the princess's condition, …
(TransSp-Ex1)

Nominalizations represent a step further in impersonalisation, since the actional component is obscured, and the event is presented as an 'abstract thing'. Nominalization, in Langacker's words (2000: 86), represents conceptual reification of the way an event is construed: "the event or process as a whole is construed as an abstract thing and is profiled by the nominal expression". Choices in the level of specificity in describing the actions may also result in the use of abstract nouns denoting events, rather than intentional actions.

(16) a. …but the situation worsened as the night wore on leading to *reports* of her death at 5am this morning. (TransSp-N1)
b. … *the persistent hounding* of the princess and her privacy *by* photographers.'
(TransSp-N2)

3. Impersonalisation strategies in Spanish

3.1 Constructions with *Ser* and *Estar*

The periphrastic construction with *ser* is found in process passives (a). The construction with *estar* codes the stative passive or object resultative construction (b), and the subjective resultative (c). In some cases, motion, *ir* (d), and result verbs, *quedar* (e), are found in non-prototypical and resultative passives (Marín Arrese, 1995).

(17) a. … *y yo había* ***sido*** *salvada del naufragio …* (B4.10)
… and I had been saved from the shipwreck…
b. **Estaba** prohibida la lectura de periódicos y, … (B6.E1.)
Reading newspapers was forbidden, and…
c. ***Está*** hinchado.
It is swollen.
d. … *la propuesta* ***va*** *dirigida contra …* (CSv04)
…the proposal is directed against …
e. *Ha* ***quedado*** *demostrado que …*
It has been proved that …

Suppression of the volitional agent in the periphrastic passive is also a matter of degree, as the following examples show.

(18) a. *Estas teorías han sido ampliamente alentadas* ***por*** *el padre de Dodi, el multimillonario Mohamed al Fayed.* (*El País*, 15/12/2006)
These theories have been widely encouraged by Dodi's father, the multimillionaire Mohamed al Fayed.

b. *El informe ha sido elaborado durante tres años* ***bajo la dirección de*** *sir John Stevens, ex jefe de Scotland Yard,…* (*El País*, 15/12/2006)
The report has been worked on for three years under the supervision of sir John Stevens, former head of Scotland Yard,…

c. *Ambos extremos nunca han sido confirmados.* (TransE-P4)
Both facts have never been confirmed.

3.2 Constructions with *Se*

The middle marker (MM) *se* is attested in examples representing all the situation types comprising the middle domain in modern Spanish (Gómez Torrego, 1992; Marín Arrese, 2003). Two-form languages typically use 'heavy marking' for reflexive and reciprocal situations and 'light marking', *se*, for middle situation types and other related types (Kemmer, 1993). Spanish uses 'light marking', the MM *se*, co-occurring with 'heavy marking' in the form of reflexives, *a sí mismo* (to oneself), and reciprocals, *el uno al otro* (one another). With indirect reflexive and indirect middle, only the 'light' marker *se* is used. Indirect reciprocals, naturally reciprocal events, and naturally collective actions also use only the 'light' marker se. The MM *se* is also found in body action middles, such as grooming or body care, change in body posture, nontranslational motion and translational motion. The marker *se* is likewise found in middle situations comprising mental events of emotion, cognition and perception, and emotive speech actions (Marín Arrese, 2003).

Concerning situation types related to the middle domain, the MM *se* is found in the passive middle, the impersonal, and the inchoative or anticausative construction coding non-intrinsically spontaneous events.

3.2.1 *Se-passive*

In the *se*-passive, the theme participant assumes most of the characteristic features of syntactic subjects (verbal agreement, word order). The trajector or subject is an affected entity which neither initiates nor enables the event to take place. A subtype with a clausal subject is also found.

(19) a. *Esas cosas* ***se*** *notan.*
Those things can be perceived.

b. ***Se*** *cree que las negociaciones tendrán éxito. Al menos, eso* ***se*** *cree.*
It is believed that the negotiations will be successful. At least, that is believed.

Word order in Spanish is to a great extent pragmatically determined, so that the trajector is typically found in pre-verbal position when it designates a referential, identifiable entity. In the case of non-identifiable, non-referential or generic concepts, the trajector is very often postposed. Heavy subjects also tend to be postposed for reasons of end-weight.

(20) a. *Según el ministro de Asuntos Exteriores, Josep Piqué, "la hipótesis más segura es que* ***el submarino se*** repare en Gibraltar y se vaya lo antes posible". (BSv05)
According to Foreign Secretary, Josep Piqué, "the most likely hypothesis is that the submarine will be repaired in Gibraltar and that it will leave as soon as possible".

b. *Asimismo* ***se*** *han analizado* ***documentos de la CNMV y de otros organismos o sociedades como Gescartera****…*. (CSa08)
Documents from the CNMV and other institutions or companies such as Gescartera have also been analysed …

The *se*-passive and the *se*-impersonal invoke an implicit external agent who controls the action. Though the agent typically remains implicit, there appears to be a tendency in modern Spanish, specially in the spoken language, to allow the use of an expressed agent.[8] Varying degrees of backgrounding or obfuscation of the role of agency by means of locative expressions and by the use of deictics are also found.

(21) a. *…,* ***se*** *ha mentido* ***por*** *el presidente del Gobierno respecto a la existencia de dictámenes del Consejo de Seguridad Nuclear y* ***se*** *ha puesto en riesgo a los habitantes del Campo de Gibraltar*. (BSp04)
…, the President of Government has lied (it has been lied by the president of Government) about the existence of reports by the Council for Nuclear Security and the inhabitants of the Campo de Gibraltar have been endangered.

b. *Como consecuencia de todos estos contactos,* ***se*** *estima* ***desde*** *la presidencia que pueden ser aportadas algunas ideas que resulten útiles …* (BSa09)
As a consequence of all these contacts, it is estimated by (from) the presidency that some useful ideas may be provided …

3.2.2 *Se-impersonal*

Kemmer (1993:148) notes that "the impersonal situation type subsumes both situations in which there is an affected entity […], and those in which there is no particular affected entity". In Spanish, the use of the MM *se* is found in an impersonal, or 'non-promotional' passive construction (Givon, 1990), where the verb takes an invariant 3SG form, in agreement with some non-specific, potential or generic human agent evoked. The endpoint nominal designating the affected entity is in the accusative. When the theme refers to a human, animate or personalized entity, expressed as a full NP, it is case-marked with *a*. When it is expressed pronominally, it takes the form of an object pronoun. A curious blend, (b), was found in Marín (2002) of plural verbal agreement with the theme participant, as in 'promotional' passives, together with case-marking of the animate theme participant, as in 'non-promotional' passives.

(22) a. ***Se*** *atendió* ***a*** *los enfermos.*
The sick patients were attended to.

b. *También* ***se han*** *arrestado* ***a*** *cinco activistas 'legales', entre ellos…* (CSa06)
Five illegal activists have also been arrested, among them …

The form *se* has also extended to an impersonal construction with no reference to an Endpoint participant, with generic meaning and with intransitive and copulative or semi-copulative verbs.

(23) a. **Se** vive bien en Madrid.
One lives well in Madrid
b. *¡Qué bien **se** está aquí!*
How nice it is here!/How well one is here!

3.3 Spontaneous events

Spanish makes a formal distinction between the two types of spontaneous events. Non-intrinsically spontaneous events are marked with the MM *se* (b), whereas intrinsically spontaneous events tend to receive no marking (c).

(24) a. *Maria/La llave/El viento **abrió** la puerta.*
Mary/The key/The wind opened the door.
b. *La puerta **se abrió** de repente.*
The door opened suddenly.
c. *El rosal floreció.*
The rosebush blossomed.

In non-intrinsically spontaneous events, degree of control of the theme participant may be reinforced by the inclusion of an adverbial modifier *él/ella solo/a*, *por sí solo/mismo* ('by him/herself/itself), as in English.

(25) a. *La puerta **se** abrió por sí misma.*
The door opened by itself.
b. *El barco **se** hundió por sí solo.*
The boat sank by itself.

In some cases, there is potential ambiguity between an inchoative or a passive reading, due to the continuum from external ultimate causes (implicit agent) in the *se*-passive, to direct/immediate causes and internal control in the spontaneous event (cf. Keyser and Roper, 1984; Marín Arrese, 2003).

(26) a. *El barco **se** hundió intencionadamente.*
The boat was sunk intentionally.
b. *El barco **se** hundió a causa de la explosión.*
The boat sank due to the explosion.
c. *El barco **se** hundió por sí solo.*
The boat sank by itself.

With causative verbs involving agent-oriented components (corresponding to 'non-reversible' verbs in English), and when the theme is singular (3SG verbal agreement) and inanimate (without case-marking), the construction with *se* may be interpreted either as passive or impersonal. In the latter case, the nominal element typically occupies post-verbal position.

(27) a. *El documento **se** leyó.*
The document got/was read.
b. ***Se** cortó el pan.*
The bread was/got cut.

Intrinsically spontaneous events are usually coded by an unmarked intransitive construction, with both unaccusative and unergative predicates (Perlmutter, 1978).

(28) a. *El terrorismo callejero en el País Vasco no **cesa**.* (CSa03)
Terrorism in the streets does not cease in the Basque country.
b. *…, y considera imprescindible que **desaparezca** inmediatamente la violencia que vive la zona.* (ASv02)
…, and considers that it is essential that the violence the area is experimenting should disappear immediately.
c. "*Las excusas **llegan** demasiado tarde*" (ASp07)
"Excuses arrive too late"

Variation within and across languages regarding the types of predicates found in intrinsically and non-intrinsically spontaneous events is often due to differences in the way events are construed as either basically internally or externally caused. Some examples of verbs expressing events which are conceptualized as internally caused in English, and either internally or externally caused in Spanish are the following: *rust* vs. *oxidar/se*, *wither* vs. *marchitar/se*, *swell* vs. *hinchar/se*, *rot* vs. *pudrir/se* (Mendikoetxea, 1999). These verbs describe processes inherent to the entities involved which are triggered by certain external conditions. Whereas in English the crucial characteristic is the inherent process of the entity, in Spanish the event may be construed as internally caused (a), or as involving some form of external energy input (b). The lexical causative is allowed for the expression of direct/immediate causes involving natural forces, but not ultimate causes (agents).

(29) a. *Las flores **se** marchitaron.*
The flowers withered.
b. *El calor marchitó las flores.*
*The heat withered the flowers.
c. *El jardinero/El calor hizo que **se** marchitaran las flores.*
The gardener/The heat made the flowers wither/caused the flowers to wither.

Reference to an animate agent requires the use of the causative periphrasis (c). Indirectness of causation is reflected in the use of a subordinate clause with the verb in the subjunctive rather than an infinitive form. The presence of the MM *se* also provides the notion of an autonomous process, even though it might be triggered indirectly by an external agent.

3.4 Other grammatical strategies of impersonalisation

In Spanish we also find the impersonal *uno*, and impersonal or vague uses of personal pronouns (*nosotros, vosotros, ellos*) or indefinite pronouns (*todos*). In most cases, however, we find no explicit subject element (Spanish is a pro-drop language), but simply 1PL, 2PL or 3PL agreement marking on the verb (Marín Arrese, 2002).

(30) a. "*En este país **todos** sabemos qué es lo que tenemos que hacer para …*" (CSp02)
"In this country we all know what we have to do to in order to … "
b. *… donde le **informaron** que …* (CSv03)
…where they informed him that …

As in the case of English, other impersonalisation strategies in Spanish include existentials, and nominalizations (Marín Arrese, 2002).

(31) a. *No **hubo** torturas.* (BSv03)
There was no torture/He was not tortured.

b. ***La selección** del jurado de la investigación judicial sobre la muerte de la princesa y su entonces novio Dodi Al Fayed se inició el jueves pasado en un tribunal de Londres.* (TransE-N1)
The choice of the jury for the judicial inquiry concerning the death of the Princess and her then boyfriend Dodi Al Fayed began last Thursday in a Court in London.

4. Defocusing of agency

The various situation types of reduced transitivity described above may be characterized in terms of a set of parameters: (i) the degree to which we may identify a salient cause and a salient effect, which relates to the degree of distinguishability of participants in the event; (ii) whether the source of energy input is construed as internal or external (Levin and Rappaport, 1995); and (iii) the degree of control that the thematic participant exerts.

Kemmer (1994: 211) observes that in coding transitivity distinctions the general conceptual dimension involved is that of *relative elaboration of events*, which "can be thought of as the degree to which different schematic aspects of a situation are separated out and viewed as distinct by the speaker". Subsumed within this conceptual domain is the semantic property *degree of distinguishability of participants*, that is, "the degree to which a single physico-mental entity is conceptually distinguished into separate participants" (Kemmer, 1994: 206). In the middle domain, the reflexive situation type evokes the type of event where the Initiating and Endpoint facets of the event are normally conceptualized as two distinct participants, but where the Initiator is actually acting on itself as Endpoint. Thus it evokes two semantic roles, but they are conflated in a single participant. In the case of related situation types such as the passive middle the Initiator and Endpoint roles are distinguished, though typically no reference is made to the Initiator entity. Spontaneous events are semantically middle, since the subject is not only an Endpoint element but is also construed as an initiating entity, for which I will be using the term effector[9]. In the impersonal situation type the event may or may not involve an Endpoint entity; in the latter case the situation would approximate that of ø-participant events. The following continuum of individuation of participants, based on Kemmer (1994: 209), illustrates these differences in the various situation types.

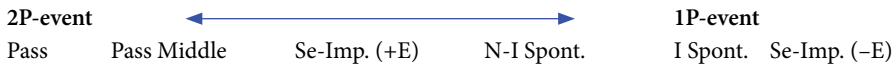


Figure 1. Degree of distinguishability of participants

In addition to the distinguishability of Initiator and Endpoint facets of the event, degree of transitivity is also inherently linked to volitional control and to source of energy input. Levin and Rappaport (1995: 94) propose the distinction between "internally vs. externally caused eventualities". Internal causation involves "causation initiated by, but also residing in, the single argument and hence dependent on its properties". Events that are internally caused may be agentive, that is, involving an intrinsic capacity of the agent, which s/he can control voluntarily (*laugh, stretch*), or they may be non-agentive, as in the case of existential changes (*appear, disappear*), or other events involving inanimate arguments where the notion of control would not apply, such as intrinsically spontaneous events (*blossom, wither*) (Kemmer, 1993). Levin and Rappaport (1995: 107) argue that in externally caused events, "the external cause argument sets the eventuality in motion, but it is not necessarily involved in seeing it through". External causation may involve agents, natural forces, instruments or circumstances, which may or may not be expressed. The inchoative or anticausative construction is characterized by "a complete lack of specification of the causing event", and as a result the cause argument is left unexpressed.

Elaborating on Levin and Rappaport's (1995) notions, I make a distinction in the construal of the event as involving different degrees of external vs. internal energy input. This feature refers to the degree to which the energy for the initiation of the event is generated either externally or internally, i.e. whether the theme participant is the initiator of the event or whether the source of energy originates in some other participant, which is not at the nucleus of the event. At one end of the continuum, the event may be construed as caused and under the control of some external agent, as in the passive. At the other end, in spontaneous events, it is conceptualised as autonomous and involving a non-salient schematic cause or as inherently generated.

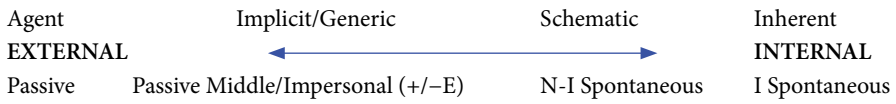


Figure 2. Source of energy input

Regarding the feature of degree of control of the theme participant, spontaneous events involve an effector theme participant, which conflates the roles of Initiator and Endpoint. In the *get*-passive, which is a form of passive middle in English, the subject is construed as a catalyser of the action. The agentless passive designates externally caused events, where the theme participant is the endpoint, a purely affected entity.

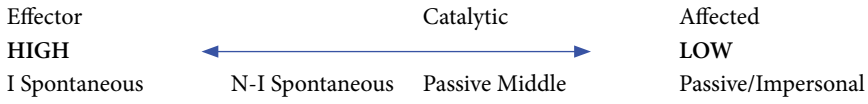


Figure 3. Degree of control of theme participant

As Enger and Nesset (2001) note, the schemas for the passive and middle systems subsume a number of functions, instantiated in the various constructions we have examined. The main characterizing features of the systems in English and Spanish may be summarized as in Tables 1 and 2 (adapted from Marín Arrese, 2003). The features are the following:

A. Degree of distinguishability of participants
B. Energy input
C. Salience of agency/cause
D. Degree of control/autonomy of theme participant
E. Role of theme participant
F. Emphasis on terminal stage of event

Table 1. Characterizing properties of passive and middle systems in English

ENGLISH	Passive	Passive GET	N-I. Spont.-Ø	I. Spont.-Ø
A	High	Medium	Low	Low
B	External	External	External/Internal	Internal
C	High	Medium/Low	Schematic	Inherent
D	Low	Medium	Medium	High
E	Affected	Catalyser	Effector	Effector
F	+	–	–	–

In the network for English the feature energy input sets apart the spontaneous situation types from the passives types. As regards the role of the theme element, in the *get*-passive the subject often functions as catalyser. This distinctive feature seems to motivate the presence of *get* in some middle situations. As we have seen, the *get*-passive is akin to the *se*-passive and to non-intrinsically spontaneous events. The emphasis on the terminal stage of the event is a characteristic feature of the *be*-passive in contrast with the other situation types.

Table 2. Characterizing properties of passive and middle systems in Spanish

SPANISH	Passive SER	Passive SE	Impersonal SE	N-I. Spont.-SE	I. Spont.-Ø
A	High	Medium	Medium	Low	None
B	External	External	External	External/Internal	Internal
C	High	Low	Opaque /Generic	Schematic	Inherent
D	Low	Low	Low/ø	Medium	High
E	Affected	Affected	Affected	Effector	Effector
F	+	–	–	–	–

In Spanish, the various situation types all share a distinctive feature, the fact that the theme participant is affected by the verbal action, that is, it is the locus of the event's effects, though in the spontaneous situation types its role as effector is more salient. The feature energy input sets the passive, *se*-passive and *se*-impersonal apart from the spontaneous situation

types. There are also distinct differences in degree of salience of the agent or cause. Finally, the focus on the terminal stage of the event distinguishes the periphrastic passive from the constructions with *se* and the intransitive coding intrinsically spontaneous events.

5. The case study

5.1 The subjects and the translation tasks

The subjects were a group of 4th and 5th year university students of English Philology, native speakers of Spanish, in a foreign language learning (FLL) context, who were asked to carry out two translation tasks, a translation into Spanish and a translation into English, of the following texts:

i. TransSp: *Di reported dead with Dodi in Paris car crash*, by Hal Austin, Peter Hooley and Daniel John, The Observer, Sunday August 31, 1997 (855 words).
ii. TransE: *Un jurado indaga cómo murió Diana de Gales*, by Mábel Galaz, El País, Tuesday, 2nd October 2007 (456 words).

The tasks for the Spanish learners of English thus involved both translating from the original text in a foreign language (English) into their native language, and from a text in their native language (Spanish) into a foreign language. The students were not given any guidance or any clue with respect to the aim of the analysis, but were simply asked to translate the whole text as a classroom exercise.

There were no *get*-passives in the English text, so the students were asked to translate 5 sentences with *get*-passives, randomly selected from my corpus of passives (Marín Arrese, 1993). Since the translation tasks were administered on different dates, the number of students differs for each task: TransSp: 23 students; TransE: 15 students; Get-pass: 17 students.

5.2 The data and the analysis

Before the translation task was administered, four sets of target sentences were selected for analysis in the texts.

5.2.1 *Be-passives*

In the English text, the target sentences for translation into Spanish are the following *be*-passives (P):

1. Diana, Princess of Wales, was reported <P1> to have died in a road crash…
2. Dodi Fayed, was also killed <P2>.
3. The accident happened as their limousine was allegedly chased <P3> through the west of Paris by paparazzi
4. Last night, the Princess was taken <P4> to the Pitie-Salpetriere Hospital …
5. The British ambassador, Sir Michael Jay, was summoned <P5> to the hospital.

6. The Harrods heir and the driver of the limousine were believed <P6> to have died instantly, according to police sources.
7. Prince Charles, …, was told <P7> of the accident,
8. In London the Prime Minister, Tony Blair, was woken <P8> to be told of the accident.
9. … that this accident was in part caused <P9> by the persistent hounding of the princess and her privacy by photographers.
10. US President Bill Clinton and his wife, Hillary, were also informed <P10> about Diana's accident.

5.2.2 *Get-passives*

As mentioned above, the following *get*-passive (GP) target sentences were selected from my corpus of passives:

1. During our visit to the zoo, the camera got mislaid <GP1>. (A3.G2, 162)
2. The old lady could have got mugged and killed <GP2>… (A12.G1, 239)
3. I got arrested <GP3> in Montreal last year, (A13.G1, 618)
4. That article never got printed <GP4> in The Bugle. (A17.G1, 94)
5. He got caught up <GP5> in the criminal-justice system in New York. (A13.G2, 621)

5.2.3 *Se-passives*

In the Spanish text, the following *se*-passives (SE-P) were selected for translation into English:

1. *… todavía se escucha* <SEP1> *la música del concierto…*
2. *Se prevé* <SEP2> *que dure entre cuatro y seis meses.*
3. *No es un proceso, ni se pronunciará* <SEP3> *condena.*
4. *La selección del jurado de la investigación judicial sobre la muerte de la princesa y su entonces novio Dodi Al Fayed se inició* <SEP4> *el jueves pasado…*

5.2.4 *Ser-passive*

The following periphrastic *ser*-passive (S-P) sentences were selected for translation into English:

1. *Unas 200 personas fueron citadas* <S-P1> *ante el tribunal …*
2. *Los llamados fueron elegidos* <S-P2> *de una lista del registro electoral …*
3. *, gran parte del cual será sufragado* <S-P3> *por los contribuyentes.*
4. *Ambos extremos nunca han sido confirmados* <S-P4>.

5.3 The analysis

For each target sentence, in each language, I listed the translation equivalents used by each student, and grouped them according to the different constructions chosen, in order to identify the preferred translation equivalents for the *be*-passives, *get*-passives, *ser*-passives and *se*-passives.

The analysis also focused on whether the choice of translation equivalents involved variation in the following features, and whether that could be attributed to differences in construal of the events designated: (i) the degree of defocusing of agency, (ii) perspective, and (iii) aspect and dynamicity.

5.3.1 *Degree of defocusing of agency*

The degree of defocusing of agency is the degree to which there is more salience of agency/cause or whether the agent evoked is generic or schematic, or no external agency is involved. In this respect, it was relevant to note the extent to which the translation equivalents were similar or different in degree of defocusing of agency with regard to the original sentences.

5.3.2 *Perspective*

Perspective involves both choice of trajector and position of theme. For each translation equivalent, an annotation was made regarding whether the translation maintained the same trajector (S-Tr) or whether a change in trajector (D-Tr) was effected. The same type of annotation was made for same (S-Th) or different (D-Th) thematic element.

5.3.3 *Aspect and dynamicity*

Finally, *aspect and dynamicity* pertain to the degree to which the event is construed as involving a more dynamic on-going process, one more akin to an active construction where there is more emphasis on some agent's activity, or whether the event is construed as completed, and the emphasis is on the terminal stage of an event. In this case, a note was made of the degree to which the same aspectual features were preserved in the translation equivalents.

6. Results and discussion

6.1 Translation into Spanish

In the first task, the translation of *be*-passives into Spanish, the choice of the *ser*-passive was the preferred translation equivalent (41.7%). However, a whole range of other constructions was also found: *se*-passive, *se*-impersonal, *se*-anticausative or inchoative (non-intrinsically spontaneous events), impersonal, intransitive (intrinsically spontaneous events), some examples of *se* coding middle situation types, active transitive, and others.

(32) 8. In London the Prime Minister, Tony Blair, **was woken** <P8> **to be told** of the accident.
   a. *Ser*-pass: …, *el Primer Ministro, Tony Blair,* ***fue*** *despertado para informarle* ..
   b. *Se*-imp: … ***se*** *despertó* ***a****l Primer Ministro… para informarle …*
   c. *Se*-middle: … *el Primer Ministro, …,* ***se*** *despertó con la noticia*
   d. Imp: … *al Primer Ministro … le despertaron para contarle …*
   e. Act.Intrans.: … *el Primer Ministro, …, despertaba con la noticia…*

In examples involving raising-constructions which cannot co-occur with passivisation with *ser* in Spanish, as in target sentence 6 in (33), a whole range of constructions were found. Translation equivalents also involved differences in choice of trajector and thematic position. In some cases, only thematic position is maintained, as in (b) and (c) below. In others neither trajector nor thematic position is maintained, as in (d).

(33) 6. The Harrods heir and the driver of the limousine **were believed** <P6> to have died instantly, according to police sources.
   a. *Se*-pass: ***Se** cree que el heredero de Harrods y el conductor … murieron …*
   b. *Se*-imp: ***Al** heredero … y **al** conductor … **se les** dio por muertos…*
   c. Imp: ***Al** heredero … y **al** conductor … **les** dieron por muertos…*
   d. Act. Trans.: *La policia cree que el heredero de Harrods y el conductor … fallecieron.*
   e. Others (Act.+Evidential): *Parece ser que el heredero de Harrods y el conductor … fallecieron …*

In the translation equivalents of the *get*-passive sentences into Spanish, we do not find such a varied range of constructions. Apart from the *ser*-passive (34.8%), we find: *se*-passive, *se*-spontaneous, impersonal, active, and others.

(34) 4. That article never **got printed** <GP4> in *The Bugle.*
   a. *Ser*-pass: *Ese artículo nunca **fue** editado…*
   b. *Se*-pass: *Ese artículo nunca **se** publicó…*
   c. Act.: The Bugle *nunca llegó a imprimir el artículo.*
   d. Others (Non-prototypical pass): *Este artículo nunca **apareció** publicado…*

As mentioned above, constructions involving a higher degree of defocusing of agency are the *se*-passive, *se*-impersonal, *se*-spontaneous, and actives coding impersonal and intrinsically spontaneous events. These constructions also allow for variation in the position of the theme element, and tend to highlight the process of change of state rather than terminal stage of an event.

The figures for the translation equivalents of the target sentences in English are shown in Table 3.

6.1.1 *Be-passive*

In general terms, the percentages indicate that the translation equivalents for the *be*-passive are somewhat lower in salience of agency/cause. There are also certain differences in the choice of perspective (trajector and thematic position), and they tend to reflect a relatively higher degree of dynamicity. More specifically, for the translation of the *be*-passive, we find the following features:

6.1.1.1 *Degree of defocusing of agency.* The choice of constructions in Spanish reflects a relatively higher degree of mystification of agency. The equivalent construction, the periphrastic ser-passive, is chosen in only 41.7% of cases. There are also some cases of translation equivalents involving a higher degree of defocusing of agency: se-passive, se-impersonal, se-spontaneous, as well as constructions designating impersonal or intrinsically spontaneous situation types (Total = 24.7%). However, there are also some translation equivalents in the active (16%), and se-middle (2.6%), where the agent is a focal

Table 3. Translation equivalents for English *Get*-passives and *Be*-passives

TRANS-SP	Total GET-P		Total BE-P	
	N	%	N	%
SER-PASS	23	34.8	98	41.7
SE-PASS	9	13.6	25	10.9
SE-IMP	0	0	15	6.5
SE-SP	10	15.1	7	3.0
IMP	17	25.7	9	3.9
SPONT.	0	0	1	0.4
SE-MID	0	0	6	2.6
ACT	4	6.0	37	16.1
OTHER	3	4.5	10	4.3
Ø	0	0	22	9.5
TOTAL	66	100	230	100
PERSPECTIVE				
S-Tr	42	63.6	164	71.3
D-Tr	24	36.3	66	28.6
S-Th	54	81.8	174	75.6
D-Th	12	18.1	56	24.3

participant. For target sentence 5, for example, we find the following translation equivalent: *Sir Michael Jay se acercó al hospital* (Sir Michael Jay went along to the hospital).

6.1.1.2 *Perspective.* Students chose to maintain the same perspective on the event in the majority of cases. Regarding the choice of trajector, in cases where the periphrastic passive was avoided, the choice of a particular construction often involved keeping the same trajector as in the English text (71.3%). In the case of choices in word order, the theme element was placed in thematic position in 75.6% of cases. In general terms, there is little variation in word order in the case of periphrastic passives, and active clause equivalents. In some cases with the *se*-impersonal the endpoint participant is marked as an accusative, but placed in thematic position, so that the same perspective is maintained.

6.1.1.3 *Aspect and dynamicity.* The choice of constructions in Spanish reflects a higher degree of dynamicity. There is considerable use of active constructions, and constructions which focus on the core process rather than on the terminal stage of the event.

6.1.2 *Get-passive*

The figures for the *get*-passive equivalents show a lower degree of salience of agency/cause, fewer differences in choice of perspective and a high degree of dynamicity. The translation equivalents for the *get*-passive are characterized by:

6.1.2.1 *Degree of defocusing of agency.* The constructions chosen as translation equivalents reflect a higher degree of defocusing of agency. The ser-passive is found in only 34.8% of cases. The total percentage of constructions designating impersonal or spontaneous situation types and se-passives is higher for the get-passives (54.4%). The active

construction, where the agent is a focal participant, is found in only 6% of cases. In general terms, the get-passive appears to be construed as involving a higher degree of defocusing of agency.

6.1.2.2 *Perspective.* A distinction is found here between trajector choice and thematic position. In cases where the periphrastic passive was avoided, the choice of a particular construction often involved choosing a different trajector from that of the English text (36.3%). As regards word order, there is a relatively higher tendency to maintain the same thematic position in the translation equivalents (81.8%).

6.1.2.3 *Aspect and dynamicity.* The choice of constructions reflects a higher degree of dynamicity. The percentage of translation equivalents using constructions which focus on the core process (64.9%) is much higher than the choice of the *ser*-passive, in which the emphasis is on the terminal stage of an event.

6.2 Translation into English

In the translation into English, the target sentences were four *se*-passives and four periphrastic *ser*-passives. The translation equivalents are crucially different in both subtypes.

In the case of the *ser*-passive, the use of the passive as an equivalent in the English translation is almost absolute. There is only one case, where the student chose to designate the event in the active.

(35) 3. (el coste)…, *gran parte del cual* ***será sufragado*** *<P3>* ***por*** *los contribuyentes.*
a. … a great part of which will be paid by taxpayers.
b. … and taxpayers will defray most of it.

The translation equivalents of the *se*-passives in English included the following constructions: passive, impersonal, intransitive (spontaneous events), active transitive, and other expressions.

(36) 1. *todavía* ***se escucha*** *<SEP1> la música del concierto…*
a. … the music …can still be heard …
b. … still one can hear the music …
c. … the music … still echoes in our ears …

(37) 4. *La selección del jurado de la investigación judicial sobre la muerte de la princesa y su entonces novio Dodi Al Fayed* ***se inició*** *<SEP4> el jueves pasado…*
a. The selection … was started …
b. The selection … started …

The translation equivalents for the two sets of target sentences in Spanish are shown in Table 4.

6.2.1 *Ser-passive*

The translation equivalents for the *ser*-passive involve a similar degree of defocusing of agency, the same choice of perspective, and an equivalent degree of dynamicity.

Table 4. Translation equivalents for Spanish *Se*-passives and *Ser*-passives

TRANS-ENG	Total SE-P		Total SER-P	
	N	%	N	%
BE-PASS	26	43.3	59	98.3
IMP	8	13.3	0	0
SPONT	14	23.3	0	0
ACT	0	0	1	1.7
OTHER	10	16.7	0	0
Ø	2	3.3	0	0
TOTAL	60	100	60	100
PERSPECTIVE				
S-Tr	45	75	59	98.3
D-Tr	15	25	1	1.7
S-Th	38	63.3	59	98.3
D-Th	22	36.6	1	1.7

6.2.1.1 *Degree of defocusing of agency.* There is equivalence in the degree of mystification of agency in the English translations.

6.2.1.2 *Perspective.* The only difference in perspective is due to the choice of an active construction by one student, thus involving a different trajector, which also affected word order.

6.2.1.3 *Aspect and dynamicity.* The degree of dynamicity is equivalent.

6.2.2 *Se-passive*

For the *se*-passive, however, we find translation equivalents involving higher defocusing of agency, considerable differences in perspective, and a higher degree of dynamicity, all of which point to a clearly different construal of the events designated by the *ser*-passive and the *se*-passive in Spanish.

6.2.2.1 *Degree of defocusing of agency.* There is a general tendency to avoid the choice of the English passive as the preferred translation equivalent; it is found in only 43.3% of cases. The choice of constructions in English involves a higher degree of defocusing of agency; there are quite a number of cases of translation equivalents involving constructions designating impersonal or spontaneous situation types (Total = 36.6%).

6.2.2.2 *Perspective.* The translation equivalents for the *se*-passive in a certain number of cases involve differences in choice of trajector (25%). In contrast with the *get*-passive, translation equivalents of the *se*-passive involve considerable variation in word order (36.6%). It must be pointed out that the trajector was postverbal in three out of the four target sentences in Spanish. Translation into English thus required changes in theme position or in trajector status.

6.2.2.3 *Aspect and dynamicity*. The choice of constructions in English for the *se*-passive involve a higher degree of dynamicity, involving the use of constructions designating impersonal or spontaneous situation types, and others.

The significantly different choices in translation equivalents in English for the *se*-passive and the *ser*-passive may be explained in aspectual terms, which also affect the degree to which the agent's activity is highlighted or defocused. In the original texts, all the examples in the Spanish *ser*-passive and the English *be*-passive are perfective; the emphasis is on the terminal stage of the event, on the construal of the event as completed. In the examples in the *se*-passive, however, we find both perfective and imperfective aspect; in general terms, the *se*-passive is perceived as akin to the active situation type, which prototypically highlights the agent's activity. The Spanish *se*-passive, like the English passive, highlights "the core process which the theme thereby undergoes" (Langacker, 2006: 127), but, unlike the English passive, the emphasis is not on the terminal stage of the event. This would explain the frequent choice of spontaneous and active situations as translation equivalents in English of the *se*-passive. A similar phenomenon is found in the case of the translation of the *get*-passive into Spanish. Though the original examples are perfective in aspect, the general tendency is to choose translation equivalents in Spanish depicting a higher degree of dynamicity, which seems to indicate that the event designated by a *get*-passive is perceived as highlighting the process of change of state.

7. Conclusion

Languages differ in the range of constructions available for the designation of events involving mystification of agency. Variation within and across languages is often due to differences in the way events are construed as involving two distinguishable participants, or whether the event is construed as either internally or externally caused, or the degree to which the theme is construed as an effecting or affected participant. Aspectual features relating to whether the emphasis is on the core process or on the terminal stage of the event are also crucial in the choice of construction.

The differences in the English and Spanish networks seem to account for the results in the study on the choice of translation equivalents for the English *be*-passive and *get*-passive, and for the *se*-passive and the periphrastic *ser*-passive in Spanish. In the translations into Spanish, students made use of a variety of constructions which in some cases reflect differences in how they perceived the construal of the event in the foreign language, and in other cases derive from strategies of avoidance of the periphrastic *ser*-passive in Spanish. The translations into English indicate a clear difference in how events designated by the *se*-passive and those designated by the *ser*-passive are construed by the students.

The *be*-passive is used in English prototypically for events which are perceived as equivalent to the uses of the *ser*-passive. But, in contrast, there is a whole range of events designated by the *be*-passive which are construed as more dynamic, and were thus coded in the active and the impersonal, as well as the *se*-passive.

The use of *get*-passives is very similar to that of *se*-passive, if we observe the distribution of the translations equivalents. *Get*-passives tend to be translated by constructions

other than the *ser*-passive, typically constructions designating a more dynamic construal, such as: impersonals, *se*-spontaneous, and *se*-passives. Similarly, events designated by the *se*-passive are also perceived as more dynamic, and are thus translated by using constructions coding spontaneous and impersonal situation types and other strategies.

The case study indicates that construal of the event in terms of the features of defocusing of agency and dynamicity appears to be crucial in the choice of translation equivalents. Choice of perspective is mostly motivated by discourse-pragmatic features and by the requirements of the codification system in each language.

Notes

1. Kemmer (1993: 269–270) identifies a series of situation types associated with the semantic range of the middle domain (all Spanish examples and some English examples are mine).

1. Reflexive situations: Direct (see oneself, *verse a sí mismo*) and indirect reflexive (bite one's nails, *morderse las uñas*), indirect middle (buy a house for oneself, *comprarse una casa*), logophoric reflexive (believe oneself to be, *creerse*) and logophoric middle (decide to, *decidirse a*).
2. Reciprocal situations: Reciprocal proper (look at each other, *mirarse el uno al otro*), indirect reciprocal (hit/beat one another, *pegarse una paliza*), naturally reciprocal events (embrace, *abrazarse*), collective actions (gather, *juntarse*).
3. Body action middles: Grooming or body care (wash (oneself), *lavarse*; get dressed, *vestirse*); change in body posture (sit down, *sentarse*); nontranslational motion (stretch, *estirarse*); translational motion (go away, *irse*).
4. Cognition middle: Mental events of emotion, cognition and perception (get angry, *enfadarse*; remember, *acordarse*), and emotive speech actions (complain, *quejarse*).

2. Coded examples used throughout the paper are from my own corpora and use the following coding systems, respectively:

1. Marín Arrese (1993):
   A: General Fiction (Texts A) discourse
   B: Academic (Texts B) discourse
   R: Resultative construction
   G: Get passive
   Numbers refer to text numbers and example number
   E.g. A18.R2 = Text A nº 18, Resultative example nº 2
2. Marín Arrese (1995) – The coding system used in the 1995 corpus is similar to the system for the 1993 corpus, with General Fiction (Texts A and B), comparable, English and Spanish texts.
3. Marín Arrese (2002):
   A, B, and C: Comparable journalistic discourse (Texts types A, B and C), English and Spanish texts. The coding system used is the following:
   E: English
   S: Spanish
   g: *The Guardian* (source)
   t: *The Times* (source)
   p: *El País* (source)
   a: ABC (source)
   v: *La Vanguardia* (source)
   Numbers refer to example number
   E.g. AEg04 = Texts A, English, *The Guardian*, example 4.

4. Marín Arrese (in press): Political discourse
   ps: Political speeches
   (st) parliamentary statements
   AB: Anthony Blair
   GB: Gordon Brown
   E.g. GBps: 296 = Gordon Brown, political speech, example nº 296

3. Langacker (1991a: 129ff) observes that three semantic variants of the perfect participial morpheme may be identified in be+PP constructions. The first type [PERF1] is formed on intransitives, typically of spontaneous events and some translational motion events, and derives stative participles. The resulting *be*+adjectival participle construction corresponds to so-called subjective resultative constructions and resultatives of translational motion events (Nedjalkov and Jaxontov, 1988).

The second type [PERF2] involves transitive verbs and additionally effects a figure/ground reversal like [PERF3], in that the landmark of the verbal stem becomes the trajector of the adjectival participle. This type is found in so-called objective resultative constructions (Nedjalkov and Jaxontov, 1988), or stative passives. The effect is to derive a stative relation; the profiled relationship is limited to the final, resultant state, that is, to the "resulting condition of the entity undergoing the change of state" (Langacker, 1991b: 203).

4. Givon (1990: 621–622) and Collins (1996: 50–51) note that the initiator role of the subject-referent is reflected in a series of features: the possibility of embedding of the *get*-passive under verbs of manipulation; the attribution of volitionality to the subject in co-occurrence with adverbs indicating the presence of a volitional entity; and the fact that *get* is more likely to occur than *be* in imperatives.

(i) a. They told him to get/*to be fired.
    b. John got treated by a bad doctor deliberately. (> John acted deliberately)
    c. "Get/*Be stuffed", answered Witcharde.

5. As regards the direction of the alternation, Haspelmath (1993) distinguishes three main types in cross-linguistic terms: causative derivation, where the inchoative pair is basic and the causative is derived (*rodar, hacer rodar*); anticausative derivation, in which the causative verb is basic and the inchoative is derived (*abrir, abrirse*); and non-directed alternation, one of which variants is the so-called labile alternation, with the same verb form in both the causative and the inchoative (*open, open; hervir, hervir*). In terms of the preferences for the direction of derivation of the different verbs, Haspelmath (1993: 105) notes that verbs expressing events normally caused externally (*cerrar*) show preference for the anticausative derivation (*cerrarse*), while verbs of spontaneous events (*wither*) will show a preference for the causative derivation (*cause to wither*).

6. According to Perlmutter (1978: 162–63), predicates determining initially unaccusative clauses are the following: (a) Predicates expressed by adjectives in English; (b) Predicates whose initial nuclear term is semantically a Patient: burn, fall, sink, … This includes the class of inchoatives, die, fill, vanish, …; (c) Predicates of existing and happening: exist, happen, and various inchoatives such as arise, result, disappear, …; (d) Non-voluntary emission of stimuli that impinge on the senses (light, noise, smell, etc.): shine, glitter, smell, …; (e) Aspectual predicates: begin, stop, continue, end, …; (f) Duratives: last, remain, survive, … (my suspension points).

7. Initially 'unergative' clauses correspond closely to active intransitive clauses or activity verbs. According to Perlmutter (1978:162), predicates determining initially unergative clauses are the following: (a) Predicates describing willed or volitional acts: work, play, speak, think, swim, walk, come, …; (b) Certain involuntary bodily processes: cough, sleep, weep, … (my suspension points).

8. In modern Spanish, especially in the spoken language, there is a growing tendency to use an expressed agent with the *se*-passive: "Por la Presidencia se han dado instrucciones para …" (attested: Telediario A3, Federico Trillo, 10 Oct 2001). The presence of a *por*-phrase coding an agent, or at least a causal entity, is already attested in the *Poema de Mio Cid* (c. 1140) (Marín Arrese,1989a):

(i) … que por vós **se** ondre hoy la cort
…that the court should be honoured by/because of you today

9. The term 'effector' is originally found in Van Valin and LaPolla (1997: 85), who define the participant role of effector as "the doer of an action, which may or may not be wilful or purposeful". In this paper, I have applied it to the case of the participant in spontaneous events where the theme role is primarily construed as initiator, thought it is also an affected role.

References

Barber, Elizabeth J. W. (1975). Voice - beyond the passive. *Berkeley Linguistics Society, 1*, 16–24.

Cameron, Carrie Anne (1990). The semantics and pragmatics of voice systems: A functional analysis. *Dissertation Abstracts International 51*(11). (UMI No. 9110952)

Collins, Peter (1996). Get-passives in English. *World Englishes, 15*(1), 43–56.

Déchaine, Rose-Marie (1995). One be. In M. den Dikken & K. Hengeveld (Eds.), *Linguistics in The Netherlands* (73–88). Amsterdam: John Benjamins.

Enger, Hans-Olav & Tore Nesset (2001). The Norwegian and Russian reflexive-middle-passive systems and Cognitive Grammar. In H. G. Simonsen & R. T. Endresen (Eds.), *A cognitive approach to the verb* (223–241). Berlin: Mouton de Gruyter.

Givon, Talmy (1990). *Syntax: A functional-typological introduction* (Vol. 2). Amsterdam: John Benjamins.

Givon, Talmy & Lynne Yang (1994). The rise of the English Get-passive. In B. Fox & P. Hopper (Eds.), *Voice: Form and function* (119–149). Amsterdam: John Benjamins.

Gómez Torrego, Leonardo (1992). *Valores gramaticales de 'se'*. Madrid: Arco Libros.

Haspelmath, Martin (1993). More on the typology of inchoative/causative verb alternations. In B. Comrie & M. Polinsky (Eds.), *Causatives and transitivity* (87–120). Amsterdam: John Benjamins.

Hatcher, Anna G. (1949). To get/be invited. *Modern Language Notes, 64*, 433–46.

Kemmer, Suzanne (1993). *The middle voice*. Amsterdam: John Benjamins.

Kemmer, Suzanne (1994). Middle voice, transitivity and the elaboration of events. In B. Fox & P. Hopper (Eds.), *Voice: Form and function* (179–230). Amsterdam: John Benjamins.

Keyser, Samuel J. & Thomas Roeper (1984). On the middle and ergative constructions in English. *Linguistic Inquiry, 15*, 381–416.

Kitawaga, Chisato & Adrienne Lehrer (1990). Impersonal uses of personal pronouns. *Journal of Pragmatics, 14*, 739–759.

Lakoff, Robin (1971). Passive resistance. *Chicago Linguistic Society, 7*, 149–62.

Langacker, Ronald W. (1991a). *Concept, image and symbol: The cognitive basis of grammar*. Berlin: Mouton de Gruyter.

Langacker, Ronald W. (1991b). *Foundations of cognitive grammar*. (Vol. 2). Stanford: Stanford University Press.

Langacker, Ronald W. (2000). *Grammar and conceptualization*. Berlin: Mouton de Gruyter.

Langacker, Ronald W. (2006). Dimensions of defocusing. In T. Tsunoda & T. Kageyama (Eds.), *Voice and grammatical relations. In honor of Masayoshi Shibatani* (113–137). Amsterdam: John Benjamins.

Levin, Beth & M. Rappaport Hovav (1995). *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, MA: MIT Press.

Maldonado, Ricardo (1999). *A media voz. Problemas conceptuales del clítico se*. Mexico: Universidad Nacional de México.

Marín Arrese, Juana I. (1989a). *Functions of Se in the Poema del Mio Cid.* Unpublished manuscript, University of Oregon.

Marín Arrese, Juana I. (1989b). *Functions of Se in El Quijote.* Unpublished manuscript, University of Oregon.

Marín Arrese, Juana I. (1993). *La Pasiva en Inglés: Un Estudio Funcional-Tipológico*. Unpublished doctoral dissertation, Editorial de la Universidad Complutense de Madrid.

Marín Arrese, Juana I. (1995). Degrees of topicality of subjects and agents in the passive in English and Spanish. Paper presented at the *European Society for the Study of English, 3rd Conference*, 8–12 September 1995, University of Glasgow, Glasgow.

Marín Arrese, Juana I. (1997). Cognitive and discourse-pragmatic factors in passivisation. *Atlantis, 19*(1), 203–218.

Marín Arrese, Juana I. (1999). Conceptualization of events, semantic relations between constructions and topology: a catastrophe theoretic study of get and be. *Journal of English Studies, 1*, 97–117. [Special issue: Cognitive Linguistics].

Marín Arrese, Juana I. (2002). Mystification of agency in passive, impersonal and spontaneous situation types. In: J. Marín Arrese (Ed.), *Conceptualization of events in newspaper discourse: Mystification of agency and degree of implication in news reports* (31–54). Madrid: Universidad Complutense.

Marín Arrese, Juana I. (2003). The middle domain in English and Spanish: Middle and related situation types. In C. Molina, M. Blanco, J. Marín Arrese, A. L. Rodríguez & M. Romano (Eds.), *Cognitive linguistics in Spain at the turn of the century: Vol. I. Grammar and Semantics* (229–252). Madrid: AELCO & Universidad Autónoma de Madrid.

Marín Arrese, Juana I. (2009). Effective vs. epistemic stance, and subjectivity/intersubjectivity in political discourse: A case study. In: A. Tsangalidis & R. Facchinetti (Eds.), *Studies on English Modality*. In honour of Frank R. Palmer (23–52). Berlin: Peter Lang.

Mendikoetxea, Amaya (1999). Construcciones inacusativas y pasivas. In: I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española: Vol. 2* (1573–1629). Madrid: Espasa Calpe.

Nedjalkov, Vladimir P. & Sergej Je. Jaxontov (1988). The typology of resultative constructions. In V. Nedjalkov (Ed.), *Typology of resultative constructions* (3–62). Amsterdam: John Benjamins.

Perlmutter, David M. (1978). Impersonal passives and the unaccusative hypothesis. *Proceedings of the Berkeley Linguistics Society, 4*, 157–89.

Shibatani, Masayoshi (1985). Passives and related constructions: A prototype analysis. *Language, 61*, 821–848.

Shibatani, Masayoshi (1988). Introduction. In: M. Shibatani (Ed.), *Passive and voice* (1–8). Amsterdam: John Benjamins.

Smith, Carlota S. (1970). Jespersen's 'Move and Change' class and causative verbs in English. In M. A. Jazayery, E. C. Palome & W. Winter (Eds.), *Linguistic and Literary Studies in Honor of Archibald A. Hill: Vol. 2. Descriptive Linguistics* (101–109). The Hague: Mouton.

Svartvik, Jan (1966). *On voice in the English verb*. The Hague: Mouton.

Van Valin, Robert & Randy LaPolla (1997). *Syntax*. Cambridge: Cambridge University Press.

CHAPTER 12

Inside the attriter's mind

A comparative exploration of the cognitive constraints in Dutch L1 attrition in an L2 English environment and advanced Dutch L1 acquisition

Merel Keijzer
Utrecht University

1. Introduction

To answer the question what language is, researchers have traditionally looked at imperfect language use in a variety of contexts, ranging from language use in aphasia through L1 and L2 acquisition to the birth of Creoles. It is only when things 'are not quite right' that a window on the grammar in the speaker's mind can be provided (see Corder, 1967). Such deviations are best explained from a cognitive perspective, since they are invariably the result of reduced cognitive capacity or stem from competition: in L1 acquisition, for instance, linguistic and cognitive maturation go hand in hand, while in L1 attrition in emigrants changes are set in motion by a second language competing for processing space. While cognitive approaches are commonly found in acquisition, using notions such as frequency, item-based learning, and form and function (see Ellis, this volume), they are not often found to explain forms of language breakdown, such as language attrition found in emigrants.

In recent years, this has started to change, as insights from cognitive psychology are now tentatively applied in explaining language development across the lifespan, starting with language acquisition and ending in language erosion in the elderly (cf. Hansen, 2001; de Bot, 2007). Typical for these approaches is the stance that language users go through stages of growth and decline throughout their lives, caused by changing cognitive capacities or changing environmental factors.

This paper aims to expand on this recent tradition by comparing the language breakdown in the L1 attrition of Dutch emigrants in an Anglophone L2 environment and advanced stages of L1 Dutch-speaking children. It has been suggested that the parallels between language growth and decline constitute mirror symmetries: language attrition takes place in the reverse order of language acquisition; those features that are acquired late in childhood are also vulnerable in attrition. This idea is captured in the regression hypothesis (Jakobson, 1941). According to this hypothesis, the language output of the

émigrés and advanced acquirers used in this study will show parallels. With all this in mind, this study is guided by two questions. The first is whether such mirror symmetries indeed occur. In other words, does regression obtain? The second, even more important, question is – if regression is found – to what extent cognitive explanations can be given for this phenomenon. More specifically, can cognitive notions like item-based learning, analogy and entrenchment (see Tomasello, 2003; Ellis, this volume) explain regression? In order to answer these questions, this study focuses on the attrition and acquisition of morphology and syntax, as these two areas generally show clear and stepwise paths of development in acquisition, which can easily be compared to attrition processes.

2. Cognitive explanations for the parallels between language attrition and acquisition

The few studies that have investigated the mirror symmetries between language acquisition and dissolution (e.g. Berko-Gleason, 1982; Jordens et al., 1986, 1989; Schmid, 2002) have not attempted to find explanations for their observations. Rather, they have treated regression as a theoretical framework in its own right. However, it is not whether regression occurs, but rather "when and under what conditions its predictions hold true, and what the causal mechanisms are" (Hansen, 1999: 150).

Linguistic deviations attested in attrition and acquisition are most plausibly explained as constraints operating inside the language users' minds. As such, they could be approached from a cognitive or usage-based perspective: it has the advantage of language development being seen as intricately linked to language input. Children build their linguistic repertoires in interaction with their environment through intention reading and pattern finding (Tomasello, 2003: 3–4). It can then be argued that, by contrast, a lack of interaction with the L1 can lead to attrition in emigrants. More importantly, usage-based explanations are cognition-based, involving constructs such as item-based learning, analogy and entrenchment/preemption. As these latter three notions can also have explanatory power in a comparison of cognitive constraints in attrition and acquisition, they will be more elaborately discussed in turn below.

2.1 Item-based learning

In item-based learning, children build their grammars on a word-specific basis. The first instances of grammatical marking typically revolve around a single verb, a so-called verb island. Specific verbs may have only one instance, exemplified in (1) or can have as many as four instantiations (as in (2a–d)) (taken from Tomasello, 2003: 117).

(1) cut – X
(2) a. draw-X
    b. draw – X – on – X
    c. draw – X – for –X
    d. X – draw – on – X

Item-based learning can be used to capture the parallels between acquisition and attrition, because neither attrition nor acquisition affects the whole language system. Instead, linguistic resources are more likely affected on an item-specific basis (Jarvis, 2003: 99). For example, single verbs and their corresponding open slots may be lost in emigrants, but this does not lead to an inability to converse in Dutch, as substitutes are likely to be available.

2.2 Analogy

Drawing analogies is essentially the next step in the acquisition process after item-based learning. It is at this stage that children start transferring item-specific knowledge to novel contexts to create abstract linguistic structures, such as transitives, intransitives, causatives and resultatives. All generalizations at this stage are likely based on a critical number of item-based structures, although it is less clear where the cut-off point exactly is (Marchman & Bates, 1994). Explaining mirror symmetries in child language and language breakdown in emigrants on the basis of analogy is best done in terms of exemplar mass: it is likely that certain acquired features are acquired late because they are infrequent in the input and do not reach a necessary threshold to be used in novel contexts. Similarly, a lack of L1 input in attrition might lead to a drop below the critical threshold, especially in the case of features that are relatively infrequent, leading to loss of a previously acquired feature. An example is the Dutch passive voice, which can be formed both on the basis of transitive verbs (e.g. *hij wordt gewassen* – 'he is washed') and, less commonly, on the basis of intransitives (e.g. *er wordt gedanst,* lit. there is being danced – 'people are dancing'). The latter category is acquired relatively late, quite possibly because intransitive passives are not frequently found in the input and abstractions cannot easily be formed on the basis of exemplars. In the light of the regression hypothesis, the prediction would then be that intransitive passives are lost before transitive passives.

2.3 Entrenchment/preemption

In usage-based approaches to language development, generalizations to novel contexts do not only stem from analogy, but are also likely to arise from the principle that the more a child hears a verb in a particular construction, the less likely he or she is to use that same verb in other contexts, because it has become so ingrained as a function of frequency; a process known as entrenchment (Tomasello, 2003: 180; see also Ellis, this volume). If, however, adults systematically use alternatives for that verb in ways that are communicatively similar, the child may infer that the generalization he or she has made so far is unconventional. It is this process that preempts the generalization and is consequently referred to as preemption (Tomasello, 2003: 178). For example, the child may have – wrongly – assumed that the verb *disappear* is used in simple transitive constructions: *he disappeared the rabbit.* However, if the child becomes aware of the fact that adults structurally use constructions as *He made the rabbit disappear,* he or she may infer that the generalization was inappropriate and subsequently preempts it (Keijzer, 2007: 27). At first glance, it would appear

that the notion of entrenchment cannot be united with the regression hypothesis, because entrenchment implies that it is not what is learned first that is retained in attrition, but what is learned best[1]. In many cases, however, first and best tend to coincide, as features that are acquired first are often entrenched due to their frequency.

Preemption can also have explanatory power in relation to regression. It plays a role in more advanced stages of acquisition, because older children are generally more sensitive to those parts of the input that contradict established entrenchments (Tomasello, 2003: 180). In attrition, on the other hand, L2-based alternatives may be introduced in the L1 that can subsequently preempt existing generalizations. For example, Dutch émigrés in Anglophone L2 environments may experience difficulties with respect to word order in subordinate clauses. While Dutch places the verb at the end of subclauses, English verbs follow the subject in such cases (see Keijzer, 2007: 128–134). Due to influence from English, Dutch émigrés might analyze a sentence as (3a) as correct, while (3b) is in fact the standard Dutch structure.

(3) a. **omdat hij is wakker*
'because he is awake'
b. *omdat hij wakker is*
because he awake is
'because he is awake'

This new generalization might itself become entrenched and cannot be preempted on the basis of input from Dutch, because it is generally not available.

The explanatory value of these cognitive constructs has been questioned in the past (Tomasello, 2003: 181). In this study, they were therefore tested in an experimental research design.

3. Methodology

3.1 Subjects

Three groups of subjects were included in the present study: 45 first-generation Dutch emigrants in Anglophone Canada, 45 matched control subjects in the Netherlands and, finally, a group of 35 secondary school children in the Netherlands (between 13 and 14 years old).

45 Dutch emigrants (mean age: 66.4) were contacted in the greater London area in Ontario, Canada, through various organizations, such as Dutch churches, Dutch-Canadian societies and the Dutch vice-consulate. A number of selection criteria applied, the most important of which were that subjects were at least 15 years old upon emigration (to rule out incomplete L1 acquisition) and that they should not be language professionals (e.g. language teachers or translators). This resulted in a group of 21 male and 24 female participants, who had all stayed in Canada between 21 and 57 years (mean length of stay: 43.5 years). They all indicated (on a sociolinguistic questionnaire) that they used English as their main language of everyday communication, although 44.4% of all subjects

indicated that, in addition, they still used Dutch regularly – defined as a few times a week to daily. The other 55.6% of subjects said that their use of Dutch was confined to a few times a month or year.

These emigrants were compared to a control group of 45 Dutch adults in the Netherlands (mean age 66.2). They were matched to the subjects in Canada on a one-to-one basis regarding the extralinguistic variables of gender, age, educational level, and region of birth and upbringing. In some cases, the control subjects were siblings of the Dutch-Canadians. Fewer selection criteria applied to this group, but they too could not be language professionals. Furthermore, all subjects invariably used Dutch in their everyday lives, but it could not be guaranteed that all subjects were monolingual. As English-medium television programs are subtitled rather than dubbed in the Netherlands and as English is also a mandatory subject in the Dutch secondary school system, all subjects had a basic to intermediate command of English.

Finally, a group of 35 second graders of a Dutch secondary school were included in the research design.[2] They were either 13 or 14 years old (mean age: 13.9). Since attrition has been reported to be a subtle phenomenon, especially in those subjects who migrated at a post-puberty stage (Hansen, 2001), parallels are only likely to be revealed in advanced stages of language development. The expectation concerning the adolescents included in the design was that they had virtually completed their language development, but that they might nevertheless show variation where mature native grammars do not. The children (20 male and 15 female) attended three different, intermediate-level schools spread across the Netherlands. The most important selection criterion that applied to this group was that all participants were monolingual Dutch speakers, defined as their home language being the same as the language spoken at school. All children received foreign language education at school (English, French and German) and all had access to English-medium television and computer games.

3.2 Materials and procedure

The test battery that was used in this study presented a combination of controlled language tasks (none of which were timed), self-assessment measures and a narrative to elicit free spoken data (Keijzer, 2004). Only the language tasks that were most essential to the research set-up and the narrative are discussed here.

As the main focus of this study was on morphology and syntax, two language tasks were included that tapped into these two domains. To assess the subjects' morphological proficiency, subjects were presented with a modified wug test. First created as a measure of internal morphological rules in young Anglophone children (Berko, 1958), the task has been used in many different settings and for many different languages since. The wug test employs the constructs of nonsense words and sentence completion to elicit evidence about morphological rule productivity. In its original set-up, inflection was elicited through visual clues: children were presented with, for instance, a picture of a fantasy creature and were told the name of the creature (*this is a wug*). The subsequent picture would show two such creatures and the child was encouraged to, for instance, pluralize the name he or she had just learned (*now there is another one. There are two of them.*

There are two…). In the present study, a modified version was used: no pictorial support was provided, as no young children were included in the design. Instead, sentences to be completed were presented both orally and in their written form. Subjects were then asked to complete the sentence by inflecting the nonce word, exemplified in (4) below.

(4) You can have one *trag* but if there are two of them, you have two…[3]

What counted as correct (a score of 1) was the inflected form of the word from which the nonsense item was derived. For example, the plural of *trag* (see (4) above) should be *trag-g-en* by analogy of the existing *vlag* ('flag'), which pluralizes in *vlag-g-en* (for an overview of the principles that determine Dutch plural formation and possible analogies that could be drawn for the pseudowords, see Keijzer, 2007: 47–49). In this way, the following morphological features were investigated: plural inflection, agentive formation, article selection, adjectival inflection and diminutive formation for noun phrase morphology and, for verb phrase morphology, simple present tense, simple past tense, past participles, auxiliary selection and future tense (see Keijzer, 2007, chapters 3 and 4).

Subjects' syntactic proficiency, on the other hand, was assessed by means of a grammaticality judgment task, where constructions are presented whose grammaticality has to be judged. Grammaticality judgments can provide information about which structures language users find acceptable and which ones they do not. Such information cannot always be easily obtained in less explicit tasks and certain structures may also be avoided in free spoken data (Altenberg & Vago, 2004: 105). The most common format of the grammaticality judgment task is an absolute one where subjects are presented with a sentence and have to indicate whether they think it is correct or not, occasionally supplemented by a third option of 'don't know'. (Sorace, 1996: 396). The format of the grammaticality judgment task has come under close scrutiny in recent years. Much of the controversy stems from the argument that what subjects are asked to do is very unnatural. Language users do not normally assess the grammaticality of sentences, raising the issue of task complexity (see Robinson, this volume). Still, the task was employed in the present study, mainly for its ability to explicitly test constructions that would not necessarily present themselves in the free spoken data. The relatively easy administration and scoring method of the test were seen as additional advantages. All grammaticality judgments used in this study were absolute, and subjects were presented with a single sentence at a time. These were presented in written form, but were also read aloud. Subjects were then asked to rely on their intuitions and to indicate as quickly as possible whether the construction was correct, incorrect or whether they did not know. Whenever subjects judged a construction to be incorrect, they were asked to provide an alternative formulation. An example of a sentence that formed part of the task is provided in (5).

(5) *Vervolgens hoorde ik auto's toeteren*
Then hear-1SG.PST I car-PL honk-INF
'Then I heard cars honk'
0 *Incorrect, het moet zijn…*('incorrect, it should be…')
0 *Ik weet het niet* ('I don't know')
0 *Correct* ('correct')

Where subjects had judged grammatical structures as correct or had provided a valid alternative for ungrammatical structures, they were awarded a score of 1. In all other cases (i.e. where subjects had not been able to respond to an item at all; where they had judged an ungrammatical sentence as correct or where they had provided a correction that did not result in a grammatical sentence), a score of 0 was assigned. The features examined in this way were negation, passive constructions, the V2 rule, subordination and discontinuous word order (see Keijzer, 2007, chapter 5 for an overview).

In addition to these specific tasks, subjects were also asked to watch a clip from the silent Charlie Chaplin movie *Modern Times* and were asked to retell what they had just seen. In that way, free spoken data was elicited that could be compared to the controlled language task data. Narratives are popular in linguistic research (see Trebits & Kormos, 2008) and this particular film has been used in previous linguistic investigations. In this study, 10 minutes of the movie were played to each subject. The researcher did not leave the room, but did not actively engage in watching the clip with the subject either. Subjects were free in their method of retelling. Where events were missing in the narratives, subjects were reminded and prompted to relate the omitted parts. The whole narrative was recorded on minidisk, resulting in approximately 5 to 10 minutes of film retelling for each subject. Each speech sample was then transcribed orthographically and converted to the standard CHILDES (Child Language Data Exchange System) format that is compatible with the CLAN (Computerized Language Analysis) program, enabling automated searches in the speech samples (see MacWhinney, 2000, for more information about this program and transcription format). The free speech samples were tagged for non-standard occurrences of all the features that were also formally tested as part of the wug test and grammaticality judgment task.

4. Results

Table 1 presents the total scores of all three groups on the wug test and the grammaticality judgment task. The score on the wug test has been divided into noun phrase and verb phrase morphology.

Table 1. Total scores on the noun phrase and verb phrase parts of the wug test and the grammaticality judgment Task (N = 125)

	NP wug (max = 70)	**SD**	**VP wug (max = 50)**	**SD**	**GJT (max = 26)**	**SD**
1: attriters (n = 45)	50.18	8.07	40.47	6.12	17.98	3.44
2: controls (n = 45)	62.86	2.87	47.76	1.91	20.42	2.60
3: acquirers (n = 35)	53.23	4.34	43.63	2.67	17.14	2.91
Mean	55.42	5.09	43.95	3.57	18.51	2.98

All differences in scores were subjected to an ANOVA test of variance. The difference in scores on the noun phrase part of the wug test, first of all, were found to be significant: $F_{(2,119)} = 57.733$, $p < .001$, $\eta^2 = .49$. Not only did the emigrants obtain a significantly lower score than the control subjects, the children also performed on a level significantly below the controls ($p < .001$ in both cases). There was no significant difference between the attriters and the acquirers, however. In other words, the emigrants and learners made as many mistakes, notably more than the controls. For the verb phrase wug test, too, there was a significant effect of condition ($F_{(2,122)} = 35.750$, $p < .001$, $\eta^2 = .37$), but a different tendency was revealed: the emigrants still scored significantly lower than the control subjects and the controls were also significantly better than the children ($p < .001$ in both instances), but the learners were now also found to outperform the attriters ($p < .01$). Finally, the score discrepancies on the grammaticality judgment task were significant: $F_{(2,118)} = 12.864$, $p < .001$, $\eta^2 = .18$. More importantly, the same pattern as for the noun phrase wug test was attested here, with the attriters and acquirers both scoring significantly lower than the controls ($p < .001$ in both cases), but not diverging significantly from each other.

It would thus appear that mirror symmetries were found in the case of the noun phrase wug test and the grammaticality judgment task, but not for the verb phrase part of the wug procedure. Closer inspection of the individual linguistic features that were included in each of the tasks, through a MANOVA test of variance, revealed a slightly different picture, however. For the noun phrase part of the wug test, it was plural inflection, agentive formation, article selection and diminutive formation that elicited mirror symmetries between the attriters and acquirers. In other words, the attriters and acquirers produced significantly lower scores on all these features than the control subjects in the Netherlands, but at the same time did not diverge from each other. The same was true for simple past tense, past participles and future tense on the verb phrase wug test. In the case of the grammaticality judgment task, finally, significant parallels between the emigrants and the adolescents were found for the features of negation and passive constructions. Thus, the overall verb phrase wug test may have revealed no parallels between the attriters and children, but was apparently distorted by two of its features that did not elicit any mirror symmetries, while the rest of the phenomena tested as part of the wug test (adjectival inflection, simple present tense and auxiliary selection) did reveal such convergences. Similarly, the effect found for the grammaticality judgment task as a whole was apparently caused by strong effects found for only two of the features examined by means of this testing device. The other features of V2, subordination and discontinuous word order, which can collectively be grouped under the header of word order in main and subordinate clauses, did not result in mirror symmetries between the attriters and acquirers.

Also important to know is that the predictor variable of education was found to play a role. For both the noun phrase and verb phrase part of the wug test, subjects with a higher educational background tended to produce more standard inflections than those who were less highly educated ($F_{(4,117)} = 4.581$, $p < .005$, $\eta^2 = .14$ for the noun phrase part of the wug test and $F_{(4,120)} = 5.663$, $p < .001$, $\eta^2 = .16$ for the features grouped under verb phrase morphology). Similarly, subjects who had a higher educational background typically produced more accurate judgments on the grammaticality judgment task than subjects with a lower educational background: $F_{(4,116)} = 4.969$, $p < .005$, $\eta^2 = .15$.

The attriters and acquirers did not only produce similar scores on the morphological and syntactic aspects discussed above; the deviations they produced were also highly similar. For example, in agentive formation, the more frequent way of formulating the agentive in Dutch (by means of the suffix *-er*) was overgeneralized in the case of the emigrants and learners, as opposed to the controls, resulting in nonsense items like **frindel-er* and **menker-er.* On the basis of the Dutch phonotactic rules of agentive formation (see Keijzer, 2007: 54–55), *frindelaar* and *menkeraar* were expected, which were in fact most common in the data of the control group. Similarly, the emigrants and adolescents, more so than the control subjects, frequently inflected nonce past tense items according to the strong conjugation, resulting in forms like **vrak* or **vrok* as the past tense of the stem *vruk*. Following the phonotactic cues of this stem (see Keijzer, 2007: 84–87), the expected past tense was *vruk-te,* a weak form. As a final example, significantly more attriters and acquires than controls accepted passive structures as the one provided in (6) below, presented as part of the grammaticality judgment task.

(6) **Critici hadden al gezegd dat er veel gegroeid was door het Turkse elftal.*
Critics have-PL.PST already say-PTCP that there much grow-PTCP be-SG.PST
by the Turkish team
'critics had already said that the Turkish team had improved greatly'

The Dutch *groeien* ('to grow') is an intransitive verb. Only intransitives that are telic (have a natural endpoint) and are clearly controlled by an agent can be passivized in Dutch (see Keijzer, 2007, 116–117). Especially the agentivity factor does not apply in the case of 'to grow', rendering the passive in (6) ungrammatical. It needs to be stressed that the morphological and syntactic mistakes described here were also found in the data of the control condition, although markedly less.

In addition, it is important to indicate that there was not always a one-to-one relationship between the outcomes of the controlled language tasks and the free spoken data. To give an example, while relatively many deviations were attested in the controlled language task data of the attriters and acquirers in the realm of agentives, simple past tense and passives, virtually no deviant forms were found in their spontaneous speech samples collected as part of the film retelling task. This could have been caused, at least partly, by avoidance strategies on the part of the emigrants and children. In fact, there were significantly less agentives in the speech of these two groups than in the free data of the controls. Moreover, although producing hardly any word order-related mistakes in the grammaticality judgment task, the attriters did show many deviant word orders in their film retellings that were not found in the speech of the children or the controls, exemplified in (7) to (10).

(7) **en dan Charlie Chaplin staat op*
'and then Charlie Chaplin stands up'
(8) **dan zij kookt een maaltijd*
'then she cooks a meal'
(9) **toen zei die man dat hij had geen geld*
'then said that man that he had no money'
(10) **omdat de kioskeigenaar ziet niet de politie*
'because the kiosk owner does not see the police'

The first two sentences are incorrect, because the verb in Dutch main clauses always comes in second position, obeying the so-called V2 rule, and no more than one constituent can precede it. That rule is violated here (see Keijzer, 2007: 121–123). In the case of (9) and (10), both subordinate clauses, the verb continues to be placed in second position, while verbs in Dutch subordinates typically appear in final position (Keijzer, 2007: 128–130). As can be seen from the English glosses, such word order is perfectly acceptable in English. On the whole, the results provide evidence for regression, as quite a number of parallels were revealed between the attriters on the one hand and the acquirers on the other; not only did the attriters and acquirers produce more deviant forms than the controls, but the deviations were also of the same type.

5. Discussion

Now that the occurrence of regression has been established, the question is whether these parallels can be explained by means of the cognitive notions of item-based learning, analogy and entrenchment/preemption.

To start with the construct of item-based learning, it can be said that the acquirers and attriters both experienced similar linguistic difficulties, but their overall communicative competence in Dutch was not impaired. They were all perfectly capable of relating the plot of a story as part of the film retelling task. In other words, advanced acquisition and L1 attrition are item-specific and never involve the whole grammatical system (cf. Jarvis, 2003: 99). In addition, the obvious problems of the acquirers and attriters with nonsense items itself presents evidence for item-based learning. Until the age of 3;0, children typically cannot transfer knowledge they have acquired on an item-specific basis to novel contexts (see 2). The available evidence from studies that have looked at children's treatment of nonsense words supports this (Tomasello, 2003: 119). The present study suggests that older children and émigrés also have difficulty transferring morphological distinctions to novel words.

This also relates to the concept of analogy. As the wug test showed, advanced learners and attriters are able to draw analogies, reflected in their ability to inflect nonsense words at all. However, the control subjects were better at drawing analogies than the adolescents and émigrés. They produced an overall higher score on the wug test. Furthermore, some analogies were easier to make than others. For example, article selection in the nonsense *degin,* based on *begin* ('start') was perceived as easier than in the case of *leinde,* formed by analogy of *einde* ('finish'), presumably because the modification in the latter is more dramatic than in the former (only changing the initial letter versus adding an initial consonant). In the first case, subjects supplied the article *het* ('the') relatively easily, but their choice of determiner for *leinde* was more variable, with the controls showing the least variation. Evidence that language users work on the basis of analogy was also found in the syntactic domain. Intransitive passives (as in *er werd gedanst* – there become-SG.PST dance-PTCP – 'people are dancing'), for example, were found to be more problematic than passive forms of transitive verbs (exemplified in *hij werd gewassen* – he become-SG.PST wash-PTCP – 'he is washed'). While this was true for all subjects, the acquirers and attriters

produced significantly lower scores than the control subjects. Perhaps the input of the children and emigrants did not meet the criterion of a critical mass of exemplars of intransitive passives, resulting in a reduced ability to form abstractions in relation to this feature. This does not mean that the learners and attriters were not able to form intransitive passives at all. More frequent forms, for example *er werd gedanst* – 'people are dancing', were more often judged correctly than the more sporadic **er was gegroeid* – 'people are growing'. More frequent items are apparently more salient, and are acquired earlier and lost later than less salient items.

Saliency leads to entrenchment. This study has found that purely syntactic phenomena, like V2 and discontinuous word order, were easier than features conditioned on more than one level, such as plural inflection or simple past formation, which are governed by both morphological and semantic properties: forming the correct plural begins with the realization that one entity may have several instantiations (see Keijzer, 2007: 265–266 for a more detailed discussion). Perhaps this causes syntax to be more ingrained, or entrenched, than morphological features. More importantly, those features that revealed significant parallels between the adolescents and emigrants were invariably less frequent phenomena, such as irregular plural inflection or strong past tense forms rather than regular plurals or the past tense of weak verbs. This indicates that more frequent and entrenched features are less problematic than more sporadic phenomena, and this effect is, apparently, more clearly felt for subjects who are in the final stages of their language construction or are experiencing language dissolution than for mature, non-attrited speakers.

The narratives of the subjects in Canada furthermore suggest that English L2 may function as a source of preemption in relation to V2 and subordination, as English word order was found to influence Dutch structures. The same subjects did not reveal any difficulties in relation to these two features on the formal grammaticality judgment task, indicating that the preemption process had not been completed. The linguistic confusion that results from competition between entrenchment and preemption was thus clearly felt in the data of the attriters. In sum, cognitive notions can be used to explain the mirror symmetries between the fluctuating language systems of both L1 attriters and acquirers.

6. Conclusion

Although this study has been one of the most consistent ones to observe regression on a larger scale, it is not without its limitations. In particular, problems arise regarding the research design where the artificialness of some of the tasks may have impacted on the results. Moreover, the exact effect of task complexity was not taken into consideration either (see Robinson, this volume). Still, this study has shown that general cognitive limitations prevent full utilization of the language systems in both populations, although these limitations most likely stem from different sources: a lack of cognitive maturation on the part of the acquirers and competition from a second language in the attriters. In both cases, resources are limited and trade-offs are made. Interestingly, this results in the same deviations. This is not to say that the deviations will always remain the same for both populations; in recent years language is more and more seen as a dynamic system

(see also Spivey, 2007; Lowie et al., 2008) that is constantly changing due to internal restructuring and external influences. It may well be that in a number of years all irregularities will have been smoothed out of the L1 learners' repertoires and that the Dutch proficiency of the émigrés will have gotten worse – or better as a function of older age and a preoccupation with the past, including a return to the language of the past. What all of this does show is that cognitive constraints and outside factors interact to form a language system and that, at certain points in its development, the systems of different language users greatly resemble each other, which in turn grants us a peek inside the language user's mind. To learn more about how language functions 'when things are not quite right' (see (1)), it is desirable for studies to also compare outcomes of other fluctuating language systems, such as L2 acquisition, L2 attrition, bilingualism and multilingualism, language contact, pidgin and Creole varieties and diachronic language change (Hansen, 1999: 4). Although a start has been made in the present research project, it is the study of language "during its unstable or changing phases that is an excellent tool for discovering the essence of language itself" (Slobin, 1977: 185).

Notes

1. The hypothesis that it is not what is learned first, but what is learned best is sometimes seen as an opposing theory of regression (Pitres' rule, 1895), but it is most commonly viewed as another interpretation of regression.

2. A group of 35 second graders was included as opposed to 45 subjects in both the Dutch-Canadian and control group, for practical reasons: due to time constraints it proved impossible to test more than 35 adolescents. However, SPSS, the statistical software program used, can weigh cases to eliminate this difference.

3. For the sake of convenience, this sentence is presented in English. The original text was Dutch.

References

Altenberg, E. P. & Vago, R. M. (2004). The role of grammaticality judgments in investigating first language attrition: A cross-disciplinary perspective. In M. S. Schmid, B. Köpke, M. Keijzer & L. Weilemar (Eds.), *First Language Attrition: Interdisciplinary perspectives on methodological issues* (105–129). Amsterdam & Philadelphia: John Benjamins.

Berko, J. (1958). The child's learning of English vocabulary. *Word, 14* (1), 150–177.

Berko-Gleason, J. (1982). Insights from child language acquisition for second language loss. In R. D. Lambert & B. F. Freed (Eds.), *The Loss of Language Skills* (13–23). Rowley, MA: Newbury House.

Bot, K. de (2007). Dynamic systems theory, life-span development and language attrition. In B. Köpke, M. S. Schmid, M. Keijzer & S. Dostert (Eds.), *Language Attrition: Theoretical perspectives* (53–68). Amsterdam & Philadelphia: John Benjamins.

Corder, S. P. (1967). The significance of learners' errors. *International Journal of Applied Linguistics, V-4,* 162–169.

Ellis, N. (this volume). Constructing a second language: Learning as a function of frequency, frequency distribution, form and function. Talk presented at the 33rd International LAUD Symposium: Cognitive approaches to second/foreign language processing: Theory and pedagogy. 10–13 March 2008, Landau, Germany. This volume.

Hansen, L. (1999). Not a total loss: The attrition of Japanese negation over three decades. In L. Hansen (Ed.), *Second Language Attrition in Japanese Contexts* (142–153). Oxford: Oxford University Press.

Hansen, L. (2001). Language attrition: The fate of the start. *Annual Review of Applied Linguistics, 21,* 60–73.

Jakobson, R. (1941). *Kindersprache, Aphasie and allgemeine Lautgesetze.* Uppsala: Almqvist.

Jarvis, S. H. (2003). Probing the effects of the L2 on the L1: A case study. In V. Cook (Ed.), *The Effects of the Second Language on the First* (81–120). Clevedon: Multilingual Matters.

Jordens, P., Bot, K. de, Os, C. van & Schumans, J. (1986). Regression in German case marking. In B. Weltens, K. de Bot & T. van Els (Eds.), *Language Attrition in Progress* (159–174). Dordrecht: Foris.

Jordens, P., Bot, K. de & Trapman, H. (1989). Linguistic aspects of regression in German case marking. *Studies in Second Language Acquisition, 11* (2), 179–204.

Keijzer, M. C. J. (2004). First language attrition: A cross-linguistic investigation of Jakobson's regression hypothesis. *International Journal of Bilingualism, 8* (3), 398–292.

Keijzer, M. C. J. (2007). *Last in First out? An investigation of the regression hypothesis in Dutch emigrants in Anglophone Canada.* Utrecht: Lot publications.

Lowie, W., Verspoor, M. & Seton, B. (2008). Conceptual representations in the multilingual mind. Talk presented at the 33rd International LAUD Symposium: Cognitive approaches to second/foreign language processing: Theory and pedagogy. 10–13 March 2008, Landau, Germany. This volume.

MacWhinney, B. (2000, 3rd). *The CHILDES Project: Tools for analyzing talk.* Mahwah, NJ: Lawrence Erlbaum Associates.

Marchman, V. A. & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language, 21* (2), 339–366.

Pitres, A. (1895). Etude sur l'aphasie chezs les polychottes. *Revue de Médicine, 15,* 873–899.

Robinson, P. (2008). Task complexity, intentional reasoning demands, L2 speech production, learning and syllabus design. Talk presented at the 33rd International LAUD Symposium: Cognitive approaches to second/foreign language processing: Theory and pedagogy. 10–13 March 2008, Landau, Germany. This volume.

Schmid, M. S. (2002). *First Language Attrition, Use and Maintenance: The case of German Jews in Anglophone countries.* Amsterdam & Philadelphia: John Benjamins.

Slobin, D. I. (1977). Language change in childhood and history. In J. Macnamara (Ed.), *Language Learning and Thought* (185–214). New York, NY: Academic Press.

Sorace, A. (1996). The use of acceptability judgments in second language acquisition research. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of Second Language Acquisition* (375–409). San Diego, CA: Academic Press.

Spivey M. J. (2007). *The Continuity of Mind.* New York, NY: Oxford University Press.

Tomasello, M. (2003). *Constructing a Language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Trebits, A. & Kormos, J. (2008). Working memory capacity and narrative task performance. Talk presented at the 33rd International LAUD Symposium: Cognitive approaches to second/foreign language processing: Theory and pedagogy. 10–13 March 2008, Landau, Germany. This volume.

SECTION III

Cognitive language pedagogy

Classroom studies with applications for teaching

CHAPTER 13

Situating and distributing cognition across task demands

The SSARC model of pedagogic task sequencing

Peter Robinson
Aoyama Gakuin University, Japan

1. Introduction: Cognition, settings, abilities and the SSARC model of pedagogic task sequencing

Complex cognition in any domain of thought and endeavour, such as mathematics or music (as when preparing a year-end financial audit of a large trading company, or performing Chopin's second piano concerto at the Albert Hall), is inevitably grounded and embodied in attempts to perform simple tasks (adding single digit numbers, doing two or five finger exercises) in unschooled or schooled settings where performance on them is scaffolded by support from caregivers, peers, and teachers. Across settings for, and courses of, *instruction*, from these simple beginnings, tasks are staged to increase in complexity for learners in what are judged to be manageable ways, as informed, for example, by mathematics or music learning theory, or theories of situated action through talk in instructed second language learning settings.

Learners in any domain progress differently, some falling behind, some excelling in learning as tasks they are presented with increase in complexity. In part this is because learners differ in the cognitive abilities they have that are drawn on in performing tasks in a domain (e.g., verbal working memory capacity in the case of mathematics, auditory acuity in the case of music). Successful adaptation to the learning and performance demands of tasks, and further progress along the route from novice to expert that *instructional sequences* of tasks prescribe is, to this extent, a result of the *interaction* between the demands educational tasks place on learners and the strengths and weaknesses in abilities that they bring to them. When learners' strengths in abilities are matched to the cognitive processing demands instructional tasks place on them – and when learners are above the necessary threshold of interest in, and motivation to perform instructional tasks – their prospects for success on them become clearer. Consequently, many have argued, educational research and curriculum planning should seek to identify the complementary 'match' between learner abilities, interests, motivations and task demands (see e.g., Ackerman, 2003; Altman, 1972; Cronbach & Snow, 1977; Dornyei, 2002, 2005; Snow, 1987, 1994;

Robinson, 2002a, 2005b, 2007b; Shuell, 1980) in order to facilitate progress for individuals in instructional programs. This complex, adaptive-interactionist research agenda serves as the backdrop to the present chapter which focuses on a key aspect of pedagogic decision-making, for any course of instruction, in any setting, and for any learner – *syllabus design*, and the theoretical rationale for it. Narrower in focus still, this chapter addresses issues in syllabus design for *second language* (L2) learners, though the adaptive-interactionist backdrop sketched above, and the issues of task-person-ability-matching it raises, will be addressed throughout and specifically accommodated in the final section of this chapter.

Syllabus design

Syllabuses serve the external (to the learner) purposes of coordinating practical decisions about articulating instruction and language testing within any one classroom setting, and across classrooms, institutions, local and national sectors and agencies, countries and cultures. More importantly (for the learner, and ultimately too for the various stakeholders in effective language education) syllabuses are also an external representation of two *theoretical* decisions made about how to optimally stimulate the internal cognitive processes that lead to language development – decisions about what *units* should be used to organize L2 instruction, and about the *sequence* in which the units are to be presented to learners. For many years approaches to instruction at private and public levels adopted language-based syllabuses, characterizing 'units' in terms of variously described language structures, patterns, notions, functions, lexical items, propositional relations, etc., and classroom learning events were organized around teaching and practicing these, explicitly, in some form of serial or cyclical order, based on various notions of complexity, frequency, usefulness, coreness, and others (see e.g., Crombie, 1985; Ellis, 1993; Fries, 1952; Trim & North, 2001; Wilkins, 1976; Valdman, 1978, 1980; van Ek, 1976; Willis, 1990; and see Richards, 2001; Robinson, 2009a; and White, 1988, for overviews).

In contrast, beginning with the Procedural Syllabus described by Prabhu (1987) task-based approaches to language instruction have assumed that tasks are more preferable units of syllabus design (for rationales see e.g., Candlin, 1984; Long, 1985, 1989, 2006; Long & Crookes, 1992; Long & Robinson, 1998; Prabhu, 1987; Robinson, 1996, 2001b, 2009a; Skehan, 1998; Skehan & Foster, 2001; Van den Branden, 2006) and that learner language develops predominantly incidentally, while 'doing' tasks in the L2 (see Lesgold, 2001; Schank, 1999). Consequently, there has been a growing amount of empirical research over the last twenty years into these assumptions – in particular, into the effects of design characteristics of tasks on the accuracy, fluency and complexity of L2 speech production; the amount of interaction incidental learning they lead to; and the extent to which Focus-on-Form (Doughty, 2001; Doughty & Williams, 1998; Long, 1991; Long & Robinson, 1998) is needed to enhance and promote the effectiveness of the task-based language learning environment (see R. Ellis, 2003; Mackey, 2007; and Samuda & Bygate, 2008, for recent summaries). As has frequently been pointed out, however, (see e.g., Long, 2006, 2007, 2009; Robinson, 2001a, 2001b, 2005a, 2007a; Robinson & Gilabert, 2007) much of this research remains disparate – lacking an accepted *taxonomy* of task characteristics that can be both cumulatively researched for their individual and combined effects on learner

language and language learning, and also related to pedagogic decision-making about task sequencing in a principled way.

Mapping pedagogic task design to real-world task demands – Constraints on a task taxonomy

In recent years educational research outside the field of L2 instruction has made concerted efforts to apply methods for cognitive task analysis (see e.g., Hoffman & Militello, 2009; Hollnagel, 2003) to principled decision making about task design and the design of instructional sequences, stressing that such decision making must be motivated both by evidence from empirical research into task effects on learning, and by an instructional-design *theory* (see e.g., Reigeluth, 1999) that describes how tasks can be sequenced, and also related to learning objectives and assessment procedures, and profiles of learners' abilities. In the approach to these issues that I will be describing, as in many other instructional-design theories, (see Gagne, Wolar, Golas & Keller, 2005; Ledford & Sleeman, 2002) task-based L2 syllabus design follows from an analysis of the real-world target-tasks learners need to perform in the L2 on exit from instructional programs (Fine, 1974; Long & Crookes, 1992; Long 2005; McCormick, 1979; Williams, 1977). The demands of these tasks are analyzed, and then initially simple, but progressively more complex pedagogic versions are sequenced to gradually approximate the real-world target task demands.

This procedure requires an initial specification, and *taxonomy* of the task characteristics (and categories of these) available to the task and syllabus designer for creating L2 pedagogic tasks. This specification cannot be an exhaustive *simple description* of all the ways in which pedagogic tasks may differ (see Fleishman & Quaintance, 1984; Sokal, 1974). This would result in a potentially endless and so unmanageably extended list of characteristics, offering no guidance for task and syllabus designers aiming to draw on it (why would it matter for language *learning*, for example, that some tasks take on average one minute to perform, others two, or three, or three and a half, or five or thirty minutes, etc.?) Rather, the taxonomy must be selective of those categories and characteristics that are; (a) most effective in *promoting learning and performance* (i.e., they are selected following some theoretical SLA motivation); (b) of most *utility val*ue to task designers attempting to promote the abilities needed for real-world task performance in classrooms (i.e., they are selected because they map pedagogic task demands to needed real-world task performances in a coherent way); (c) and are *operationally feasible* for task and materials designers (i.e., the categories and components can be used in a consistent way by a wide variety of materials designers, thereby ensuring comparability of design decisions in one context, and program, with those in others). A taxonomy of pedagogic task characteristics must therefore satisfy this three-way mapping constraint of task characteristics to *learning processes*, *target task analyses* and *operational consistency* if it is to be of optimal use for task and syllabus design purposes (Robinson, 2007a, 2009a). In the second section of this chapter I propose such a taxonomy but principally address in the remainder of what follows only how it meets the constraint of mapping task characteristics to *learning processes* when used in conjunction with a model for operationalizing them in the design of pedagogic tasks and task sequences.

Situating and distributing task cognition

Language learning is a dynamic process resulting in the emergence of increasingly complex interlanguage systems over time (see De Bot, Lowie & Verspoor, 2007; Eisenstein, 1989; Ellis, 1998, 2008, this volume; Ellis & Larsen-Freeman 2006; Ellis & Robinson, 2008; MacWhinney, 2001, 2008; Munoz, 2006; Ortega & Byrnes, 2008; Robinson & Ellis, 2008a, 2008b) – albeit systems which for individual L2 learners often exhibit non-linear trajectories in (morphological, lexical, syntactic, etc.) domains of accurate and complex L2 use, when sampled across different windows of time (days, weeks, months, etc.) and normatively compared to native-speaker L1 use in these domains (see e.g. Cook, 2003; Huebner, 1983; Kellerman, 1985; Larsen-Freeman, 2006; Long, 1991; Ortega & Byrnes, 2008; Sato, 1990). Slow quantitative increases in accuracy and complexity in a domain of L2 use can also be followed by sudden qualitative changes, which occur in many cases in predictable sequences in such areas as the acquisition of negation, question forms, temporal and spatial reference and morphology (Larsen-Freeman & Long, 1991; Ortega, 2009; Perdue, 1993a, 1993b; Robinson, Cadierno & Shirai, 2009). Alternatively, increases in accuracy and complexity in a domain of L2 use can slow to a temporary, or permanent halt (Han & Odlin, 2006).

A rationale for situating learning opportunities within task demands and distributing these opportunities over *sequences* of tasks therefore addresses specific issues raised more generally by dynamic/complex systems theory – not only with respect to how the development of learner language can be expected to proceed, and so be appropriately measured across different time-scales (Van Geert, 2008; Verspoor, Lowie & Van Dijk, 2008) – but also with respect to the design of the task-based environment intended to promote its emergence. Instructional designers attempt to facilitate learning by scaling down the complexity of real world performative environments, and presenting simple, more manageable versions of it to learners (Ehret, Gray & Kirschenbaum, 2000) and then subsequently descaling, and increasing the real world veridicality of tasks that learners perform. A taxonomy of task characteristics, and categories of these, for such design purposes could therefore be called, using a term from dynamic/complex systems theory, the *'phase-space'* for task design, representing all the possible ways in which pedagogic tasks can differ. Within this phase-space, actual *pedagogic task design* is the result of mapping the coordinates of specific real-world tasks identified by a needs analysis (say, the need to explain to a superior what caused a conflict in the workplace) to all the parameters of tasks specified by the classification system and taxonomy as available for systematic manipulation. So individual pedagogic tasks will have their own *'parameter-space'* which is adjusted, gradually, to increase in real world veridicality and complexity. For example – in ways which I will describe below – the parameter space for one sequence of pedagogic tasks could be +/− planning time (Ellis, 2005; Skehan, 1998) and −/+ intentional reasoning demands (Ishikawa, 2008; Robinson, 2007c), where + planning time and − intentional reasoning demands represents the maximally scaled down and simple version of a pedagogic task.

The Cognition Hypothesis

The fundamental pedagogic claim of the Cognition Hypothesis (Robinson, 2001b, 2003b, 2005a) is that distributing optimal task-based L2 language use and learning opportunities over time, i.e., task *sequencing,* is done by designing and having learners perform tasks

simple on all the relevant parameters of task demands first, and then gradually increasing their cognitive complexity on subsequent versions. Such sequences allow for *cumulative* learning, since each task version differs in only small respects from the previous one, but also introduces an incremental *increase in the conceptual and communicative challenge* of the task which prompts learners to adjust and expand their interlanguage resources to meet it, thereby creating the conditions for L2 development. In the theory I will be describing, increasing task complexity by, for example, requiring reasoning about the intentional states that cause people to perform actions (+ intentional reasoning) versus simply describing their actions (– intentional reasoning) has the potential to *direct* learner attention to, and promote 'noticing' (Schmidt, 2001) and internalization of those aspects of linguistic code which can be used to meet these complex task demands (e.g., cognitive state terms such as 'think' and 'wonder', 'doubt', and the complex syntactic complementation that accompanies their use; 'X wonders if Y', 'X doubts that Y believes Z', etc.) (see Astington & Baird, 2005; Bartsch & Wellman, 1995; Lee & Rescorla, 2002; Lohman & Tomasello, 2003; Nixon, 2005). On the other hand, for example, removing planning time (– planning time) increases cognitive complexity but simply disperses attentional resources over many dimensions of tasks with no particular linguistic correlates. I argue increasing complexity on the former *resource-directing dimensions* of task demands promotes attention to form-function/concept mappings, thereby leading to interlanguage development, and on the latter *resource-dispersing dimensions* it promotes increasing automatic access to current linguistic resources. Both are important (Bialystok, 1994; Robinson, 2003a, 2005a, 2007a). With this distinction in mind, there are only two instructional-design *principles* for task sequencing in the theoretical-model I describe, making this an operationally feasible proposal for task-based syllabus design.

Task Sequencing Principle 1. *Only the cognitive demands of tasks contributing to their intrinsic conceptual and cognitive processing complexity are sequenced*
Following this principle, for example, tasks that do not require intentional reasoning are performed before those that require it. In contrast, the interactive demands of tasks (such as whether they require one-way or two-way information exchange) are replicated each time pedagogic task versions are performed so as to help ensure deep semantic processing (Craik & Lockhart, 1972; Hulstijn, 2001, 2003) rehearsal in memory (Robinson, 2003a) and elaboration and successful transfer of the particular 'schema' for interactive or monologic task performance to real-world contexts of use (Schank, 1999; Schank & Abelson, 1977). The theory is thus parsimonious, placing the sole emphasis for task sequencing on Task Complexity.

Task Sequencing Principle 2. *Increase resource-dispersing dimensions of complexity first (e.g., from + to – planning time), and then increase resource-directing dimensions (e.g., from – to + intentional reasoning)*
The rationale for this can be described in the following way. First (Step 1), tasks simple on all dimensions are performed (e.g. + planning, – intentional reasoning). Task performance thus draws on the simple, stable (SS) '*attractor state*' of current interlanguage (cf. Larsen-Freeman & Cameron, 2007; van Geert, 2008). Next (Step 2), complexity on

resource-dispersing dimensions is increased (e.g., – planning, – intentional reasoning). This promotes speedier access to, and so automatization (A) of, the current interlanguage system (cf. Bialystok,1994; DeKeyser, 2001, 2007; Robinson, 1997; Robinson & Ha, 1993; Segalowitz, 2003). Finally (Step 3), complexity on both resource-dispersing and resource-directing dimensions is increased (e.g., – planning, + intentional reasoning). This promotes restructuring (R) of the current interlanguage system and the development of new form-function/concept mappings along resource-directing dimensions of task demands (cf. Andersen, 1984; Andersen & Shirai, 1996; Doughty, 2001; Karmiloff-Smith, 1992; Li & Shirai, 2000; Meisel, 1987; Sato, 1990; von Stutterheim & Klein, 1987; von Stutterheim & Nuse, 2003) and introduces maximum complexity (C) destabilizing the current interlanguage system (cf. Long, 2003). Increasing task complexity by sequencing shifts in task demands induces (in the theory proposed here) similar shifts in the structure of interlanguage resources used to accomplish them (Robinson, 2005a).

The SSARC model

The steps described briefly above constitute the 'SSARC model' for increasing L2 pedagogic task complexity, and can be represented in the following way, where i = current interlanguage state; e = mental effort; 's' = simple task demands; 'c' = complex task demands; rdisp = resource dispersing dimensions of tasks; rdir = resource directing dimensions of tasks; and ^{n} = potential number of practice opportunities on tasks, which are determined *in situ* by teachers observing pedagogic task performance by individuals, dyads and groups:

Step 1. SS = i × e ('s'rdisp) + ('s'rdir) ^{n}
Step 2. A = i × e ('c'rdisp) + ('s'rdir) ^{n}
Step 3. RC = i × e ('c'rdisp) + ('c'rdir) ^{n}

The remaining sections of this chapter expand on this introduction to, and rationale for, the SSARC model. The following section illustrates a framework describing task characteristics (and their taxonomic structure) and further illustrates how they can be combined in the principled order described by the SSARC model during decision-making about task design and task sequencing. The final section relates this framework for describing task characteristics to individual differences in cognitive abilities and affective factors, and then describes how research into the interaction of learner abilities and task demands during successful task performance and learning may coherently proceed.

2. Scaling the world to classroom contexts: The Triadic Componential Framework for task classification and sequencing

As described above, the fundamental pedagogic claim of the Cognition Hypothesis is that pedagogic tasks should be sequenced for second language learners only on the basis of increases in their cognitive complexity. The pedagogic claim is motivated by a number of ancillary theoretical claims from second language acquisition (SLA) research and is supported by some recent empirical findings into them. These theoretical claims are that

increasing the cognitive complexity of task demands leads to more accurate and complex L2 production, and also to greater amounts of interaction, noticing of forms made salient in the input and uptake of corrective feedback, on complex versus simpler task versions (for various studies theoretically and/or empirically addressing these claims see Gilabert, 2005, 2007; Gilabert, Baron & Llanes, 2009; Ishikawa, 2007, 2008; Kuiken & Vedder, 2007; Michel, Kuiken & Vedder, 2007; Niwa, 2000; Nuevo, 2006; Rahimpour, 1997, 1999; Révész, 2007; Robinson, 1995, 1996, 2001a, 2003b, 2005a, 2007c; Robinson & Gilabert, 2007). The taxonomic Triadic Componential Framework for further researching these claims, and for operationalising task design and task sequencing decisions, distinguishes in its taxonomic inventory three *categories* of task characteristics.

Task Complexity

Task *Complexity* concerns cognitive factors affecting their intrinsic cognitive challenge (e.g., doing simple addition versus calculus) and is a consequence of the relatively lesser or greater demands tasks make on conceptualization, attention, memory and reasoning processes during task performance. Figure 1 lists characteristics contributing to the cognitive complexity of second language learning tasks that I have argued are of two *kinds* (Robinson, 2001a, 2003b, 2005a), resource-directing and resource-dispersing characteristics. *Resource-directing* variables make conceptual/ communicative demands which direct learner attention and effort at conceptualization in ways that the linguistic L2 system can help them meet. Understanding *how* this may be so, and then adopting appropriate *operational measures* of the effects of these resource-directing dimensions of task demands on learner *language* and language *learning* for the purpose of researching whether it *is* so, requires a linguistic theory commensurate with this claim. Cognitive Linguistics is an approach to theorizing the conceptual-linguistic interface during language performance, and so can be applied, I think, to describing the domains in which L1 patterns of conceptualization and thinking-for-speaking (Slobin, 1996, 2003) need to be reconfigured to suit the linguistic constructions and morphological systems the L2 makes available for conceptual expression in language (Achard, 2008; Ellis, N. 2003, 2008, this volume; Odlin, 2008; Pavlenko & Jarvis, 2007; Talmy, 2000, 2008; Tyler, 2008). As such it is an approach that will be important to draw on in researching the effects of task characteristics on L2 production during instructed (and also naturalistic) L2 learning (see Achard & Niemeier, 2004; Berman & Slobin, 1994; Kormos, 2006; Perdue, 1993a, 1993b; Pütz, Niemeier & Dirven, 2001a, 2001b; Robinson & Ellis, 2008a, 2008b; Slobin, 1993, 1996, 2003).

For example, in L2 English, tasks which require complex reasoning about the intentional states that motivate others to perform actions can be expected to draw heavily on the use of cognitive state terms for reference to other minds – *she suspected, realized*, etc. – and in so doing orient learner attention to the complement constructions accompanying them – *suspected that, wonders whether*, etc. – so promoting awareness of, and effort at, complex L2 English syntax (see e.g., Deissel, 2004; Lee & Rescorla, 2002; Godfroid et al., this volume; Lohman & Tomasello, 2003; Robinson, 2007c). Similarly, in L2 English, tasks which require complex spatial reasoning, and articulation of this in describing how to move, and in what manner, from point A to point E, by way of intermediary landmark points B, C and D, etc., can be expected to draw heavily on the use

Task Complexity (Cognitive factors)	*Task Condition* (Interactive factors)	*Task Difficulty* (Learner factors)
(Classification criteria: cognitive demands)	(Classification criteria: interactional demands)	(Classification criteria: ability requirements)
(Classification procedure: information-theoretic analyses)	(Classification procedure: behavior-descriptive analyses)	(Classification procedure: ability assessment analyses)
a. **Resource-directing variables** making cognitive/conceptual demands	a. **Participation variables** making interactional demands	a. **Ability variables** and task-relevant resource differentials
+/– here and now	+/– open solution	h/l working memory
+/– few elements	+/– one-way flow	h/l reasoning
–/+ spatial reasoning	+/– convergent solution	h/l task-switching
–/+ causal reasoning	+/– few participants	h/l aptitude
–/+ intentional reasoning	+/– few contributions needed	h/l field independence
–/+ perspective-taking	+/– negotiation not needed	h/l mind/intention-reading
b. **Resource-dispersing variables** making performative/procedural demands	b. **Participant variables** making interactant demands	b. **Affective variables** and task-relevant state-trait differentials
+/– planning time	+/– same proficiency	h/l openness to experience
+/– single task	+/– same gender	h/l control of emotion
+/– task structure	+/– familiar	h/l task motivation
+/– few steps	+/– shared content knowledge	h/l processing anxiety
+/– independency of steps	+/– equal status and role	h/l willingness to communicate
+/– prior knowledge	+/– shared cultural knowledge	h/l self-efficacy

Figure 1. The Triadic Componential Framework for task classification – categories, criteria, analytic procedures, and design characteristics (from Robinson, 2007a)

of constructions for describing motion events. Such tasks have the potential to promote awareness of *lexicalization patterns* in L2 English for describing these motion events, in which motion and manner are typically conflated on verbs and paths are concatenated in a series of satellites – themselves expressing a variety of locational meanings, and which may differ from the way in which prepositional path satellites semantically divide up paths of motion and spatial location in the learner's L1 (see e.g., Berman & Slobin, 1994; Cadierno & Robinson, 2009; Filipović, 2007; Filipović & Vidaković, this volume; Tyler & Evans, 2003).

In yet a different conceptual domain, tasks requiring reference to events happening now, in a shared context (Here-and-Now) orient learner attention to morphology for conveying tense and aspect in the present, compared to events requiring much more cognitively demanding reference to events happening elsewhere in time and space (There-and-Then), requiring as this does, greater effort at conceptualization (since events are not visually available in a shared context) and greater demands on memory. Cognitively less demanding Here-and-Now tasks orient learners to the prototypical meanings which progressive aspect (prototypically used to refer to activities) and past tense morphology (prototypically used to refer to achievements) accompany (Anderson & Shirai, 1996; Shirai, 1999). In contrast, on cognitively and conceptually more demanding There-and-Then tasks learners are increasingly prompted to use tense and aspect morphology to mark non-prototypical meanings (e.g., states for past tense, and accomplishments for progressive) (see Robinson, Cadierno & Shirai, 2009, for this finding).

The dimensions of resource-directing cognitive complexity illustrated in Figure 1 also describe well-known developmental trajectories in the mastery of concepts in childhood e.g., from concepts of the Here-and-Now to the There-and-Then (Behrens, 2001); from simple belief desire psychology to the ability to attribute intentional mental states to others (Bartsch & Wellman, 1995; Wimmer & Perner, 1983); and from the mastery of first and second person perspectives on events to the ability to take multiple perspectives on an event, or location (Carassa, Aprigliano & Geminiani, 2000; Cornell, Heth & Alberts, 1994). Conceptual development along these trajectories ushers in many changes and additions to the child's linguistic system, such as the ability to use past tense, and deictic expressions for referring to absent objects; the ability to use mental state verbs, and their syntactic complementation when referring to the intentions that cause behavioral actions; the ability to use non topological, axis-based referring expressions when describing locations, etc. Since languages grammaticize, lexicalize and syntacticize these conceptual domains in similar and also differing ways (Bowerman & Levinson, 2001; Ringbom, 2006), then increasing the conceptual demands of tasks from simple to complex along these dimensions constitutes, I have argued, a 'natural order' for remapping conceptualization to linguistic expression in the L2, and for noticing where this mapping differs from that in the L1, so leading to interlanguage development (see Robinson, 2003b, 2005a; Robinson & Gilabert, 2007).

In contrast to resource-directing variables, *resource-dispersing* task characteristics make performative and procedural demands on cognition, but increasing the complexity of the cognitive demands these characteristics of tasks make does not direct learner attention and effort at conceptualization to any particular aspects of language code. For

example, making a task more performatively complex by removing planning time for it simply disperses learner attention over many linguistic and non-linguistic aspects of the task (see Ellis, 2005; Skehan, 1998). Similarly, making a task complex by adding a secondary, or third task demand (answering the phone, while monitoring a TV screen, and also listening for a child crying in a separate room) also disperses attention over many non-linguistic aspects of the task. Performing increasingly complex versions of tasks on these dimensions promotes not *noticing* of language code and interlanguage *development* of new linguistic, conceptual form-function mappings, but rather consolidation and fast real-time access to existing interlanguage resources. In Bialystok's terms (1994), increasing complexity along resource-directing dimensions promotes greater *analysis*, and representational redescription of L2 conceptual-linguistic knowledge and form-function mappings, while increasing complexity along resource-dispersing dimensions promotes greater *control* over, and faster access to, existing interlanguage systems of knowledge.

Task Complexity and task sequencing

An example is given in Figure 2 of how the dimensions of task complexity just described can be selectively combined in an increasingly complex sequence of pedagogic task versions, following the SSARC model described previously. In this example, the intended target task, real world performance is giving passenger directions in the L2 to a driver on how to find a location, using a road map, while driving through an unknown area. The first pedagogic task version is simple on all the relevant dimensions of complexity the pedagogic version draws on (SS, simple/stabilizing interlanguage). Planning time is provided, and this version is performed while using a map on which the route has already been marked (so the speaker has only to describe the route, not think up the route while speaking). The map is of an area the speaker has prior knowledge of, and it is a map of a small area with few roads, buildings and other landmarks. On subsequent versions, the task is made more complex on resource-dispersing dimensions (A, automatizing access to interlanguage). So on the second version the task is performed without planning time, and on the third version there is no planning time and the route is not marked on the map. Additionally on the fourth version the map used is of an unknown area. Across these versions, therefore, the performative demands on accessing current L2 knowledge are gradually increased, promoting speedier access to it. On the final version the task is also made complex on the resource-directing dimension, so that it is performed with an authentic map of an unknown area, describing many more roads, buildings and landmarks which consequently have to be referred to and distinguished between using a greater number and variety of complex nominal expressions, involving the use of relative clauses (*the building which…, the road we turn left at, etc.*) and concatenated adjective-noun-location constructions (*the first narrow one way street on the left*, etc.). This promotes restructuring and complexifying (RC) of the learner's current level of interlanguage in order to meet the demands of the task, compared to previous versions.

The proposal that increasing the resource-directing demands of tasks leads to qualitative changes in the language learners produce in performing them has received support in a number of (but by no means all) studies of the dimensions of these demands described in Figure 1. Although full review of these findings is outside the scope of this chapter (see

Garcia-Mayo, 2007; Robinson, 2010; and Robinson & Gilabert, 2007 for recent collections of empirical studies in this area), one dimension that has perhaps been most frequently studied is the Here-and-Now/There-and-Then dimension of task complexity. For example, using general measures of speech production, Robinson (1995) found that when compared to production on a Here-and-Now task requiring participants to narrate a series of events illustrated by pictures as they were looking at them with the researcher, production on a There-and-Then task where narration was based on memory of absent pictures was more accurate (using measures of Target Like Use of articles) and of greater lexical density (using measures of Type-Token ratio). Gilabert (2005, 2007) reported similar findings, with greater lexical density, and more self-repair (indicating greater sensitivity to accuracy) on There-and-Then compared to Here-and-Now tasks. Ishikawa (2007) found that on written There-and-Then tasks there was greater syntactic complexity (S Nodes per T-unit) and also accuracy (% Error Free T-units) compared to Here-and-Now tasks. Robinson, Cadierno and Shirai (2009) found, as previously mentioned, greater use of past tense and progressive aspect to mark non-prototypical meanings on There-and-Then tasks, compared to Here-and-Now tasks. Finally, Cadierno and Robinson (2009) found that for learners with a typologically similar L1 to L2 English (i.e., Danish) complex There-and-Then tasks elicited more target-like lexicalization patterns for reference to motion (in which motion and manner were conflated on the verb and path expressed separately as a satellite, or series of satellites) compared to performance on Here-and-Now tasks. This brief summary illustrates an additional advantage of the theoretically motivated taxonomy of task characteristics described in Figure 1, since it not only guides systematic operationalizations of pedagogic decision making about task design and sequencing, but also serves as a focus for SLA research into the effects of task characteristics on speech production and learning, allowing cumulative findings for task characteristics such as Here-and-Now/There-and-Then to emerge across studies.

Dimensions of complexity	**Simple** **1**	**2**	**3**	**4**	**Complex** **5**
planning time (before speaking)	+	–	–	–	–
single task (route marked)	+	+	–	–	–
prior knowledge (a familiar area)	+	+	+	–	–
few elements (a small area)	+	+	+	+	–
	(**simplified** data/map)				(**authentic** data/map)

Figure 2. Five versions of a direction-giving map task that increase in complexity following the sequencing principles of the SSARC model

Task Conditions

The taxonomy of task characteristics listed in Figure 1 distinguishes Task Complexity from *Task Conditions*. Characteristics described above, contributing to the intrinsic cognitive complexity are proposed to be the sole basis of sequencing decisions, following the procedures for sequencing described in the SSARC model. Task Conditions, in contrast, describe two categories of task characteristics that affect the nature and amount of interaction between participants in performing real world tasks and pedagogic task versions of these. The SSARC model proposes that these characteristics are identified as relevant to pedagogic task performance on the basis of the initial behavioral needs analysis of target task performance during the first stage of instructional design, and are then held constant each time pedagogic versions of target tasks are performed in increasingly cognitively complex sequences. The rationale for this is that replicating the interactive demands of target tasks each time simple to more complex pedagogic task versions are performed most effectively embeds in *memory*, and leads to *rehearsal* and *elaboration* of the *scripts and schemata* for interactive task performance that the target task requires be accomplished. This proposal draws on the rational for case-based reasoning, dynamic memory, and schema learning and elaboration proposed by Schank (see Schank & Abelson, 1977; Schank, 1999; Schank, Berman & MacPherson, 1999).

Task characteristics listed under the category Task Conditions in Figure 1 are of two kinds. Participation variables make differing *interactional* demands on task performance, such as whether the task requires information transmission from one person to another (+ one way) as when leaving a message on an answerphone, or reciprocally, both ways (+ two way) as when filling in a form requesting a bank transfer face-to-face (see Pica, Kanagy & Falodun, 1993). Similarly, task participation may require learners to identify a single correct solution (+ closed) or leave them free to propose any solution (+ open) as in the case of exchanging opinions in the L2 about how best to regulate greenhouse gas emissions, or about which movie they would prefer to see. Participant variables make *interactant* demands, and these are a consequence of similarities and differences between participants in interactional role, gender, proficiency, shared background knowledge and other factors (see e.g., Plough & Gass, 1993; Yule & MacDonald, 1990).

Task Difficulty

The third category in Figure 1, *Task Difficulty*, concerns not characteristics of *tasks*, but the *abilities and affective factors* which learners bring to task performance and learning. Two learners differing in the abilities contributing to aptitude for mathematics will find performing calculus, for example, more or less difficult than each other, helping explain *between participant variation* in success on any one task. This contrasts with Task Complexity, as I have described it, which concerns the intrinsic differences in the cognitive demands any two tasks place on an individual. Whether learners are high or low in aptitude for mathematics, for any *one* learner calculus will always be intrinsically more complex than simple addition, helping explain *within participant variation* in relative success on the two tasks (see Robinson, 2001a, 2001b; Spilsbury, Stankov & Roberts, 1990; Wood, 1986).

The SSARC model of task sequencing proposes that only characteristics of task complexity are manipulated during task design for learners, since only these characteristics can be mapped operationally from target task description to pedagogic task design. Additionally, learner strengths and deficits in the ability and other factors described in Figure 1 are difficult to measure prior to participation in instructional programs, and so information about them can play no role in *a priori* syllabus design of the kind I have been describing, which maps target task analyses to increasingly complex pedagogic task versions. Wood (1986) describes essentially the same approach: 'If we wish to separate individual and task effects then we should logically expect to describe tasks independently of individuals who perform the task. … However, the 'task qua task' approach leaves the researcher with the difficult problem of identifying appropriate … analytic dimensions for the description of task characteristics' (Wood, 1986: 62). The characteristics contributing to task complexity in Figure 1 are a synthesis of my own and other proposals for such analytic dimensions. However, for reasons described in the introduction to this chapter, it will be important to map demands of tasks along these dimensions of complexity to the abilities and affective factors promoting success on them so that *in situ*, courses of instruction can use this information to match learners to tasks and sequences of them that they are optimally suited to (complex task sequences they will find least 'difficult'), and also to support learning and task participation by learners low in the abilities and other factors any particular task or task sequence draws on (see Robinson, 2001c, 2002a, 2005b, 2007b, 2009b).

3. Mapping Task Complexity/Condition – Task Difficulty interactions

It is not yet clear what the ability and affective factors are that contribute to perceptions of Task Difficulty, and so both promote and mitigate successful performance on the simple and complex task characteristics listed under the category of Task Complexity in Figure 1. Some of the factors listed under Task Difficulty have begun to be researched for their effects on task-based learning, but others have not. Essentially, therefore, the selective listing of the ability and affective Task Difficulty factors in Figure 1 is intended as a starting point – a rationalization, and conceptual synthesis of research and findings in cognitive and educational psychology that is judged to be relevant to, and have implications for, this Task Complexity–Task Difficulty–interaction research agenda.

Cognitive ability variables listed in Figure 1 include some that have been shown to have robust influences on success during L2 instruction in general, the most notable being 'aptitude' (see e.g., Carroll, 1981; Robinson, 2002c, 2005b, 2009b; Skehan, 1989, 2002 for discussion). But it is clear from actual tests of L2 aptitude (e.g., the Modern Language Aptitude Test, see Carroll & Sapon, 1959) that what is measured by them are a number of quite distinct but positively intercorrelated subtests of cognitive abilities (e.g., rote memory capacity; phonemic sensitivity) which contribute to a superordinate factor, labeled *aptitude*, in much the same way that tests of intelligence (e.g., the Wechsler Adult Intelligence Scale) have many different subtests of verbal and non-verbal, or crystallized and fluid abilities that together contribute to a superordinate *G*, or *general intelligence* factor

(see Deary, 2000; Nyborg, 2003; Sternberg, 1990, for discussion). But general intelligence is a factor *inferred* (not understood as a domain independent, cognitive construct) from the 'positive manifold', or consistent positive intercorrelations of scores on the many subtests contributing to an overall IQ score. Intelligence researchers working within what is called '*G*-theory' generally accept that there are subcomponents of intelligence, such as emotional intelligence (*G*e) and motivational intelligence (*G*m) (see Bar-On & Parker, 2000), as well as fluid Intelligence (*G*f) and crystallized intelligence (*G*c) (see Mackintosh, 1998) as did Carroll himself in his synthesis of findings from research into the effects of cognitive abilities on learning and performance across wide domains of intellectual activity (Carroll, 1993).

In what follows, then, in synthesizing the research into cognitive abilities and affective factors, and proposing how they may be related to the demands of task characteristics listed in Figure 1, I will take the position (in line with Snow, 1987, 1994) that there are multiple 'aptitudes' for performing tasks that are a consequence of the fact that different task characteristics draw on different sets, or complexes of cognitive abilities and affective factors. Task aptitudes are person-in-situation 'transactions' between *some* cognitive abilities and affective dispositions that can be expected to be drawn on by *some* task characteristics, but not by others. There is no one 'aptitude' factor that will uniformly and positively predict success on tasks with each of the characteristics Figure 1 identifies. In what follows I speculate on what these aptitudes for task-based learning and performance might be.

Cognitive abilities and Task Complexity

As is well known, when the complexity of a variety of cognitive tasks used in psychological research increase in the complexity of their demands on learners, so too the positive correlation of successful performance on these tasks and scores on measures of a range of cognitive abilities increases (see Snow, Kyllonen & Marshalek, 1984). Individual differences in the cognitive abilities learners bring to pedagogic tasks have their clearest effects on learning and performance on complex versions (Robinson, 2005a, 2007b, 2007c). Yet it is in performing complex versions of tasks that progress in educational programs hinges, and so it is particularly important to research and understand what these complex task demand, ability interactions are. Four of these are nominated as worthy of further research below.

Working memory *capacity* is one ability factor that has been increasingly researched in the SLA literature, showing for example generally positive correlations with incidental learning (Niwa, 2000; Robinson, 2002b, 2005b, 2005c; Williams, 1999) and in successful uptake of implicit negative feedback delivered by recasts (Mackey, Egi, Philp, Fuji & Tatsumi, 2002), both of which task-based learning environments encourage and provide opportunities for. I would argue working memory capacity (and there are many measures of this, see Cowan, 2005) is likely implicated *specifically* in successful performance on There-and-Then tasks which require learners to hold in memory a description of some event, while verbalizing it *concurrently*, and also in performance on tasks requiring dual, *simultaneous* performance of subtasks (answering a phone call while monitoring a TV screen). Related to this latter dimension of task complexity, task-switching (see Monsell, 2003 for measures of this) involves the executive *control* ability to switch back and forth

quickly and efficiently from the demands of one component of a task to another component (e.g., from monitoring a fuel gauge to talking to a co-driver) and is also likely positively related to the abilities drawn on in performing dual-tasks, as opposed to single tasks, as these characteristics are described in Figure 1. Similarly, there are many measures of abilities drawn on in successfully reasoning about causal relations (see Lohman, 2000; Stanovitch, 2000) and these are likely related positively to successful performance on tasks making causal reasoning demands listed in Figure 1 (such as explaining why a bridge fell down in a thunderstorm, or why a marketing campaign will result in greater sales revenue). More recently research has begun into measuring the cognitive abilities affecting the extent to which people are able to successfully attribute intentions and mental states to others, and to reason from this to a conclusion about why others performed certain actions (Goldman, 2006; Langdon, Coltheart, Ward & Catts, 2002; Malle, 2004). These measures of intentional reasoning ability are clearly likely to be related to success on complex L2 tasks high in the intentional reasoning demand characteristic nominated in Figure 1 (such as explaining why four or more people in the workplace fell into dispute over a breach of office protocol, and as a consequence why subsequently their personal relationships with each other changed).

Task Complexity		Task Conditions	
	i. Resource-directing, developmental dimensions	**i. Participation, interactive factors**	
T		**A**	
H	+/– here and now	**F**	+/– open solution
I	+/– few elements	**C**	+/– one way flow
N	–/+ spatial reasoning	**O**	+/– convergent solution
K	–/+ causal reasoning	**N**	+/– few participants
I	–/+ intentional reasoning		+/– few contributions needed
N	–/+ perspective taking	**1**	+/– negotiation not needed
G			
	(Gc)(e.g., IDs in reasoning)		**(Ge, Gm)** (e.g., IDs in tolerance of ambiguity)
	ii. Resource-dispersing, performative dimensions	**ii. Participant, interlocutor factors**	
R		**A**	
E	+/– planning time	**F**	+/– same proficiency
A	+/– prior knowledge	**C**	+/– same gender
C	+/– single task	**O**	+/– familiar
T	+/– task structure	**N**	+/– shared content knowledge
I	+/– few steps		+/– equal status and role
N	+/– independency of steps	**2**	+/– shared cultural knowledge
G			
	(Gf) (e.g., IDs in flexible attention, task switching)		**(Ge, Gm)** (e.g., motivational intensity, control, anxiety and self-efficacy)

Figure 3. Mapping Task Complexity/Task Condition – Task Difficulty interactions to '*G*' and aptitudes for learning and performance

In general then, I would expect individual differences in cognitive abilities, nominated in Figure 1 as contributing to perceptions of Task Difficulty, to interact with characteristics of tasks contributing to their Task Complexity in Figure 1 – inhibiting or promoting successful adaptation to, and so successful learning and performance on tasks having these characteristics – and I have identified above, four such potential interactions worthy of future research. But a broader summary of the potential for learner ability, task complexity interactions is possible. Figure 3 describes it, and I present this broader summary here.

Broad summary of potentials for Task Complexity–Task Difficulty interactions
Along resource-directing dimensions of Task Complexity, task performance requires the cognitive abilities drawn on when *engaging in complex thought* (i.e., THINKING) in different *conceptual* domains. I would argue that these cognitive abilities will be closely related to measures of crystallized intelligence (*G*c) operationalised in intelligence tests, such as the ability to reason about a domain. In contrast, along resource-dispersing dimensions of Task Complexity, tasks also require that learners *act on their thoughts* in complex *performative* environments (i.e. REACTING) and I would argue that the cognitive abilities contributing to this are likely to be related to measures of fluid intelligence (*G*f) operationalised in intelligence tests, such as the ability to switch attention rapidly between concurrent task demands – an ability involving measurable differences in processing speed and the divisibility of attention.

Affective factors and Task Conditions
While the ability factors nominated in Figure 1 are likely most influential when performing tasks differing in complexity, affective (personality) and conative (motivational) factors (what Snow, 1987, called jointly AFCON factors) nominated in Figure 1 are likely more influential in performing tasks under different Task Conditions and the demands they impose. Figure 3 also illustrates this broad relationship. During pedagogic task performance in complex interactional and interpersonal environments, adaptation to the various *participation* and *participant* characteristics nominated in Figure 1 is facilitated by greater self regulation and motivational intensity (*G* m factors), and emotional control (a *G* e factor) (see e.g., Collis & Messick, 2001; Dornyei, Csizer & Nemeth, 2006; Forgas, 2001; MacIntyre, 2002; Matthews & Deary, 1998; Mayer, Salovey & Caruso, 2000; Snow & Farr, 1987). For example, when the solution to a task learners are performing is indeterminate and not fixed (+ open) as opposed to determinant and fixed (+ closed) then individual differences in measures of emotional control, such as openness to experience and tolerance of ambiguity (Costa & Macrae, 1985; Furnham & Ribchester, 1995) may predict more, or less, successful engagement in task participation to meet these goals (with those more open to experience and more tolerant of ambiguity adapting better to participation in open tasks, and vice versa). Similarly, when one participant in a task is at a lower level of proficiency than his or her partner, then those lower proficiency learners with greater motivational intensity and control (see Dornyei, 2005), lower output anxiety (see MacIntyre, 2002) and greater self-efficacy (Bandura, 1997) may make greater efforts to participate, and so benefit more from the L2 interaction, than those lower in motivational

intensity, self-efficacy and higher in output anxiety (see for related recent work in these areas Dornyei, 2002; Oshina, 2009; Robinson, 2007c; and Sheen, 2008).

4. Summary and conclusions: Cognitive abilities, cognitive processes and task sequencing – The need to research interactions

The research agenda described in this chapter places Cronbach's (1975) call for a unification of the two approaches to psychology – correlational studies of the effects of individual differences on performance, and experimental studies of the effects of tasks and conditions of exposure on learning – in contemporary perspective. Multiplicative interaction models will be needed to thoroughly address, quantitatively, the twin issues for educational research I have raised, about how to sequence pedagogic task characteristics and also how to map components of task complexity to individual differences between learners so as to optimize learning and performance for groups and individuals. Such statistical models are commonly applied to research into interactions in other domains, such as political science (see Brambor, Clark & Golder, 2006), which could be adopted to serve these purposes. Qualitative research too, into learner perceptions of task difficulty in response to different task demands, will be needed, such as coding task participants' responses to questions about task performance and the difficulties they encountered and non-participant observation of task performance on tasks at various proposed levels of complexity. Such research could lead to a much needed metric for assessing second language learning task workload, across various pedagogic tasks and settings for performance, and there are analogues for this in ergonomics research such as Parks and Boueck's (1989) *Time Line Analysis* or North and Riley's (1989) *Workload Index* that could be usefully followed.

The need for interactionist research in educational contexts, such as that profiled in this chapter, has long been voiced (see e.g., Cronbach & Snow, 1977; Snow, Federico & Montague, 1980) and I return to one of these voices in conclusion. Writing thirty years ago Thomas Shuell raised many of the same issues I have addressed above: 'In order to improve performance on the overall task, the learner may have to identify and become proficient in performing certain subcomponents of the task that are a prerequisite to performance on the overall task' (Shuell, 1980: 281). For 'overall task' here, read 'target real-world tasks' as I have described them, and for 'subcomponents' read the components of task complexity and task conditions proposed in Figure 1. Schuell continues; 'A theory of instruction would be concerned with specifying – probably by means of a task analysis – what those important subcomponents are and the sequence in which the student should practice on the various subcomponents' (Shuell, 1980: 281). The SSARC model I have described, in conjunction with the taxonomic framework illustrated in Figure 1, aims to provide such a specification. Concerning the nature of individual differences Shuell notes: 'Differences between learners are virtually limitless, and it is possible to define or describe these in various ways. In considering the role of individual differences … in adapting instruction to the needs of individual students, some consideration must be given to the types of individual differences that are most appropriate for these concerns. Criteria must be developed that will permit us to determine which individual differences are important and which are

trivial' (Shuell, 1980: 289). Researching the interactions of the factors nominated as contributing to Task Difficulty in Figure 1, with performance on tasks having one or another characteristic of Task Complexity or Condition is a coherent way to begin to develop such criteria, I have argued. And finally, Shuell comments: 'There are a variety of ways that matches between the learner's characteristics and the optimal learning environment for that student can be made. Before any effective matching can be done, however, it must be possible to classify the characteristics of both students and learning environments in ways that will permit a meaningful match. Taxonomies of both relevant individual differences and appropriate task environments are required' (Shuell, 1980: 297). Such a taxonomy of task and learner characteristics has been described in this chapter. Whether it can be used, following the SSARC model, to optimally sequence pedagogic tasks for learners, and to optimize instruction for individual learners on tasks, and across task sequences, are issues pedagogic practice and further research will only be able to fully resolve together.

References

Achard, M. (2008). Teaching construal: Cognitive pedagogical grammar. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, (pp. 431–455). London: Routledge.

Achard, M. & Niemeier, S. (2004). (Eds.), *Cognitive Linguistics, Second Language Acquisition and foreign language teaching.* Berlin: Mouton de Gruyter.

Ackerman, P. (2003). Aptitude complexes and trait complexes. *Educational Psychologist, 38*, 85–93.

Altman, H. (1972). (Ed.), *Individualizing the foreign language classroom*. Rowley, MA: Newbury House.

Andersen, R. (1984). The one-to-one principle of interlanguage construction. *Language Learning, 34,* 77–95.

Andersen, R., & Shirai, Y. (1996). The primacy of aspect in first and second language acquisition – The pidgin-creole connection. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition,* (pp. 527–570). San Diego, CA: AcademicPress.

Astington, J., & Baird, J. (2005). (Eds.), *Why language matters for theory of mind*. Oxford: Oxford University Press.

Bandura, A. (1997). *Self-efficacy: the exercise of control.* New York: W. H. Freeman & Co.

Bar-On, R., & Parker, D. A. (2000). *The Handbook of emotional intelligence*. San Francisco, CA: Jossey Bass.

Bartsch, K., & Wellman, H. (1995). *Children talk about the mind*. Oxford: Oxford University Press.

Behrens, H. (2001). Cognitive-conceptual development and the acquisition of grammatical morphemes: The development of time concepts and verb tense. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development*, (pp. 450–474). Cambridge: Cambridge University Press.

Berman, R., & Slobin, D. (1994). (Eds.), *Relating events in narrative: A crosslinguistic developmental study.* Mahwah, NJ: Erlbaum.

Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition, 16*, 157–168.

Bowerman, M., & Levinson, S. (2001). (Eds.), *Language acquisition and conceptual development*. New York: Cambridge University Press.

Brambor, T., Clark, W., & Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political Analysis, 14*, 63–82.

Cadierno, T., & Robinson, P. (2009). Language typology, task complexity and the development of L2 lexicalization patterns for describing motion events. *Annual Review of Cognitive Linguistics*, *7*, 246–277.

Candlin, C. (1984). Syllabus design as a critical process. In C. J. Brumfit (Ed.), *General English syllabus design*, (pp. 29–46). Oxford: Pergamon.

Carassa, A., Aprigliano, A., & Geminiani, G. (2000). Describers and explorers: A method for investigating cognitive maps. In S. O'Nuallain (Ed.), *Spatial cognition,* (pp. 33–43). Amsterdam: John Benjamins.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude,* (pp. 83–118). Rowley, MA: Newbury House.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Carroll, J. B., & Sapon, S. M. (1959). *Modern Language Aptitude Test*. Washington, DC: Second Language Testing Incorporated.

Collis, J., & Messick, S. (2001). (Eds.), *Intelligence and personality: Bridging the gap in theory and measurement.* Mahwah, NJ: Erlbaum.

Cook, V. (2003). (Ed.), *Effects of the second language on the first.* Clevedon: Multilingual Matters.

Cornell, E., Heth, D., & Alberts, D. (1994). Place recognition and way finding by children and adults. *Memory and Cognition, 22*, 633–643.

Costa, P. T. & McCrae, R. R. (1985). *The NEO Personality Inventory Manual.* Odessa, FL: Psychological Assessment Resources, Inc.

Cowan, N. (2005). *Working memory capacity*. New York: Psychology Press.

Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal learning and Verbal Behavior, 11*, 671–684.

Crombie, W. (1985). *Process and relation in discourse and language learning*. Oxford: Oxford University Press.

Cronbach, L. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist, 30,* 116–127.

Cronbach, L., & Snow, R. (1977). *Aptitudes and instructional methods*. New York: Irvington.

de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory to second language acquisition. *Bilingualism: Language and Cognition, 10*, 7–21.

Deary, I. (2000). *Looking down on human intelligence: From psychometrics to the brain*. Oxford: Oxford University Press.

Deissel, H. (2004). *The acquisition of complex sentences*. Cambridge: Cambridge University Press.

DeKeyser, R. M. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction*, (pp. 125–151). Cambridge: Cambridge University Press.

DeKeyser, R. M. (2007). (Ed.), *Practicing in a second language: perspectives from applied linguistics and cognitive psychology.* New York: Cambridge University Press.

Dornyei, Z. (2002). The motivational basis of language learning tasks. In P. Robinson (Ed.), *Individual differences and instructed language learning*, (pp. 137–158). Amsterdam: Benjamins.

Dornyei, Z. (2005). *The psychology of the language learner.* Mahwah, NJ: Erlbaum.

Dornyei, Z., Csizer, K., & Nemeth, N. (2006). *Motivation, language attitudes and globalisation: A Hungarian perspective*. Clevedon: Multilingual Matters.

Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction*, (pp. 206–257). Cambridge: Cambridge University Press.

Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition*, (pp. 197–262). New York: Cambridge University Press.

Ehret, B., Gray, W., & Kirschenbaum, S. (2000). Contending with complexity: Developing and using a scaled world in applied cognitive research. *Human Factors, 2*, 8–23.

Eisenstein, M. (1989). (Ed.), *The dynamic interlanguage.* Mahwah, NJ: Erlbaum.

Ellis, N. C. (1998). Emergentism, connectionism and second language learning. *Language Learning, 48,* 631–664.

Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. Doughty & M. Long (Eds*.), The handbook of second language acquisition,* (pp. 63–103). Oxford: Blackwell.

Ellis, N. C. (2008). Usage-based and form-focused instruction: The associative learning of constructions, learned inattention and the limited endstate. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, (pp. 372–405). London: Routledge.

Ellis, N. C. & Larsen Freeman, D. (2006). Language emergence: Implications for Applied Linguistics – Introduction to the special issue. *Applied Linguistics, 27*, 558–589.

Ellis, N. C. & Robinson, P. (2008). An introduction to Cognitive Linguistics and Second Language Acquisition. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, (pp. 3–24). London: Routledge.

Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly, 27*, 91–113.

Ellis, R. (2003). *Task-based language teaching and learning*. Oxford: Oxford University Press.

Ellis, R. (2005). (Ed.), *Planning and second language task performance*. Amsterdam: Benjamins.

Filipović, L. (2007). *Talking about motion – A crosslinguistic investigation of lexicalization patterns*. Amsterdam: Benjamins.

Filipović, L. & Vidaković, I. (this volume). Typology in the L2 classroom: Second language acquisition from a typological perspective.

Fine, S. P. (1974). Functional job analysis: An approach to technology for manpower planning. *Personnel Journal*, *11*, 813–818.

Fleishman, E. A. & Quaintance, M. K. (1984). *Taxonomies of human performance: The description of human tasks.* New York: Academic Press.

Fries, C. C. (1952). *The structure of English*. New York: Harcourt Brace and Company.

Furnham, E., & Ribchester, T. (1995). Tolerance of Ambiguity: A review of the concept, its measurement and applications. *Current Psychology*, *14*, 179–199.

Gagne, R. M., Wager, W. W., Golas, K. C. & Keller, J. M. (2005). *Principles of Instructional Design.* (5th edition). Belmont, CA: Thomson, Wadsworth.

Garcia Mayo, M. P. (2006). (Ed.), *Investigating tasks in formal language learning.* Clevedon: Multilingual Matters.

Gilabert, R. (2005). *Task complexity and L2 oral narrative production*. Unpublished Ph.D dissertation, University of Barcelona, Department of Applied Linguistics, Spain.

Gilabert, R. (2007). The simultaneous manipulation along the planning time and +/− Here-and-Now dimensions: Effects on oral L2 production. In M. P. Garcia Mayo (Ed.), I*nvestigating tasks in formal language learning*, (pp. 44–68). Clevedon: Multilingual Matters.

Gilabert, R., Baron, J., & Llanes, M. A. (2009). Manipulating cognitive complexity across task types and its influence on learners' interaction during oral performance. *International Review of Applied Linguistics, 47,* 367–396.

Goldman, A. (2006). *Simulating minds*. Oxford: Oxford University Press.

Han Z., & Odlin, T. (2006). (Eds.), *Studies of fossilization in second language acquisition.* Clevedon: Multilingual Matters.

Hoffman, R., & Militello, L. (2009). *Perspectives on cognitive task analysis*. New York: Psychology Press.

Hollnagel, E. (2003). (Ed.), *Handbook of cognitive task design*. Mahwah, NJ: Erlbaum.

Huebner, T. (1983). *A longtitudinal analysis of the acquisition of English*. Ann Arbor, MI: Karoma.

Hulstijn, J. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of rehearsal, elaboration and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction*, (pp. 349–381). Cambridge: Cambridge University Press.

Hulstijn, J. (2003). Incidental and intentional learning. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition*, (pp. 631–678). Oxford: Blackwell.

Ishikawa, T. (2007). The effects of increasing task complexity along the –/+ Here-and-Now dimension. In M. P. Garcia Mayo (Ed.), *Investigating Tasks in Formal Language Learning,* (pp. 136–156). Clevedon: Multilingual Matters.

Ishikawa, T. (2008). *Task complexity, reasoning demands and second language speech production*. Unpublished Ph.D. dissertation, Aoyama Gakuin University, Department of English, Tokyo, Japan.

Karmiloff-Smith, A. (1992*). Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Kellerman, E. (1985). Input and second language acquisition theory. In S. Gass & C. Madden (Eds.), *Input in second language acquisition,* (pp. 345–353). Rowley, MA: Newbury House.

Kormos, J. (2006). *Speech production and second language acquisition.* Mahwah, NJ: Erlbaum.

Kuiken, F., & Vedder, I. (2007). Cognitive task complexity and linguistic performance in French L2 writing. In M. P. Garcia Mayo (Ed.), I*nvestigating tasks in formal language learning*, (pp. 117–135). Clevedon: Multilingual Matters.

Langdon, R., Coltheart, M., Ward, P. & Catts, S. (2002). Disturbed communication in schizophrenia: The role of pragmatics and poor theory-of-mind. *Psychological Medicine, 32*, 1273–1284.

Larsen-Freeman, D. (2006). The emergence of accuracy, fluency and complexity in the oral and written production of five Chinese learners of English. *Applied Linguistics, 27,* 590–619.

Larsen-Freeman, D., & Cameron, L. (2007). *Complex systems and applied ling*uistics. Oxford: Oxford University Press.

Larsen-Freeman, D., & Long, M. H. (1991*). An introduction to second language acquisition research.* London: Longman.

Ledford, B., & Sleeman, P. (2002). *Instructional design: System strategies*. Connecticut, VM: Information Age Publishing.

Lee, C. E., & Rescorla, L. (2002). The use of psychological state terms by late talkers at age 3. *Applied Psycholinguistics, 23*, 623–641.

Lesgold, A. (2001). The nature and methods of learning by doing. *American Psychologist*, *56*, 964–973.

Li, P., & Shirai, Y. (2000). *The acquisition of lexical and grammatical aspect*. Berlin: Mouton de Gruyter.

Lohman, D. (2000). Complex information processing and intelligence. R. Sternberg (Ed.), *Handbook of Intelligence*, (pp. 285–340). New York: Cambridge University Press.

Lohman, H., & Tomasello, M. (2003). The role of language in the development of false belief Understanding: A training study. *Child Development*, *74*, 1130–1144.

Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modeling and assessing second language acquisition* (pp. 77–99). Clevedon, Avon: Multilingual Matters.

Long, M. H. (1989). Task, group, and task-group interactions. *University of Hawai'i Working Papers in ESL, 8,* 1–25.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspectives* (pp. 39–52). Amsterdam: Benjamins.

Long, M. H. (2005). (Ed.), *Second language needs analysis*. Cambridge: Cambridge University Press.

Long, M. H. (2006). *Problems in SLA*. Mahwah, NJ: Erlbaum.

Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly, 26*, 27–56.

Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition*, (pp. 15–41). New York: Cambridge University Press.

MacIntyre, P. (2002). Motivation, anxiety and emotion in second language acquisition. In P. Robinson (Ed.), *Individual differences and instructed language learning*, (pp. 45–64). Amsterdam: Benjamins.

Mackey, A. (2007). (Ed.), *Conversational interaction in SLA*. Oxford: Oxford Unversity Press.

Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning*, (pp. 181–210). Amsterdam: Benjamins.

Mackintosh, N. J. (1998). *IQ and human intelligence*. Oxford: Oxford University Press.

MacWhinney, B. (2001). The competition model: the input, the context, and the brain. In P. Robinson (Ed.), *Cognition and second language instruction*, (pp. 69–90). Cambridge: Cambridge University Press.

MacWhinney, B. (2008). A unified model. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, (pp. 441–371). London: Routledge.

Malle, B. (2004). *How the mind explains behavior*. Cambridge, MA: MIT Press.

Matthews, G. & Deary, I. (1998). *Personality traits.* Cambridge: Cambridge University Press.

Mayer, J. D., Salovey, P., & Caruso, D. (2000). Models of emotional intelligence. In R. J. Sternberg (Ed*.) Handbook of intelligence,* (pp. 396–420). New York: Cambridge University Press.

McCormick, E. J. (1979*). Job Analysis: Methods and Applications*. New York: AMACOM Publishing.

Meisel, J. (1987). Reference to past events and actions in the development of natural second language acquisition. In C. Pfaff (Ed.), *First and second language acquisition processes,* (pp. 206–224). Cambridge, MA: Newbury House.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7* (3), 134–140.

Munoz, C. (2006). (Ed.), *Age and the rate of second language acquisition*. Clevedon: Multilingual Matters.

Niwa, Y. (2000). *Reasoning demands of L2 tasks and L2 narrative production: Effects of individual differences in working memory, intelligence and aptitude*. Unpublished M.A. dissertation, Aoyama Gakuin University, Department of English, Tokyo, Japan.

Nixon, S. (2005). Mental state verb production and sentential complements in four-year-old children. *First Language, 25*, 19–39.

North, R., & Riley, V. (1989). WINDEX: A predictive model of operator workload. In G. McMillan, D. Beevis, E. Sala, M. Strub, R. Sutton & L. Van Breda (Eds.), *Applications of human performance models to systems design*, (pp. 81–89). New York: Plenum.

Nuevo, A. (2006). *Task complexity and interaction: L2 learning opportunities and interaction.* Unpublished Ph.D dissertation, Georgetown University, Washington DC, USA.

Nyborg, H. (2003). (Ed.), *The scientific study of general intelligence*. New York: Pergamon.

Odlin, T. (2008). Conceptual transfer and meaning extensions. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, (pp. 306–340). London: Routledge.

Ortega, L. (2009). Sequences and processes in language learning. In M. H. Long & C. Doughty (Eds.), *Handbook of language teaching*. Oxford: Blackwell.

Ortega, L., & Byrnes, H. (2008). (Eds.), *The longtitudinal study of advanced L2 capacities*. New York: Routledge.

Oshina, Y. (2009). *Recasts and anxiety: Uptake of pseudo-cleft constructions and plural forms*. Unpublished M.A. dissertation, Aoyama Gakuin University, Department of English, Tokyo, Japan.

Parks, D., & Boueck, G. (1989). Workload prediction, diagnosis and continuing challenges. In G. McMillan, D. Beevis, E. Sala, M. Strub, R. Sutton & L. Van Breda (Eds.), *Applications of human performance models to systems design*, (pp. 47–64). New York: Plenum.

Pavlenko, A., & Jarvis, S. (2007). *Crosslinguistic influence in language and cognition*. New York: Routledge.

Perdue, C. (1993a). (Ed.), *Adult language acquisition: Cross-linguistic perspectives, Vol. 1: Field methods.* Cambridge: Cambridge University Press.

Perdue, C. (1993b). (Ed.), *Adult language acquisition: Cross-linguistic perspectives, Vol. 2: The results*. Cambridge: Cambridge University Press.

Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language teaching and research. In G. Crookes & S. Gass (Eds.), *Tasks in language learning: Integrating theory and practice* (pp. 9–34). Clevedon, Avon: Multilingual Matters.

Plough, I. & Gass, S. (1993). Interlocutor and task familiarity: Effects on Interactional structure. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 95–122). Clevedon, Avon: Multilingual Matters.

Prabhu, N. S. (1987). *Second Language Pedagogy*. Oxford: Oxford University Press.

Pütz, M., Niemeier, S., & Dirven, R. (2001a). (Eds.), *Applied Cognitive Linguistics 1: Theory and language acquisition.* Berlin: Mouton de Gruyter.

Pütz, M., Niemeier, S., & Dirven, R. (2001b). (Eds.), *Applied Cognitive Linguistics I1: Language pedagogy.* Berlin: Mouton de Gruyter.

Rahimpour, M. (1997). *Task condition, task complexity and variation in oral L2 discourse.* Unpublished Ph.D. dissertation, University of Queensland, Brisbane, Australia.

Rahimpour, M. (1999). Task complexity and variation in interlanguage. In N. Jungheim & P. Robinson (Eds.), *Pragmatics and pedagogy: Proceedings of the 3rd Pacific Second Language Research Forum, Vol. 2.,* (pp. 115–129). Tokyo: PACSLRF.

Révész, A. (2007). Focus on form in task-based language teaching: recasts, task complexity, and L2 learning. Unpublished Ph.D. dissertation. Columbia University, New York, USA.

Riegeluth, C. M. (1999). (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory, Volume II*. Mahwah, NJ: Erlbaum.

Ringbom, H. (2006). *Cross-linguistic similarity in foreign language learning.* Clevedon: Multilingual Matters.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning*, *45*, 99–140.

Robinson, P. (1996). Connecting tasks, cognition and syllabus design. In P. Robinson (Ed.*), Task complexity and second language syllabus design: Data-based Studies and Speculations*, (pp. 1–16). Brisbane: University of Queensland Working Papers in Applied Linguistics.

Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced and rule-search conditions. *Studies in Second Language Acquisition*, *19*, 223–247.

Robinson, P. (2001a). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*, 27–57.

Robinson, P. (2001b). Task complexity, cognitive resources, and syllabus design: A triadic framework for investigating task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction,* (pp. 287–318). Cambridge: Cambridge University Press.

Robinson, P. (2001c). Individual differences, cognitive abilities, aptitude complexes, and learning conditions in SLA. *Second Language Research*, *17*, 368–392.

Robinson, P. (2002a). Learning conditions, aptitude complexes and SLA: A framework for research and pedagogy. In P. Robinson (Ed.), *Individual differences and instructed language learning*, (pp. 113–133). Amsterdam: Benjamins.

Robinson, P. (2002b). Individual differences in intelligence, aptitude and working memory during adult incidental second language learning: A replication and extension of Reber, Walkenfeld and Hernstadt (1991). In P. Robinson (Ed.), *Individual differences and instructed language learning*, (pp. 211–266). Amsterdam: Benjamins.

Robinson, P. (2002c). (Ed.), *Individual differences and instructed language learning.* Amsterdam: Benjamins.

Robinson, P. (2003a). Attention and memory during SLA. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition*, (pp. 631–678). Oxford: Blackwell.

Robinson, P. (2003b). The Cognition Hypothesis, task design and adult task-based language learning. *Second Language Studies, 21,* (2), 45–107. <http://www.hawaii.edu/sls/ uhwpesl/ 21 (2)/Robinson.pdf>

Robinson, P. (2005a). Cognitive complexity and task sequencing: A review of studies in a Componential Framework for second language task design. *International Review of Applied Linguistics, 43,* 1–32.

Robinson, P. (2005b). Aptitude and second language acquisition. *Annual Review of Applied Linguistics, 25,* 45–73.

Robinson, P. (2005c). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition, 27*, 235–268.

Robinson, P. (2007a). Criteria for classifying and sequencing pedagogic tasks. In M. P. Garcia Mayo (Ed.), *Investigating tasks in formal language learning,* (pp. 7–27). Clevedon: Multilingual Matters.

Robinson, P. (2007b). Aptitudes, abilities, contexts and practice. In R. M. DeKeyser (Ed.), *Practice in second language learning: Perspectives from applied linguistics and cognitive psychology,* (pp. 256–286). Cambridge: Cambridge University Press.

Robinson, P. (2007c). Task complexity, theory of mind, and intentional reasoning: Effects on speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics, 45*, 193–214.

Robinson, P. (2009a). Syllabus design. In M. H. Long & C. Doughty (Eds.), *Handbook of language teaching*, (pp. 294–310). Oxford: Blackwell.

Robinson, P. (2009b). Individual differences, aptitude complexes, SLA processes and the development of language learning aptitude tests. In M. Pawlak (Ed.), *Individual differences in second language learning and teaching: Achievements, prospects and challenges,* (pp. 89–109). Poznan: Adam Mickievicz University Press.

Robinson, P. (2010). (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance.* Amsterdam: Benjamins.

Robinson, P., Cadierno, T., & Shirai, Y. (2009). Time and motion: Measuring the effects of the conceptual demands of tasks on second language production. *Applied Linguistics*, *30*, 533–554.

Robinson, P., & Ellis, N. C. (2008a). (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition.* London: Routledge.

Robinson, P. & Ellis, N. C. (2008b). Cognitive Linguistics, Second Language Acquisition and instruction – Issues for research. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, (pp. 489–546). London: Routledge.

Robinson, P., & Gilabert, R. (2007). (Eds.), *Task complexity, the Cognition Hypothesis and second language instruction*. [Special issue] *International Review of Applied Linguistics, 45*(3). Berlin: Mouton DeGruyter.

Robinson, P., & Ha, M. (1993). Instance theory and second language rule learning under explicit conditions. *Studies in Second Language Acquisition*, *15*, 413–438.

Samuda, V., & Bygate, M. (2008). *Tasks in second language learning.* London: Palgrave MacMillan.

Sato, C. (1990). *The syntax of conversation in interlanguage development.* Tubingen: Gunter Narr.

Schank, R. (1999). *Dynamic memory revisited*. New York: Cambridge University Press.

Schank, R., & Abelson, R. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Erlbaum.

Schank, R., Berman, T., & Macpherson, K. (1999). Learning by doing. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory, Vol. II*, (pp. 161–182). Mahwah, NJ: Erlbaum.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction*, (pp. 1–32). Cambridge: Cambridge University Press.

Segalowitz, N. (2003). Automaticity and second language acquisition. In C. Doughty and M. H. Long (Eds.), *Handbook of second language acquisition,* (pp. 382–408). Oxford: Blackwell.

Sheen, Y. (2008). Recasts, language anxiety, modified output and L2 learning. *Language Learning, 58*, 833–874.

Shirai, Y. (1999). The prototype hypothesis of tense aspect acquisition. In P. Robinson (Ed.), *Representation and process: Proceedings of the 3rd Pacific Second Language Research Forum, Vol 1,* (pp. 151–164). Tokyo: PACSLRF.

Shuell, T. (1980). Learning theory, instructional theory and adaptation. In R. Snow, P. Federico & W. Montague (Eds.), *Aptitude, learning and instruction, Vol. 2: Cognitive process analyses of learning and problem solving*, (pp. 277–302). Mahwah, NJ: Erlbaum.

Skehan, P. (1989). *Individual differences in second language acquisition.* London: Arnold.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2002). Theorizing and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning*, (pp. 69–94). Amsterdam: Benjamins.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction,* (pp. 183–205). Cambridge: Cambridge University Press.

Slobin, D. (1993). Adult language acquisition: A view from child language study. In C. Perdue (Ed.), *Adult language acquisition: Crosslinguistic perspectives, Vol. 2: The results*, (pp. 239–252). Cambridge: Cambridge University Press.

Slobin, D. (1996). From 'thought and language' to 'thinking for speaking'. In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity, (*pp. 70–96.) Cambridge: Cambridge University Press.

Slobin, D. (2003). Language and thought online: Some cognitive consequences of linguistic relativity. In D. Gentner & S. Godlin Meadow (Eds.), *Language in mind: Advances in the investigation of language and thought,* (pp. 157–191). Cambridge, MA: MIT Press.

Snow, R. E. (1987). Aptitude complexes. In R. E. Snow & M. J. Farr (Eds.), *Aptitude*, *learning and instruction*, *Vol. 3: Conative and affective process analysis*, (pp. 11–34). Hillsdale, NJ: Erlbaum.

Snow, R. E. (1994). Abilities in academic tasks. In R. J. Sternberg & R. K.Wagner (Eds.), *Mind in context: Interactionist perspectives on human intelligence*, (pp. 3–37). New York: Cambridge University Press.

Snow, R. E., & Farr, M. J. (1987). (Eds.), *Aptitude, learning and instruction Vol. 3: Conative and affective process analysis*. Hillsdale, NJ: Erlbaum.

Snow, R. E., Federico, P., & Montague, W. (1980). (Eds.), *Aptitude, learning and instruction Vols 1 & 2*. Hillsdale, NJ: Erlbaum.

Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence*, (pp. 47–103). Hillsdale, NJ: Erlbaum.

Sokal, R. R. (1974). Classification: Purposes, principles, progress, prospects. *Science, 185*, 1115–1123.

Spilsbury, G., Stankov, L., & Roberts, R. (1990). The effects of a task's difficulty on its correlation with intelligence. *Personality and Individual Differences, 11*, 1069–1077.

Stanovitch, K. (2000). *Who is rational? Studies of ndividual differences in reasoning*. Mahwah, NJ: Erlbaum.

Sternberg, R. (1990). *Metaphors of mind: Conceptions of the nature of intelligence*. New York: Cambridge University Press.

Talmy, L. (2000). *Towards a cognitive semantics, Vol. 1: Concept structuring systems.* Cambridge, MA: MIT Press.

Talmy, L. (2008). Aspects of attention in language. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, (pp. 27–38). London: Routledge.

Trim, J., & North, B. (2001). (Eds.), *Common European Framework of Reference for Languages: Learning, teaching and assessment*. Cambridge: Cambridge University Press.

Tyler, A. (2008). Cognitive linguistics and second language instruction. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, (pp. 456–488). London: Routledge.

Tyler, A., & Evans, V. (2003). *The semantics of English prepositions: Spatial scenes, embodied meaning and cognition*. Cambridge: Cambridge University Press.

Valdman, A. (1978). Communicative use of language and syllabus design. *Foreign Language Annals, 11,* 567–578.

Valdman, A. (1980). Communicative ability and syllabus design for global foreign language courses. *Studies in Second Language Acquisition, 3,* 81–96.

Van den Branden, K. (2006). (Ed.), *Task-based Language Education*. Cambridge: Cambridge University Press.

Van Ek, J. (1976). *The threshold level for modern language learning in schools*. London: Longman.

Van Geert, P. (2008). The Dynamic Systems approach in the study of L1 and L2 acquisition: An introduction. *Modern Language Journal, 92,* 179–199.

von Stutterheim, C. & Klein, W. (1987). A concept-oriented approach to second language studies. In C. Pfaff (Ed.), *First and second language acquisition processes*, (pp. 191–205). Rowley, MA: Newbury House.

von Stutterheim, C., & Nuese, R. (2003). Processes of conceptualisation in language production: Language specific perspectives and event construal. *Linguistics, 41*, 851–888.

White, R. (1988). *The ELT curriculum: Design, management, innovation*. Oxford: Blackwell.

Wilkins, D. (1976). *Notional syllabuses*. Oxford: Oxford University Press.

Williams, J. (1999). Attention, memory and inductive learning. *Studies in Second Language Acquisition, 21*, 1–48.

Williams, R. G. (1977). A behavioral typology of educational objectives for the cognitive domain. *Educational Technology*, *17*, 418–431.

Willis, D. (1990). *The lexical syllabus: A new approach to language teaching*. London: Collins.

Wood, R. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes, 37*, 60–82.

Wimmer, H., & Perner, T. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.

Yule, G., & MacDonald, M. (1990). Resolving referential conflicts in L2 interaction: The effect of proficiency and interactive role. *Language Learning*, *40,* 539–556.

CHAPTER 14

Typology in the L2 classroom

Second language acquisition from a typological perspective

Luna Filipović and Ivana Vidaković
University of Cambridge

1. Introduction

This paper shows how insights from cognitive linguistics in general and a cognitive linguistic typology in particular can be implemented in applied cross-linguistic studies of second language acquisition. Talmy's (1985) typology of languages is the chosen framework for our analysis. We contrast the lexicalization patterns of English and Serbian which are used to map universal cognitive categories within the domain of motion onto surface expression.[1] We demonstrate how the two languages differ within this particular typological framework and highlight the effects of the relevant typological differences on the ways languages are learned and taught. We argue that including the typological perspective espoused here in translation and second language teaching provides new ways in which to tackle some persistent difficulties for learners and practitioners in those fields.

2. Typological foundations

2.1 Talmy's typology

Talmy (1985, 2000) put forward a semantic typology of languages based on how they map cognitive notions onto words and constructions. This is a particularly beneficial way of contrasting languages because it emphasizes the universality of human linguistic behaviour while providing the necessary basis for teasing the language-specific features apart from the universally shared ones. Talmy singled out the crucial event components in a universal cognitive domain (motion), which are lexicalized across languages. However, when speakers use language-specific devices to express those universal event components, additional meanings can be conveyed. As a result, certain lexicalization patterns facilitate the presence of one kind of information over another (e.g. Path over Manner of motion), the patterns are then reinforced, with consequences potentially reaching beyond

the systemic differences and habitually affecting the content of expressions and language use in general. According to Talmy, a motion event has four basic components: *Motion* (presence of motion), *Figure* (the moving object), *Ground* (the reference-point object with respect to which the Figure moves) and *Path* (the course followed by the Figure with respect to the Ground). A typical example of an expression of a motion event would be: 'The man ran into the building'. In this sentence, 'the man' expresses the Figure, 'the building' is the Ground, and the Path is expressed by the particle 'into'. The verb root ('run') itself conflates *Manner* and Motion.

Talmy's typology divides all the languages into two major groups. Languages that characteristically map the *core schema of the event* onto the verb are *verb-framed languages* and those that do it out of the verb via "satellites" are defined as *satellite-framed languages*.[2] Satellites are defined as "certain immediate constituents of a verb root other than inflections, auxiliaries, or nominal arguments" (Talmy 1985: 102), like prefixes in Slavonic languages or adverbs like 'in' as in 'He walked in without saying a word'. The representational example of the relevant contrasts in the satellite-framed/verb-framed dichotomy would be the following examples in English and Spanish respectively:

(1) a. The man ran into the garden.
b. El hombre entró en el jardín corriendo.
('The man entered the garden running')

In English, the preposition 'into', conveys the core information about the Path of the moving Figure, whereas in Spanish it is the verb itself, 'entrar' ('enter'), that conveys this piece of information. Note also that supporting information about manner of motion is conveyed by the verb in English and by the gerundive 'corriendo' ('running') in Spanish. These patterns are the most predominant ones in English and Spanish respectively. This distinction is reflected in both lexical and combinatorial resources of the two languages. English has a large collection of verbs of motion that convey Manner ('walk', 'run', 'crawl', 'fly', etc.), combinable with a large collection of directional satellites ('in', 'up', 'to', 'across', etc.). They are easy to insert in the construction environment of a motion expression and have higher frequency than directional verbs. In Spanish, the verbs of inherent directionality ('entrar' – 'enter', 'bajar' – 'descend', 'subir' – 'ascend', etc.) are the preferred means of expression. Manner verbs are used only in restricted circumstances and the lexicon with regard to manner verbs is conspicuously less developed in this language (cf. Slobin 1997).[3]

2.2 The notion of satellites

It is essential to point out at this stage why Talmy's notion of satellite needed broadening and, in the end, a redefinition. A satellite in Talmy's terms refers to those elements that were traditionally known as adverbs. This becomes more obvious when we consider what Talmy insists on, namely that satellites for expressing Path should be distinguished from prepositions (and he focuses on analysing satellites only). It may be so because this sometimes seems to be a fuzzy area in English. Talmy insists that no confusion can occur in most Indo-European languages, where the two forms have quite distinct positional and grammatical characteristics. For example, in Latin, Classical Greek and Russian, the

satellite is the prefix to the verb while the preposition accompanies the noun and governs its case. English, perhaps alone among Indo-European languages, regularly places satellite and preposition next to each other in a sentence. Nevertheless, according to Talmy (1985), there are still ways in which the two kinds of forms – satellites and prepositions – are distinguishable. We shall explain why, even though it is sometimes possible, it is of no relevance for this kind of research.

Talmy contends that it is only a preposition that will disappear when the Ground nominal is omitted: a satellite remains (Talmy 1985: 105). Then, the two classes of forms do not have identical memberships: there are forms with only one function or the other. For example, 'together', 'apart', and 'forth' are satellites that never act as prepositions, while 'from', 'at', and 'toward' are prepositions that never act as satellites. He contends that the directional particles that follow the verb and are not followed by a noun are considered to be satellites, as opposed to prepositions, which are followed by a noun and cannot follow the verb on their own. This shows that Talmy follows the traditional way of distinguishing between (Path) prepositions and adverbs. However, basing the distinction on whether a noun phrase follows or not is nowadays considered syntactically inadequate. Huddleston (2000) and Huddleston & Pullum (2002) emphasize that is necessary to consider syntactic functions of a word, such as complement and modifier, in order to be able to assign it to a class of prepositions or adverbs.

Talmy (1985) offers evidence in support for his satellite/preposition distinction, which we think is not persuasive. He contends that English has a special feature, detected also for example in Mandarin Chinese (ibid.). There is a number of forms like 'past' that behave like ordinary satellites when there is no final nominal, as in: 'I saw him on the corner, but I just drove past', but appear without any preposition when there is a final nominal, as in 'I drove past him'. Forms like these have properties of both a satellite and a preposition. According to Talmy the distinction could be made on the basis of stress. He claims that they receive a heavy stress as a satellite, whereas as prepositions, they receive a light stress. Because of its special behaviour, according to Talmy (ibid.: 106) a form like 'past' might be considered "a coalesced version of satellite plus a preposition – a satellite-preposition".[4] However, while in 'He drove in', the last word (the adverb or Talmy's satellite 'in') is tonic, in the example 'He drove in through the gate', the last word, i.e. 'gate' gets the tonic stress. Although according to Talmy 'in' should be a satellite here as well, it does not get the heavy (tonic) stress. On the other hand, in the example provided by Talmy that is supposed to illustrate the fact that when an element is used as a preposition it does not have a heavy (tonic) stress, we clearly see that this is not the case. In 'He drove past him', 'past' would normally have a tonic stress, unless the speaker wanted to emphasize the pronoun at the end, e.g. 'He drove past *him*, not *her*'. Thus, we can completely disregard this argument.

Furthermore, as Croft et al. (2008) argue, what matters is which grammatical form encodes the 'framing' or result event. Croft et al. (ibid.) rightfully point out that if we follow Talmy's definition of satellite strictly, it is neither the verb nor the satellite that frames but rather a preposition as in 'The bird flew over the roof'. In other words, *anything that is not a verb root but encodes the result component of motion (i.e. Path) will be referred to as a satellite*. This is in line with Slobin's (2008) recent redefinition of the typological dichotomy

by defining languages as PIV (path in verb) and PIN (path in non-verb), which matches Talmy's distinction of verb-framed vs. satellite-framed but avoids the problematic issue of how to delimit the category of satellites.

Numerous studies have indicated that Talmy's typological differences could be relevant to applied domains of linguistics such as translation (cf. Filipović 2006, 2008, Slobin 1996, 1997, 2000), language acquisition (cf. Filipović 2007a, 2008, Hohenstein 2001, Naigles et al. 1992, Slobin 1997, 2000, 2006, Vidaković 2006) or other related disciplines, such as psycholinguistics (cf. Gennari et al. 2002, Malt et al. 2003) and forensic linguistics (Filipović 2007b). Our focus at present is on how these typological findings could improve the understanding of what happens in the process of second language acquisition for English and Serbian learners of Serbian and English respectively.

3. Intratypological contrasts

3.1 English vs. Serbian

The two languages were classified as members of the same type within the typology. In other words, Serbian lexicalization pattern is comparable to that of English:

(2) Čovek je utrčao u zgradu.
Man AUX.3SG into-run:PST.PFV.3SG.M into building:ACC
'The man ran into the building.'

We wanted to test the possible predictions regarding acquisition of languages that have the same lexicalization pattern that stem from the typology and see if advantage of a typological perspective could be found in this context. An intratypological study had been carried out in order to inquire whether these theoretical assumptions regarding languages that are classified as the same type (satellite-framed) can be verified on the basis of extensive and attested corpus data. Literary translation and electronic corpus data have been used for this purpose (cf. Filipović 2002, 2007a). Although the similarities predicted by the typology are found between English and Serbian lexicalization patterns (e.g. substantial use of manner verbs), a number of differences also seem to emerge. In the Serbian translation of the English text the information about Manner is omitted on a number of occasions, very much like in Spanish on which Slobin's (1996, 1997) reports, or a less specific description is given compared to the one in the English original. For instance, in more than a half of the examples of verbalized motion events in our data that come from translation from English (40 out of 70), the information about the manner of motion in Serbian has been omitted, simplified (e.g. by using a less complex verb), or expressed in another way, usually by an adverbial (cf. Filipović 1999, 2002, 2003, 2006, 2007a). Another interesting point is that the most frequent manner verbs in Serbian are those prefixed by *deictic prefixes*, which results in the additional piece of information being habitually present in Serbian but absent from English (cf. example (3)). We illustrate the use of verbs in constructions in Serbian motion expressions below:

(3) Dotrčao je uz stepenice u devojčinu sobu.
To-the-speaker/scene-run:PST.PFV.3SG.M aux.3SG up stairs:ACC into girl's room.
'He ran up the stairs into the girl's room.'

(4) Prelazio je ulicu teturajući se.
Cross:PST.IPFV.3SG.M AUX.3SG street.ACC staggering REFL
'He was staggering across the street.'

(5) Izašle su iz kuće veselo poskakujući.
Exit:PST.PFV.3PL.F AUX.3PL out of house:GEN cheerfully skipping
'They skipped out of the house cheerfully.'

We refined the classification of motion events using the most salient spatial and temporal notions of *boundary* and *change* respectively and proposed a network of *situation types*, defined as *linguistically relevant features of events most likely to be lexicalized across languages*. This enabled us to propose a set of instructions that can be used as guidelines for lexicalization of motion events that could have a formative role in foreign language learning and teaching. They are based on the essential information on spatio-temporal change that constitutes each situation, which is illustrated in Table 1:

Table 1. Situation types in motion expressions

Spatial → Temporal ↓	**Boundary-crossing**	**Boundary-reaching**	**Non-boundary-crossing**
Change-occurred	a. He limped into the house.	c. He was running towards the house.	–
Moment-of-change	b. She was swaying into the house.	He was arriving at the house.*	–
No-change	–	–	d. They strolled/ were strolling in the park.

* This situation type is expressed by using verbs of arriving and leaving (cf. Levin 1993 for classes of verbs) and not expressed by motion verbs strictly speaking. They do not constitute the central part of motion expressions in our study. For further discussion see Filipović (2007a).

The * marking in Table 1 indicates that this situation type is possible, but it is not further considered here for the following reasons: (a) it is not a motion situation type strictly speaking, and it is not expressed by motion verbs, but rather with verbs of *arriving* and *leaving*, which are considered to be a separate category (cf. Levin 1993), and (b) it could be subsumed under other situation types if motion verbs are used (i.e. it becomes a non-boundary-crossing/no-change type). In boundary-reaching/moment-of-change expression 'I was arriving at Kings Cross station when my phone rang', the noun phrase that follows the preposition refers to *goal* rather than a boundary that is reached. Furthermore, if we try to have an example where the object of the preposition actually refers to a boundary, as in 'I was arriving at your door when my phone rang', we can see that the construction sounds rather unusual. Thus, this situation type is not really possible unless we have a very loose idea of what a boundary is (e.g. is "Kings Cross station" a boundary

that somebody reaches/arrives at?). If we try to use *a manner of motion verb* to express boundary-reaching/moment-of-change, we can see that the situation type referred to is actually that of non-boundary-crossing/no-change type (still directional, not locational): 'I was running to your door when my phone rang'. Therefore, the boundary-reaching/moment-of-change situation type does not play a prominent role in our situation type templates, but we acknowledge the possibility of this spatio-temporal combination in the context of arriving/leaving, which is related to, but not a part of the focus in our network of situation types in motion events per se.

While native speakers, translators and language learners will be instructed to use manner verbs in English in all circumstances since they are the most typical, natural means to express different kinds of motion in combination with directional prepositions, the guidance for Serbian can be summarized along the following lines:

i. use either a perfective manner (if licensed as in (6a) and (6d)) or a perfective directional verb (as in (6b)) in change-occurred situations,
ii. imperfective directional verb for moment-of-change situations as in (6c) and
iii. imperfective manner verb for no-change situations as in (6e); (see Filipović 2007a for further details).[5]

The following examples illustrate the expressions of different situation types in Serbian:

(6) a. He limped into the house.
Ušepao je u kuću.
Into-limp:PST.PFV.3SG.M AUX.3SG into house:ACC

b. He entered the house limping.
Ušao je u kuću šepajući.
Enter:PST.PFV.3SG.M AUX.3SG into house:ACC swaying

c. She was swaying into the house.
Ulazila je u kuću:ACC njišući se.
Enter:PST.IPFV.3SG.F AUX.3SG into house swaying REFL

d. He ran to the house.
Dotrčao je do kuće.
To-the-speaker/scene-run:PST.IPFV.3SG.M AUX.3SG to house:GEN

e. They strolled/were strolling in the park.
Šetali su u parku.
Stroll:PST.IPFV.3PL.M AUX.3PL in park:LOC

As illustrated in the examples (6a) to (6e), both prefixed and unprefixed manner verbs are used in motion expressions in Serbian. The unprefixed manner verbs combine freely with directional particles, just like in English. When it comes to the prefixed verbs, in Serbian, the direction is expressed both in the prefixes, which are morphologically bound to the verb root in the form of bound morphemes, and in the prepositions that follow the prefixed verb. The prefixes and the prepositions can, but need not be of the same phonetic appearance. Originally, the prefixes were prepositions (some of them adverbs), but they now form a compound with the verb root, adding their prepositional meaning to the basic meaning of the verb root (Grickat 1966/67, Stevanović 1989:434; Stanojčić et

al. 1989: 151).[6] There is a large number of these prefixes, but these are the ones that cover most of the relevant spatial relations (Filipović 2007a: 14–15):

DO- **do**trčati (run to smb./sth.), **do**šetati (walk to smb./sth.), etc.; Move to a speaker/scene/location in the manner of action described by the verb root.

OD- **od**juriti (dash from smb./sth.), **o**tšetati (stroll from smb./sth.), etc.; Move from a speaker/scene/location in the manner descibed by the verb root.

PRE- **pre**trčati (run across or over), **pre**skočiti (jump across or over), etc.; Move across or over in the manner described by the verb root.

U- **u**leteti (fly into), **u**skočiti (jump into), etc.; Move inwards in the manner described by the verb root.

IZ- **i**skočiti (jump out), **i**strčati (run out), etc.; Move outwards in the manner described by the verb root.

PRO- **pro**trčati (run through), **pro**teturati (stagger through), etc.; Move through a space in the manner described by the verb root.

POD- **pot**rčati (run under), **pod**leteti (fly under), etc.; Move towards a space under something in the manner described by the verb root.

NA- **na**trčati (run onto), **na**skočiti (jump onto), etc.; Move onto a space in the manner described by the verb root.

All these prefixes also have many other meanings. They can be employed to signify only a change of aspect (without adding any other component of meaning such as direction; this is called "pure perfectivization"; cf. Grickat 1966/67); they are also used to turn non-transitive into transitive verbs, and also to express a semantically completely different kind of action from the one expressed by the verb root. Here are some examples:

a. 'trčati' (an intransitive verb: 'to run') vs. '**pre**trčati' (a transitive verb: either 'to run across', or 'to run the whole length of sth.'). For example: Trčao je ceo dan. ('He ran all day'.) vs. **Pre**trčao je ulicu. (meaning either 'He ran across the street' or 'He ran the whole length of the street').
b. 'mutiti' (an imperfective verb; 'to mix' (e.g. ingredients)) vs. '**iz**mutiti' (a perfective verb; 'to finish mixing');
c. 'čitati' (an imperfective verb; 'to read') vs. '**pro**čitati' (a pefective verb; 'finish reading after having read the whole or specified parts (of a book, for example))
d. 'staviti' (a perfective verb; 'to put') vs. '**sa**staviti' (a perfective verb; 'to join' (things together)).

Serbian, like English, has verbs that conflate *Motion* and *Path*, i.e. directional verbs, such as 'ući' ('enter'), 'izaći' ('exit'), 'popeti se' ('ascend'), 'sići' ('descend'), 'proći' ('pass'), etc., which can be perfective (the ones just quoted) and imperfective ('ulaziti' – 'enter', 'izlaziti' – 'exit', 'penjati se' – 'ascend', 'silaziti' – 'descend', 'prolaziti' – 'pass', etc.), just like manner verbs.[7] However, this difference in directional verbs is not drawn by prefixation and the stems are felt to be monomorphemic. Their distribution in both perfective and imperfective forms, unlike that of manner verbs, is not restricted in any way and this is why they are the preferred option in some situation types in Serbian. Prefixed perfective manner verbs used in narration in Serbian are restricted in terms of *morphological blocking* and

combinatory potential, as originally outlined in Filipović 2007a. These two morphosyntactic processes restrict the freedom of verb-prefix combinations and the combinability of prefixed manner verbs with directional prepositions in motion expressions respectively. As a result, the use of manner verbs is much more restricted compared to that of manner verbs in English. In the case of (6a), it is possible to use a prefixed manner verb. However, the use of such verbs in this situation type is limited. For example, if we wanted to express a motion event in a different spatial configuration, as in 'The man limped across the street', we would not be able to use a manner verb because an adequately prefixed one is not licensed (*prešepati-'across-limp'). A perfective directional verb would be used instead as in: 'Prešao je ulicu šepajući' ('He crossed the street limping'; cf. also the pattern in (6b)).

Corpus data also showed that directional verbs are much more formal and less frequent in English, while they represent the statistically favourite means on the whole for lexicalizing motion events in Serbian (cf. Filipović 2007a, Vidaković 2006). For example, the verb 'trčati' ('to run') cannot be prefixed by a prefix indicating upward movement, needed in the example (3). A deictically prefixed verb has to be used instead. Similar constraint is noticed in (5), where an adequately prefixed manner verb is not licensed and a directional verb is used.

Another constraint operates in (4) and it is also illustrated in (6c). Prefixed manner verbs, which are also perfective and used to express boundary-crossing/change-occurred situation types, are blocked in expressions such as (4) because they cannot undergo further imperfectivization. For example, the verb 'ušepati' (into-limp.PFV) would have had the imperfective form *'ušepavati' (into-limp.IPFV) which is not licensed and therefore a directional verb is used as in (6c)). This process is termed *morphological blocking* (Filipović 2006, 2007a). Unprefixed manner verbs (e.g. 'šepati'(limp.IPFV)) cannot be used either in such situations because the meaning of an expression would no longer be that of moment-of-change, but rather of a no-change situation (cf. (6e)). The morphological complexity of manner verbs and the (more or less restricted) potential of different prefixes to be followed by different prepositions in a manner expression condition the use of manner verbs in Serbian (Filipović 2007a). For example, deictically prefixed manner verbs, unlike those prefixed otherwise, have higher frequencies in both dictionary and corpus data because of their *combinatory potential* and the ability of a single deictic verb to accumulate numerous prepositions. For example, the verb 'dotrčati' ('to-the-speaker-scene-run') can be used to express the whole Path as in: '**Do**trčao je preko brda, dolina, kroz šumu u selo.' ('To-the-speaker-scene-ran across the fields, valleys, through the forest into the village'). A verb prefixed by the prefix POD- ('under') however can only be followed by one preposition, namely 'pod' ('under'), as in '**Pod**leteo je pod krov.' ('Under-flew under the roof'). Combinatory potential is measured by the number of different prepositions that can follow the prefixed manner verb. The combinatory potential makes it possible to explain why certain prefixed manner verbs have much higher frequencies in both corpus and dictionary data (Filipović 2007a). All the prefixes in Serbian have their place on the prefix cline based on their combinatory potential (see Filipović 2007a for a full account). Consequently, the two deictic prefixes OD- ('from-the-speaker/scene') and DO- ('to-the-speaker/scene') are on the top of the cline because they can combine with all the prepositions and can also be followed by more than one preposition at a time, thus providing the most economical

lexicalization of complex motion events. The prefix POD- ('under') is at the bottom of the cline because verbs prefixed by POD- can only combine with its phonetically equivalent preposition ('pod' – 'under') and cannot be followed more than one preposition. The morphosyntactic restrictions reflected in the two principles, morphological blocking and combinatory potential, make it possible to account for the differences between native speakers and learners lexicalization data and make predictions about the relevant issues in the acquisition process for both English and Serbian L1 groups.

English and Serbian show similar behaviour when non-boundary-crossing is expressed and restricted similarity in boundary-crossing/change-occurred. In boundary-crossing/moment-of-change, the means of expression in the two languages differ conspicuously. We shall now focus on boundary-crossing/change-occurred situation type because this situation type has been the crucial one for establishing the typology (cf. Aske 1989, Slobin 1996, 1997). It is also the most frequent one in narratives because most of the reports about motion events contain emphasis on the end result of the observed event and information about a kind of change that has occurred within the perceived environment (cf. Filipović 2007a). We shall analyze the following combinations of expression in the two languages respectively: English: Manner verb + Path PP vs. Serbian: Manner verb + Path PP or Path verb + Path PP (+ Manner adjunct). The data discussed in the next section will shed light on how learners learn to express motion in an L2 and they will also confirm the following central claims: (i) The use of Manner verbs is restricted in Serbian because of prefixes; (ii) Path verbs are statistically favourite means for lexicalizing motion events in Serbian, unlike in English. (iii) There is a significant tendency to omit or simplify Manner in Serbian.

4. Learning to express Motion in an L2

In this section, some of the aspects of learning how to talk about motion in a second (foreign) language beyond the beginning stage will be discussed. The aims are to tease out the factors influencing L2 acquisition and to show how these factors interact. The present study is bidirectional, examining the acquisition of English by Serbian native speakers and the acquisition of Serbian by English native speakers. Bidirectionality allows one to pinpoint whether a tendency is a product of L1 influence (only) or whether the tendency occurs in all learner groups, irrespective of different L1–L2 pairings, and is, thus, a product of a (potentially) universal factor or interlanguage strategy. The origin of potentially universal learners' strategies is sought in general cognitive capacities and non-linguistic knowledge about the world.

4.1 Participants

The informants are two control groups of fifteen native speakers each, one of them consisting of native speakers of English (EC) tested in their native tongue and another one of native speakers of Serbian (SC) tested in their mother tongue; three groups of English

native speakers learning Serbian: lower intermediate (NELI), upper intermediate (NEUI) and advanced learners (NEADV), fifteen per level, and three groups of Serbian native speakers learning English: lower intermediates (NSLI), upper intermediates (NSUI) and advanced learners (NSADV), fifteen per level. All informants are adults. At the testing time, the NELI had been learning Serbian for four years, the NEUI for nine years and the NEADV for 21 years on average. The period of formal tuition that these three groups had undergone was similar in length ranging between one year and one year and a half on average. The rest of the overall learning period had been spent mostly in learning the language through informal exposure, or through on and off self-tuition. The NSLI had been learning English for eight years, the NSUI for ten years and the NSADV for sixteen years on average before the test took place.

Unlike English learners of Serbian, Serbian learners of English had all been learning English only through formal tuition. All English learners of Serbian have spent some time in a country where Serbian or Croatian is an official language (Serbia/Montenegro, Croatia, Bosnia): the NELI two years, the NEUI five years and the NEADV sixteen years on average. On the other hand, Serbian learners of English spent considerably less time in English-speaking countries: none of the lower intermediate learners of English had ever lived outside Serbia and Montenegro or communicated in English outside the classroom; four NSUI had lived in the UK for one or two months; the NSADV had lived in the UK or Canada for four months on average (Vidaković 2006). The tuition differences between English and Serbian learners may have influenced their performance on the present tasks; their possible influence will be noted in the data analysis. However, linguistic rather than tuition style factors are the most likely to dominantly influence their L2 acquisition since plenty of English and Serbian characteristics in the domain of expression of motion are not taught at school nor mentioned in course/grammar-books.

4.2 Stimuli

Forty-six black-and-white drawings, each representing a motion event, were used as elicitation material. Each of the forty-six items was designed to elicit both Manner (*to walk, to run, to jump, to dance, to limp, to crawl, to cycle, to skateboard*) and Path of motion (*into (a house), out of (a house), across (a path, a street), up and down (the stairs)*) in boundary-crossing/change-occurred situations. Every manner of motion (*walk* – canonical Manner for humans and *run, dance*, etc. – noncanonical) was represented five times in combination with the above-mentioned five paths of motion. In addition, there are six more items where non-everyday Manner of motion (e.g. climbing, shinning) was combined with perpendicular Paths: these items represent a Figure climbing (shinning) up and down a flagpole, climbing up and down a steep cliff and climbing up and down a tree. The research subjects were interviewed individually. After the presentation of an item, subjects were asked to tell to the interviewer 'What happened?' for each item and were told they could be as explicit as they wanted. The equivalent question 'Šta se desilo? – What happened?' was used for Serbian controls and English learners of Serbian.

4.3 Hypotheses

In the spirit of the Interlanguage approach (Corder 1967, 1971, 1981, Selinker 1972) we assumed that learners' interlanguages and L2 acquisition are systematic, with certain characteristics not directly relatable to either L1 or L2 characteristics. We set out hypothesising that language-specific influences will be stronger and acquisition paths not so homogenous among the learners *beyond the beginning stage* than they are among beginners, but we also assumed that factors other than first or second language, like general problem-solving abilities or world knowledge, will bring out similarities in the interlanguages and acquisition paths of learners with different first and second languages (Vidaković 2006).[8] The latter was assumed in this paper even though the findings of the available L2 acquisition studies in the domain of spatial relations detect only L1 and/or L2 influences on learners beyond the beginning stage (Cadierno 2004, Hendriks 2005, Inagaki 2001, Schlyter 1984).

The relevant questions for our study are the following: what types of utterances will be used by learners (and controls), why, and which ones are the most frequent and why? In other words, where in utterance do learners express Manner and Path and do they sometimes choose not to express one of them at all? Before listing the hypotheses below, we will briefly summarize the possible types of utterances which are likely to be used by learners to refer to motion involving boundary-crossing:

a. utterances with little information:
   – Path only: *She went across (the road)*.
   – Manner only: *She ran*.
b. utterances with more information (Manner and Path): *She ran across (the road)*.
c. utterances with the same amount of information as (b) (i.e. with information on both Manner and Path), but with a higher structural complexity than (a) or (b) since they contain an additional (Manner) expression: *She went across (or: she crossed) the street running/quickly/in a jump.*

The research hypotheses on the distribution and supply of Manner and Path information are provided below.

4.3.1 *Hypothesis 1 (L1 and L2 influence): Distribution of Manner and Path information in an L2 utterance*

Considering that English and Serbian differ typologically in certain respects (cf. Section 3), a certain amount of language-specific influence is expected in L2 acquisition process. Owing to L1 influence, English learners of Serbian may resort to Manner verbs more frequently than to Path verbs and express Path in the prepositional phrase only ('Manner verb + Path PP'). If the constructions with Manner verbs do not outnumber those with Path verbs, English L1 influence may still be evident if English learners of Serbian use constructions with Manner verbs more frequently than Serbian controls. Serbian learners of English, on the other hand, may use constructions with Path verbs ('Path verb + Path PP + (Manner adjunct') more often than those with Manner verbs, or they may use

constructions with Path verbs more often than English controls do, due to Serbian L1 influence. L2 influence may increase with the increasing proficiency levels of English and Serbian learners.

4.3.2 *Hypothesis 2 (Economy-of-form strategy): Distribution of Manner information in an L2 utterance*

When expressing Manner, whatever the preferred pattern in an L1 or L2 is, 'Manner verb + Path PP' (satellite-framed) or 'Path verb + Path PP + Manner adjunct' (verb-framed), both English and Serbian learners may tend to express Manner in the verb and Path outside it.[9] In this way, they will be more economical by opting for the construction in which all the necessary information, Manner and Path, is packaged only in two elements: a verb and a preposition respectively. In contrast, expressing the same amount of information in the construction 'Path verb + Path PP + Manner adjunct' requires the use of an additional expression, Manner adjunct, beside a verb and a PP. Since both constructions are equally semantically transparent, and convey the same amount of information, opting for the one whose form is more economical (Manner verb + Path PP) will facilitate the processing (production/comprehension) load. We expect that learners will predominantly use this, as we term it, *economy-of-form strategy*, on all proficiency levels, *when Manner is expressed* (Vidaković 2006).[10] Naturally, the more Manner verbs the learners know the more instances of economy-of-form strategy in the learners' data. Economy-of-form strategy does not necessarily imply that Manner verbs will be the most numerous of all verbs, since learners may also choose not to express any Manner at all.

Satellite-framed languages will also be referred to as S-languages, while verb-framed languages will be referred to as V-languages. The contrasting patterns defined as 'Manner verb + Path PP' and 'Path verb + Path PP + Manner adjunct' are actually an elaboration of Talmy's pattern outline, namely *satellite-framed*: Manner verb + Path satellite vs. *verb-framed*: Path verb. We contend that this elaboration is necessary in order to highlight the whole complexity of structures that actually occur in lexicalization as well as individuate all the relevant elements such as prepositions and adjuncts. It is not enough just to say, for example, that the satellite-frame consists of Manner verb + Path satellite because in this way we disregard what goes on elsewhere in the structure (e.g. the prepositional phrase (PP)), which is essential for the understanding of the points where learner difficulty comes from. For instance, the case of the noun following the preposition may be crucial in the distinction between expressions of boundary-crossing and non-boundary-crossing, whereby different verbs or verb forms are required (cf. examples in (6a) vs. (6e) and the example (14)). The need to move from a verbo-centric approach to the issues presented in this paper is thus made evident.

4.3.3 *Hypothesis 3: Path in prefixes*

We expect that the acquisition of Path prefixes may pose a significant problem for English learners of Serbian, not so much because of L1–L2 differences (there being no Path prefixes for motion verbs in English), but rather because of the functional load of Serbian Path prefixes. Serbian Path prefixes have the twofold function of: (a) encoding Path and (b) marking perfective grammatical aspect, thereby signalling that the motion event is

finished, i.e. the boundary has been crossed (cf. example (2) above). Moreover, English learners of Serbian may find the double marking of Path in Serbian (in the prefix and a preposition) redundant and thus choose not to encode it twice. Thus, we hypothesize that English learners of Serbian will rarely prefix Manner verbs in L2 Serbian, which will, in those cases, result in an incorrect rendering of a motion event where boundary-crossing/change-occurred information will not be conveyed.

4.3.4 *Hypothesis 4: World knowledge about Motion*

According to Talmy (1985), *Path* is, universally, the most basic information in a motion situation (i.e. the *core component* of a motion event), while Manner is a subsidiary one. Given that assumption, learners may be more likely to omit Manner information and less likely to omit Path information, even when they lack the adequate means to talk about Path. We expect this to happen irrespective of L1 or L2 and due to the general world knowledge about motion.

4.4 Data analysis

Learning how to talk about motion in a second language is examined in this section in the light of the previously discussed typological framework. It will be investigated how English learners of Serbian, Serbian learners of English and English and Serbian controls distribute Manner and Path information at an utterance level. The table below shows how the information on Manner and Path is distributed in the utterances of learners and controls.

Table 2. Distribution of Manner and Path information in the utterances of learners and controls

	VPOMP*	VPOP	VPOM	VP	VMOPM	VMOP	VNOP	VM
NELI	34 (5%)**	64 (9%)	10 (1%)	24 (3%)	70 (10%)	242 (35%)	219 (32%)	27 (4%)
NEUI	74 (11%)	128 (19%)	29 (4%)	37 (5%)	46 (7%)	228 (33%)	116 (17%)	32 (5%)
NEADV	102 (15%)	200 (29%)	16 (2%)	33 (5%)	58 (8%)	243 (35%)	34 (5%)	4 (1%)
SC	217 (31%)	274 (40%)	57 (8%)	39 (6%)	13 (2%)	87 (13%)	2 (0%)	1 (0%)
NSLI	134 (19%)	239 (35%)	48 (7%)	46 (7%)	40 (6%)	150 (22%)	2 (0%)	31 (4%)
NSUI	105 (15%)	133 (19%)	49 (7%)	27 (4%)	65 (9%)	283 (41%)	9 (1%)	19 (3%)
NSADV	75 (11%)	100 (14%)	29 (4%)	29 (4%)	66 (10%)	388 (56%)	2 (0%)	1 (0%)
EC	9 (1%)	14 (2%)	8 (1%)	18 (3%)	24 (3%)	614 (89%)	2 (0%)	1 (0%)

* **VPOMP** – Path is expressed in the verb and Manner and Path are expressed outside the verb ('Ušla je u kuću trčeći.' – 'She entered into the house running'); **VPOP** – Path is both in the verb and outside it ('Sišao je niz stub.' – 'He descended down the pole'); **VPOM** – Path is in the verb and Manner is outside it ('They entered the room dancing.'); **VP** – Path is expressed in the verb ('He entered.'); **VMOPM** – Manner is expressed in the verb and outside it, while Path is expressed outside the verb only ('Beba je hodala sa rukama i nogama kao pas iz kuće.' – 'The baby walked with her hands and legs like a dog out of the house.'); **VMOP** – Manner in the verb, Path outside it ('They danced into the house.'); **VNOP** – the verb expresses motion only (e.g. move), while Path is expressed outside it ('She moved up the stairs.'), **VM** – Manner is expressed in the verb ('She ran.').

** The percentages for each category in each subject group were calculated out of the total of 690 utterances.

The above data show that *English and Serbian controls* differ considerably in the way they express Manner and Path information in an utterance. English controls behave like typical S-language speakers predominantly using a 'Manner verb + Path PP construction' (89%, cf. VMOP) and rarely omitting Manner information (5%, cf. VPOP, VP, VNOP). Serbian controls, on the other hand, mostly encode Path both in the verb (like V-language speakers) and outside it, in a prepositional phrase (like S-language speakers), (40%, cf. VPOP), expressing Manner optionally outside the verb (like V-language speakers), (31%, cf. VPOMP). Besides, Serbian controls omit Manner information much more frequently than English controls do, which is a V-language tendency (46%, cf. VPOP, VP, VNOP). While English shows all the characteristics of an S-language, Serbian exhibits characteristics of both S- and V-languages, which is a further support to Filipović's (2002, 2007a) claim that Serbian is midway on the continuum between the two language types.

4.4.1 *Distribution of Path information in an L2 utterance*

English and Serbian learner groups differ considerably from one another and many of these differences arise from the interplay of L1 and L2 influences. When they express Path, lower intermediate *English learners of Serbian* mostly rely on prepositions (77%) rather than both verbs and prepositions (14%) for Path encoding.[11, 12] This is due to the influence of L1 English while also being partly due to the economy-of-form factor (in Manner encoding) which will be explained further below. When 'disentangled' from economy-of-form factor, English L1 influence is obvious in the following: (a) English lower intermediate learners of Serbian rely only on prepositions for expressing Path much more often than the same level Serbian learners of English, who do it in 28% of the cases (cf. VMOPM, VMOP, VNOP), (b) English lower intermediate learners of Serbian prefer expressing motion in the verb and Path in the preposition (32%, VNOP), as in example (7) below, to encoding Path both in the verb and the preposition (14%, VPOMP, VPOP). The instances of expressing only motion in the verb and Path outside it are negligible among the Serbian and English controls, while being considerable only among lower and upper intermediate English learners of Serbian. Although such a tendency is very rare among English controls, it is, nevertheless, explainable by English L1 influence: in our opinion, it arises from the fact that the slot 'reserved' for Path encoding is outside rather than inside the verb in L1 English. With the increasing proficiency levels, the influence of L2 Serbian increases although the influence of L1 English is still evident. For example, English upper intermediate learners of Serbian still rely on prepositions only for Path encoding (57%) more often than both on verbs and prepositions (30%).[13] Advanced English learners of Serbian rely equally frequently on prepositions only and on both verbs and prepositions for expressing Path, which indicates a lingering but decreasing L1 influence, and an increasing L2 influence.

(7) Momak je *išao gore *kod zastave.
Guy AUX.3SG go:PST.**IPFV**.3SG.M up **at** flag:GEN
'A guy was going up to the flag.'

The interplay of language-specific influences is also observable in the way *Serbian learners of English* express Path. The construction 'Path verb + Path PP' (35%, cf. VPOP) is preferred to all other constructions by lower intermediate Serbian learners of English (cf. example (8) below), who mostly rely both on the verb and the preposition for Path encoding (54%,

cf. VPOMP, VPOP) rather than on prepositions only (30%, cf. VMOPM, VMOP, VNOP). In these respects, they are clearly influenced by L1 Serbian. Serbian learners of English at higher proficiency levels (upper intermediate and advanced) rely on prepositions only for expressing Path (NSUI 50%, NSADV 66%) much more often than both on verbs and prepositions (NSUI 34%, NSADV 25%), thus following English L2 pattern (cf. example (9)). This difference in tendencies between the lower intermediate and higher proficiency levels may have been partly motivated by English L2 influence on Path lexicalization; for another part, we suggest that it has been motivated by economy-of-form factor in Manner encoding.[14] The more Manner verbs in a learner's vocabulary and the more frequently Manner is mentioned, the more instances of economy-of form strategy and hence fewer Path verbs (Vidaković 2006).

(8) The man came up the stairs.
(9) The girl climbed down a mountain.

The above findings have shown that there is the interplay of both L1 and L2 influences across all proficiency levels. The L1 influence is the strongest among lower intermediate learners (e.g. English learners of Serbian at that proficiency level encode Path in prepositions mainly, while Serbian lower intermediate learners of English encode it frequently both in a verb and a preposition), while it decreases with the increasing proficiency levels due to a stronger L2 influence. We have also suggested that economy-of-form factor plays a part, which will be elaborated on below.

4.4.2 *Distribution of Manner information in an L2 utterance*

Examination of the distribution of Manner information in an utterance will reveal a certain amount of L1 and L2 influence too, but it will also show that English and Serbian learners exhibit a common tendency when expressing Manner in their respective L2s.

L1 influence is the strongest at the lower intermediate proficiency level and it decreases with increasing proficiency levels. For example, owing to English L1 influence, English lower intermediate learners of Serbian use the typically 'English' construction 'Manner verb + Path PP' (35%, cf. VMOP and example (10) below) more frequently than either Serbian controls (13%, cf. VMOP) or Serbian lower intermediate learners of English (22%, cf. VMOP). They also express Manner outside the verb (in a Manner adjunct), more often than English controls do (6% vs. 2%), due to Serbian L2 influence which increases among upper intermediate and advanced English learners of Serbian (NSUI: 15%, NSADV: 17%).[15]

On the other hand, Serbian lower intermediate learners of English use the construction 'Manner verb + Path PP' (VMOP) less often (22%), while using constructions with Manner adjuncts more often (26%, cf. VPOMP, VPOM and example (11) below), than the same level English learners of Serbian do, due to Serbian L1 influence (cf. Table 2 above). With the increasing proficiency levels, Serbian L1 influence decreases.

(10) Devojčica skoči *dole stepenice.
A girl jump:PST.PFV.3SG **down(wards)** stairs:NOM
'A girl … jumped down the stairs.'
(11) He crossed the street running.

Beside the above mentioned language-specific influences arising from typological differences between English and Serbian, English and Serbian learners share some *common traits*. All learners (apart from Serbian lower intermediate learners of English) prefer expressing Manner in the verb (VMOP) rather than outside it (VPOMP or VPOM), as shown in Table 2 above. In that respect, it is evident that Serbian learners of English get closer to English controls than English learners of Serbian do to Serbian controls.

The English pattern (VMOP) is already very frequent (41%) among the upper intermediate Serbian learners of English and its frequency is considerably higher than the frequency of any other construction. Advanced Serbian learners of English develop a clear preference for the English way of distributing information across the utterance: the percentage of the construction 'Manner verb + Path PP' (VMOP: 56%) surpasses the percentage of any other construction (cf. example (12) below).

(12) They danced into the house.

As far as English learners of Serbian are concerned, it is evident that 'Manner verb + Path PP' (VMOP) is the most frequent construction across the proficiency levels examined and its frequency remains around 35% from the lower intermediate to the advanced proficiency level (cf. example (13) below). The constructions in which Manner is expressed outside the verb (VPOMP, VPOM) are much less frequent. Their frequency increases from only 6% among English lower intermediate learners of Serbian to only 17% among the advanced learners.

(13) On je *hodao uz stepenice.
He AUX.3SG walk:PST.**IPFV**.3SG.M up stairs:ACC
'He was walking up the stairs.'

We suggest that holding on to the English L1 pattern, not only by the lower intermediate but also by the advanced English learners of Serbian, and a prominent departure from the L1 Serbian pattern in favour of the L2 English pattern by upper intermediate and advanced Serbian learners of English are, for a big part, the results of resorting to the (potentially) *universal, economy-of-form strategy*, *in cases where learners express Manner* (cf. Hypothesis 2). We are not inclined to attribute these differences in attainment to different types of acquisition because lexicalization of motion events in boundary-crossing situations is not a topic taught at school nor is it analyzed in textbooks and grammar books.

The learners' preference for expressing Manner inside the verb becomes even stronger when utterances referring to cycling and skateboarding drawings are excluded. In Serbian, unlike English, Instrument of motion can only be expressed outside the verb, in an adjunct (e.g. 'na biciklu/biciklom' – 'on the bike'). In order to see what subjects do when they *can* chose between expressing Manner in the verb or outside it, we excluded the utterances describing items with cycling and skateboarding motion events (cf. Table 3).

The situation among the *controls* is the following: a one-way ANOVA shows that the frequency of 'Manner verbs only' is significantly higher than that of constructions with Manner adjuncts ($F(1, 14) = 440.040$, $p = .000$) among English controls, whereas the Serbian controls significantly prefer Manner adjuncts to 'Manner verbs only' ($F(1, 14) = 4.693$, $p = .048$).

Table 3. Economy of form: Manner verb only constructions vs. constructions with Manner adjuncts (without utterances referring to 'cycling' and 'skateboarding' drawings)

	Constructions with Manner verb only		Constructions with Manner Adjunct(s)	
NELI	238	82%	54	18%
NEUI	236	74%	82	26%
NEADV	232	81%	55	19%
SC	85	35%	158	65%
NSLI	162	63%	95	37%
NSUI	282	74%	99	26%
NSADV	343	83%	69	17%
EC	487	96%	22	4%

The (English and Serbian) *learners'* preference patterns are now clear cut and considerably different than before. The results show that when English learners can choose between encoding Manner in the 'verb only' and encoding it in a construction with a Manner adjunct, they opt for a verb. This preference is strongly significant at all proficiency levels (NELI: F (1, 14) = 53.154, p = .000; NEUI: F (1, 14) 35.736, p = .000; NEADV: F (1, 14) = 74.426, p = .000). The same holds for Serbian learners of English (NSLI: F (1, 14) = 4.732, p = .047; NSUI: F (1, 14) 39.696, p = .000; NSADV: F (1, 14) = 37.860, p = .000).

In sum, when English and Serbian learners express Manner, lower intermediate, upper intermediate and advanced learners construct their interlanguages similarly by choosing to encode Manner in the verb significantly more often than in the adjunct, in spite of language-specific differences and due to the economy-of-form factor. English learners of Serbian may initially be partly influenced by L1 English, but it is the economy-of-form factor that encourages them to use 'Manner verb + Path PP' construction with a similar frequency across proficiency levels thereby hindering their progress towards the L2 pattern of expressing Manner outside the verb. Due to the same, economy-of-form, factor Serbian learners of English home in to the English L2 pattern early exhibiting a considerable progress towards the L2 pattern with increasing proficiency levels.

4.4.3 *L2 acquisition of Serbian Path prefixes*

We hypothesised earlier that English learners of Serbian will find it difficult to acquire Serbian Path prefixes (cf. Hypothesis 3). The data analysis confirmed this, both with respect to the frequency with which English learners of Serbian used prefixes and the information they expressed in those prefixes.

In the Serbian control data, all the Manner verbs are prefixed, apart from some instances of the perfective verb *skočiti (to jump)*. As far as learners' data are concerned, even though the English learners of Serbian use around three times as many Manner verbs (cf. Table 4 below), prefixed (and hence perfective) Manner verbs are extremely infrequent. Therefore, it does not come as a surprise that *the number of prefixed Manner verbs is significantly lower than the number of unprefixed ones* in the learners' data across all proficiency levels (NELI: F(1, 14) = 2419.200, p = .000, NEUI: F(1, 14) = 149.775, p = .000, NEADV: F (1, 14) = 35.400, p = .000). This is the case even though the elicitation question 'Šta se desilo?' (What happened?) prompts the use of perfective verbs. This means that when they

use Manner verbs, English learners of Serbian rarely encode boundary-crossing information in the verb, but they usually do so in a preposition, thus following their L1 pattern. In doing so, they fail to convey boundary-crossing/change-occurred information, as in the example below:

(14) Ona je *trčala u kuću.
She AUX.3SG run:PST.**IPFV**.3SG.F in(to) house:ACC
'She was running towards the house.'

Table 4. Prefixed and non-prefixed manner verbs

	NELI	NEUI	NEADV	SC
Prefixed MV	9 (3%)	25 (8%)	59 (19%)	99 (97%)
Non-Prefixed MV	330 (97%)	281 (92%)	246 (81%)	3 (3%)

When the prefixes in our data are examined for the information encoded in them, the following trends become apparent. The Serbian controls almost always (in 99% of the cases) encode Path in prefixes to Manner verbs, as evident from the Table below. The prefixes encoding target Path (INTO, OUT OF, ACROSS, UP and DOWN) are the most frequent.[16]

Table 5. Path prefixes and purely aspectual prefixes

	NELI	NEUI	NEADV	SC
Target-Path Prefix	3 (33%)	4 (16%)	36 (61%)	71 (72%)
Other Path Prefixes: OD- (away from)/DO- (to)	0 (0%)	2 (8%)	1 (2%)	27 (27%)
Purely Aspectual Prefix	6 (67%)	19 (76%)	22 (37%)	1 (1%)

In the English lower and upper intermediate learner data, the amount of Path prefixes is almost negligible, as well as their variety. It is only at the advanced level that Path, or rather, target-Path prefixes, outnumber purely aspectual prefixes. For example:

(15) Oni su isplesali iz kuće.
They AUX.3SG out-dance:PST.PFV.3PL.M out.of house:ACC
'They danced out of the house.'

Therefore, the information that the majority of prefixes in the learners' data encode is only aspectual information: purely aspectual prefixes only perfectivize Manner verbs. Using purely aspectual prefixes in boundary-crossing situations is incorrect (cf. example (16) below), since only Path prefixes (target-Path or OD-/DO- prefixes) can be employed:

(16) Oni su *za**plesali** u kuću.
They AUX.3SG **start**-dance:PST.PFV.3PL.M in(to) house:ACC

It may be the case that since the learners very rarely perfectivize Manner verbs in L2 Serbian, they have not (fully) realized the following: although the simple past form of the verb in English refers primarily to completed (perfective) action in boundary-crossing situations, it is not so in Serbian where such a form is imperfective and has to be perfectivized in order to refer to a completed action. However, differences in what past tense forms refer

to in L1 English and L2 Serbian do not explain it all. Since on many occasions the learners opted for purely aspectual prefixes rather than Path prefixes, we suggest that they find Path prefixes more difficult to learn due to their *dual function* of conveying both spatial and aspectual information. Moreover, it is likely that learners are not prone to encoding the same Path twice, both in prefix and preposition, as they may find it redundant and perhaps because they typically encode it only in one sentential element in their mother tongue. Since this element is a preposition in English, they opt mostly for prepositions in Serbian as well. We also suggest that another cause of the rare occurrence of Path prefixes in the learners' data is the *low frequency of Path prefixes in the L2 input*. Their frequency in the input is related to the typological character of Serbian as a language in between satellite- and verb-framed languages: both the written corpus data (Filipović 2002) and the oral data (Vidaković 2006) have shown that (prefixed) Manner verbs are less frequently used than Path verbs. Finally, tuition style may have played the role in the delayed acquisition of prefixes: if English learners of Serbian had been tutored instead of being mostly semi-tutored, the frequency of Path prefixes may have been higher.

4.4.4 *Supply of Path and Manner information*

So far, it has been discussed where in utterance Manner and Path are expressed. The question to be addressed now is whether Manner and/or Path are expressed at all. We hypothesized that learners may tend to supply information on Path and more readily omit information on Manner because Path is, according to Talmy (1985), a core component of a motion event, while Manner is only an external one.

Our findings have revealed that learners are rather meticulous when Path information is concerned. All learners, as well as controls, rarely omit information on Path. The percentage of Path omission never surpasses 5% (cf. Table 6). English and Serbian intermediate learners omit information on target Path more often than controls (who do it only rarely) due to limited interlanguage vocabularies (Vidaković 2006). However, despite limited vocabularies, (English/Serbian) learners still try and manage to express target Path most of the time. They all use similar interlanguage strategies, such as assigning new meanings to existing prepositions and inventing new prepositional combinations, in their attempts to render full, complete (target) Path when they lack an adequate L2 preposition, (ibid.): e.g. 'dole na drvo' – lit. 'down on the tree' instead of 'niz drvo' – 'down the tree'; 'kroz ulicu' –'through the street' instead of 'preko ulice' – 'across the street'.

Table 6. Omission of Path and Manner Information

	Path Omission (VM)	**Manner Omission (VPOP, VP, VNOP)**
NELI	27 (4%)	307 (44%)
NEUI	32 (5%)	281 (41%)
NEADV	4 (1%)	267 (39%)
SC	1 (0%)	315 (46%)
NSLI	31 (4%)	287 (42%)
NSUI	19 (3%)	169 (24%)
NSADV	1 (0%)	131 (18%)
EC	1 (0%)	34 (5%)

Unlike Path information, information on Manner is more readily omitted by far (cf. Table 6). When learners lack an adequate L2 Manner verb, they sometimes use a simpler Manner verb, a Manner adverb or a Manner prepositional phrase, but they often, especially at the lower intermediate level do not express Manner at all. Our data reveal that many cases of Manner omission were accompanied by the comments of the type: 'I do not know/have forgotten how to say to "limp"'. Such comments were the most frequent for English semantically more complex verbs like crawl, limp, skateboard, cycle and their Serbian counterparts.

The omission of Manner does not only take place when learners lack certain Manner verbs, but is also influenced by the L1/L2 tendencies of Manner omission. English and Serbian learners are similar at the lower intermediate level where they omit Manner with a similar frequency. They differ at the upper intermediate and advanced proficiency levels due to L2 influence: Serbian learners of English express Manner more frequently due to English L2 influence, while the lower frequency of Manner supply among English learners of Serbian is due to Serbian L2 influence (Vidaković 2006).

Reluctance to omit Path information and a high rate of Manner omission provide support for our hypothesis that Path is a core component of a motion event, unlike Manner, and therefore is not likely to be omitted.

5. Conclusions

The detailed scrutiny of the typological classification by Talmy resulted in a finely-grained network of situation types that help us identify accurately the points in the lexicalization process where two (or more) languages differ as well as those features that they share. We provided the reasons why the use of verbs and presence or absence of information about Manner in Serbian may differ from the English pattern even though the two languages are said to belong to the same typological group. Our experimental findings show that the interlanguages of the examined learner groups are all systems *different from either L1 or L2*, but they do exhibit *language-specific influences*, which mould those systems in varying degrees. L1 influence is normally the strongest at the lower intermediate proficiency level, while L2 influence increases with the increasing levels of proficiency. However, we have demonstrated that focusing only on L1 or L2 influence while examining this reorganization of the learner interlanguage beyond the basic variety stage is not explanatory enough. These influences are only a part of a large *bundle of factors* influencing acquisition, plenty of which are *not directly relatable to either the first or the second language*. In order to ease the processing (production/comprehension) load while communicating in their L2, *both* English and Serbian learners frequently resort to economy-of-form strategy. Owing to the dual function of Path prefixes (carrying spatial and aspectual information), low frequency input from L2 (i.e. Serbian) and the semi-guided tuition style, English learners of Serbian use Path prefixes rarely even at the advanced proficiency level. Owing to general world knowledge about motion events, all learners rarely omit Path information, as core information on a motion event, while Manner omission is more frequent.

The analyzed tendencies in English, Serbian and in L2 acquisition data also have implications for teaching. It was shown that Serbian Path prefixes are problematic for English learners partly due to their functional load. Therefore, teaching prefixes should be carried out by explicating both functions that Path prefixes have (aspect marking and Path encoding) and comparing them to prefixes that are used only as aspect markers. Even though more than one option is available in English and Serbian to refer to Manner and Path of motion, these options are not equally frequent and colloquial in English and Serbian, as opposed to 'infrequent, literary and stilted' (Talmy 1985). In order to facilitate learners' adoption of the L2 pattern, teaching should be concerned not only with what is grammatical or ungrammatical, but also with what is generally more natural in a language being learnt.

Notes

1. We use the term Serbian here, but we used sources of data that come from sources and native speakers of both Serbia and Croatia. Serbo-Croatian is justified linguistically but political reality nowadays is reflected in the fact that Serbo-Croatian is going out of use as a term. Nevertheless, our arguments apply to both the Serbian and the Croatian variant.

2. In the case of the expressions of motion, the core information is that of Path (since motion events are defined as those where a Figure changes location), and it is expressed within the verb itself in verb-framed languages, and via satellites in satellite-framed languages.

3. Manner verbs are licensed in Spanish where a motion expression refers to motion at a location (e.g. the equivalent of 'He ran in the park') not directional motion ('He ran into the park'). This point was discussed extensively in the literature in terms of telicity and boundary-crossing (cf. Aske 1989, Slobin 1996, 1997), Filipović (2003, 2006, 2007a)). Nevertheless, Slobin argues that even when possible to insert them, manner verbs in Spanish are noticeably rarely opted for and the preference for the pattern with directional verbs is pervasive (ibid.).

4. This is very reminiscent of what Bolinger (1971) termed *adpreps*, which are particles that can function both as adverbs and prepositions.

5. Perfective manner verbs are formed by adding a prefix (e.g. 'U-' meaning 'into') to an imperfective verb form ('trčati'), as in 'U' + 'trčati' = 'utrčati' ('run into'). Perfective directional verbs do not undergo the same derivational process (cf. Filipović 2007a) and therefore are not subject to the same restrictions to their use in motion expressions.

6. The reverse happened in English. Namely, directional particles were prefixally bound to the verb in Old English but later became free morphemes that follow the verb.

7. The fundamental difference is the fact that directional verbs are much more formal and less frequent in English, while they represent the statistically favourite means on the whole for lexicalizing motion events in Serbian (cf. Filipović (2007a) and also Vidaković (2006) for some experimental evidence from spoken Serbian). In other words, the pattern in the example (6b) would be preferred even if the manner verb is licensed as in (6a).

8. L2 acquisition of motion and static spatial relations by beginners was analysed and discussed in Becker & Carroll (1997).

9. Cf. examples under (b) and (c) above

10. This term was first used in Vidaković (2006).

11. For reliance on prepositions, cf. Table 2 and the figures for: VMOPM+VMOP+VNOP.

12. For reliance on both verbs and prepositions, cf. Table 2 and the figures for: VPOMP+VPOP.

13. Cf. footnotes 11 and 12.

14. To be discussed further below.

15. Cf. VPOMP, VPOM

16. Target Paths are those that the drawings were designed to elicit.

References

Aske, J. (1989). Path predicates in English and Spanish: A closer look. Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society. Berkeley: *Berkeley Linguistics Society, 1–14*.

Becker, A. & Carroll, M. (1997). *The acquisition of spatial relations in a second language.* Amsterdam/Philadelphia: John Benjamins.

Bolinger, D. (1971). *The phrasal verb of English.* Cambridge, MA: Harvard University Press.

Cadierno, T. (2004). Expressing motion events in a second language: A cognitive typological perspective. In M. Achard & S. Niemeier (Eds.), *Cognitive linguistics, second language acquisition and foreign language teaching* (13–49). Berlin: Mouton de Gruyter.

Corder, S. P. (1967). The significance of learner's errors. *IRAL Vol. V/4,* 161–170.

Corder, S. P. (1971). Idiosyncratic dialects and error analysis. *IRAL Vol. IX/2,* 147–160.

Corder, S. P. (1981). *Error analysis and interlanguage.* Oxford: OUP.

Croft, W. et al. (2008). *Revising Talmy's typological classification of complex events.* Draft. Available at: http://www.unm.edu/~wcroft/Papers/TalmyTypology-paper.pdf.

Filipović, L. (1999). *Language-specific expression of motion and its use in narrative texts.* MPhil Dissertation. University of Cambridge.

Filipović, L. (2001). Language-specific use of verbs in motion expressions. In A.Veiga et al. (Eds.), *El Verbo entre el Lléxico y la Gramática* (53–62). Lugo: Tris-Tram.

Filipović, L. (2002). *Verbs in motion expressions: Structural perspectives.* PhD Dissertation. University of Cambridge, Department of Linguistics.

Filipović, L. (2003). Moving in mysterious ways: On the nature of lexical patterning. In D. Klikovac and K. Rasulić (Eds.), *Jezik, društvo, saznanje* (103–117). Beograd: Filološki fakultet.

Filipović, L. (2006). Weaving the web of meaning. *Languages in Contrast 6/1,* 151–175.

Filipović, L. (2007a). *Talking about Motion: A crosslinguistic investigation of lexicalization patterns*. Amsterdam: John Benjamins.

Filipović, L. (2007b). Language as a witness: Insights from cognitive linguistics. *Speech, Language and the Law 14* (2), 245–267.

Filipović, L. (2008). Typology in action: Applying insights from typological contrasts. *International Journal of Applied Linguistics 18* (1), 42–61.

Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition 83,* 49–79.

Grickat, I. (1966/67). Prefiksacija kao sredstvo gramatičke (čiste) prefiksacije. *Južnoslovenski Filolog XXVII/1–2,* 185–223.

Hendriks, H. (2005). Structuring space in discourse: A comparison of Chinese, English, French and German L1 and English, French and German L2 acquisition. In H. Hendriks (Ed.), *The structure of learner varieties* (111–156). Berlin: Mouton de Gruyter.

Hohenstein, J. (2001). *Motion event similarities in English- and Spanish-speaking children.* PhD Dissertation, Yale University.

Huddleston, R. (2000). *Introduction to the grammar of English.* Cambridge: CUP.

Huddleston, R. & Pullum, G. K. (2002). *The Cambridge grammar of the English language.* Cambridge: CUP.

Inagaki, S. (2001). Motion verbs with goal PPs in the L2 acquisition of English and Japanese. *SSLA 23,* 153–169.

Malt, B. C., Sloman, S. A. and Gennari, S. P. (2003). Speaking versus thinking about objects and actions. In D. Gentner, and S. Goldin-Meadow (Eds.), *Language in mind* (81–111). MIT Press: Cambridge, MA.

Naigles, L., Fowler, A., & Helm, A. (1992). Developmental shifts in the construction of verb meanings. *Cognitive Development 7,* 403–428.

Schlyter, S. (1984). L'acquisition des verbes de déplacement/ mouvements par des adultes suédois. *Papers from the Institute of Linguistics University of Stockholm 52,* 3–43.

Slobin, D. I. (1996). Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani and S. A. Thompson (Eds.), *Grammatical constructions – Their form and meaning* (195–219). Oxford: Clarendon Press.

Slobin, D. I. (1997). Mind, code, and text. In J. Bybee et al. (Eds.), *Essays on language function and language type* (437–467). Amsterdam: John Benjamins.

Slobin, D. I. (2000). Verbalised events: A dynamic approach to linguistic relativity and determinism. In S. Niemeier and R. Dirven (Eds.), *Evidence for linguistic relativity* (107–138). Amsterdam: John Benjamins.

Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner, and S. Goldin-Meadow (Eds.), *Language in mind* (157–192). MIT Press: Cambridge, MA.

Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömqvist and L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspectives* (219–257). Mahwah, NJ: Lawrence Erlbaum Associates.

Slobin, D. I. (2006). What makes manner of motion salient? Explorations in linguistic typology, discourse and cognition. In M. Hickmann & S. Robert (Eds.), *Space in languages: Linguistic systems and cognitive categories* (59–81). Amsterdam: John Benjamins.

Slobin, D. I. (2008). *Paths between verb-framing and satellite-framing of motion events: Speculations and Questions.* Draft.

Stanojčić, Ž., et al. (1989). *Savremeni srpskohrvatski jezik i kultura izražavanja.* Beograd: Zavod za udžbenike i nastavna sredstva.

Stevanović, M. (1989 [1964]). *Savremeni srpskohrvatski jezik Vol. 1.* Beograd: Naučna knjiga.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description Vol. 3* (57–149). Cambridge: CUP.

Talmy, L. (2000). *Toward a cognitive semantics Vol. 2.* Cambridge, MA: MIT Press.

Vidaković, I. (2006). *Second language acquisition of dynamic spatial relations.* PhD Dissertation. RCEAL, University of Cambridge.

CHAPTER 15

Metaphoric competence in the first and second language

Similarities and differences

Jeannette Littlemore
University of Birmingham

1. Introduction

In recent years there has been a growing interest in the teaching of metaphor to second language learners (Danesi, 2008; Holme, 2004), which has partly been due to an appreciation of the fact that the ability to understand and use metaphors in the target language is likely to make a substantial contribution to second language proficiency. An ability to understand and produce metaphor can help learners to develop their sociolinguistic, illocutionary, grammatical, discourse and strategic competence (Littlemore and Low, 2006a, b). Despite this growing interest, there has been surprisingly little research into the extent to which language learners are able to transfer their metaphor interpretation and production skills (or behaviour patterns) from their mother tongue (L1) to the target language (L2). The only study of which I am aware is Johnson (1989) who looked at metaphor interpretation skills in a group of bilingual children. The first language of all the children was Spanish but they all attended English-medium schools. She found that the complexity of the interpretations offered in English correlated significantly with the complexity of interpretations offered in Spanish. These findings suggest that metaphor interpretation skills constitute an individual difference variable. Despite these intriguing findings, no further studies have looked at how metaphoric competence develops in the bilingual lexicon, or at how it contributes to the sort of linguistic multicompetence that L2 users acquire (Cook, 2002). Information on this matter would help language teachers to decide upon the extent to which it is necessary to teach metaphor skills in the target language, and the extent to which they can safely assume that they consist of cognitive skills that learners will simply transfer from their L1.

In order to investigate this issue, an exploratory study was conducted to investigate whether or not there is a relationship between metaphor interpretation and production behaviour in the L1 and the L2. The focus of this study was on decontextualised novel linguistic metaphors, and its findings do not tell us anything about a language learner's ability to understand and produce either conceptual metaphors (Lakoff and Johnson, 1980), or

contextualised or conventional linguistic metaphors (Cameron, 2003), but they do tell us something about the relationship and differences between L1 and L2 metaphoric processing and/or ability.

2. The development of metaphor interpretation and production skills in the first language

The ability to understand and produce metaphors develops at a relatively early stage in first language acquisition. Research on children has suggested that both the production and comprehension of linguistic metaphors are skills that increase with age. Gardner et al. (1974) carried out an experiment into children's capacities to both create and appreciate metaphors. In the first part of their experiment, the children were instructed to think of appropriate metaphorical endings for a list of twelve very short stories. In the second part the children were told to choose the "best" metaphor from a list of four metaphors containing a novel metaphor, a literal comparison, a conventional metaphor and an inappropriate metaphor. In each case, they were told that marks would be gained for novelty.

Gardner et al. observed a tendency, increasing with age, towards novel metaphoric production and preference. They also found that the highest percentages of novel metaphors were produced by the youngest participants and the oldest participants. Where these two groups differed was in the tendency of young participants to produce metaphors which were highly original but inappropriate or nonsensical. Their findings led them to claim of the oldest participants that (ibid.: 138–139):

> The greater cognitive sophistication of the older participants enables them to appreciate the links between the disparate realms... The oldest participants generally display an explicit appreciation of metaphor, often prefer metaphoric endings, have some capacity to produce vivid comparisons if not always to shift across domains, and the critical ability lacking among the youngest participants. They alone possess the metalinguistic flair of reflecting upon the words of others, and of considering what is appropriate in a given context.

However, the idea that the ability to understand metaphor does not develop until late childhood has been contested. Waggoner and Palermo (1989) asked thirty two children of three age levels (five, seven and nine years) and thirty two college students to interpret and explain metaphors describing love, hate, happiness, sadness, anger, and fear. Their results showed that even the youngest children demonstrated some ability to interpret metaphors, although they could not explain their interpretations. The appreciation of metaphor by young children is also observed by Elbers (1988) who notes children's tendencies to comment on similarities from the age of two, although they cannot explain the meaning of these similarities. The general consensus would seem to be that children do possess some ability to perceive metaphor from a very early age, but that this ability increases as they get older and that older children are more able to explain and produce metaphor[1].

Researchers have wondered whether this change is due to a developmental increase in domain-specific knowledge, linguistic ability or general cognitive capacity. Billow (1975) carried out an experiment, the results of which led him to the tentative suggestion

that metaphoric competence was related to cognitive development. In order to test this further, Johnson (1991) examined the metaphor interpretation skills of children with differing levels of language proficiency (as measured by the oral language subscale of the Woodcock Language Proficiency Battery, – Woodcock, 1980) and mental capacity (as measured by the Figural Intersections Test – Pascual Leone and Burtis, 1975). She found that mental capacity was a more important indicator of metaphor interpretation ability than language proficiency.

Support for the idea of metaphoric development as a cognitive phenomenon has also been demonstrated by Kogan (1983) who administered the Metaphoric Triads Task (MTT) to children of different ages. The MTT is a non-linguistic test comprising twenty nine triads of pictures. Each triad offers three pairing possibilities. For example, one triad offers pictures of a toddler on the grass, a brightly coloured watering can, and a rosebud. The following pairing links are possible:

> Toddler and watering can: The toddler can play with the watering can;
>
> Watering can and rosebud: One can sprinkle water on the rosebud with the watering can;
>
> Toddler and rosebud: Both are at an early stage of development

The third link is considered to be the metaphorical link. The children are asked to make as many pairings as possible between the pictures. As this test has no overt linguistic element it is hypothesized to be a test of pure cognitive development. Using this test, Kogan found a progressive increase in metaphoric comprehension across an age span extending from kindergarten to young adulthood.

The idea that the ability to interpret metaphors involves generic cognitive operations provoked two attempts to investigate various Piagetian operations as possible cognitive prerequisites for metaphoric thinking. Billow (1975) investigated the relationship between *combinational reasoning* (one of Piaget's formal operations) and metaphoric comprehension, though he was unable to find clear evidence for a relationship. Cometa and Eson (1978) found that *intersectional classification* (one of Piaget's concrete-operational operations) did correlate with an ability for metaphoric comprehension. Thus it would seem that the ability to understand metaphors is a cognitive phenomenon which develops during the concrete operations period (four-to-eleven year olds). This finding is in accordance with all the above experiments which suggest that metaphoric competence is a cognitive skill which develops during childhood. Clearly then, most types of metaphoric competence involve cognitive skills, and they develop with age.

More recent work on metaphor comprehension in the first language has also suggested that metaphor interpretation and production reflects underlying general cognitive processes (see Evans and Green, 2006). These processes include the activation of relevant domain knowledge (Giora, 2003), imagery (Li, 2002), episodic memory (Bottini, et al., 1994), analogical reasoning (Paivio and Walsh, 1993), categorization (Glucksberg et al., 2001), the use of context (Gibbs, 1994), associative fluency (Johnson and Rosano, 1993), and conceptual blending (Fauconnier and Turner, 1998). Although there is very little universal agreement over which of these processes are involved in metaphor production and comprehension, or how they interact, the fact that they are all cognitive processes

suggests that people may exhibit individual differences in these areas that are relatively stable across languages. Thus a tendency/ability to understand and produce metaphor in the L1 may be related to the same tendency/ability in the L2. That is to say, people who are good at understanding and producing metaphor, or who like to do so, in the L1, may also display the same tendency in the L2, as far as their linguistic resources permit. In the following section I review the limited amount of research that has investigated L1 and L2 metaphor comprehension, interpretation and production.

3. The relationship between metaphor interpretation and production skills in the L1 and the L2

There have been a few studies comparing the metaphoric competence of L1 speakers with that of L2 speakers. Both Trosborg (1985) and Johnson and Rosano (1993) have compared the metaphoric interpretation abilities of native speakers with those of language learners. The only overall difference observed by Trosborg between L2 students and native speakers was that the native speakers had a higher tendency than all the other groups to produce conventional metaphors. This could be because the native speakers were simply more aware of the conventional metaphors than the other groups. The difference would therefore not seem to be due to variations in cognitive processing, but simply to gaps in the knowledge of the L2 learners. Johnson and Rosano found no group differences for either of the tests of metaphor interpretation. Finally Johnson (1996) found that L2 proficiency and L2 metaphor interpretation abilities were unrelated, and that a better predictor of L2 metaphor interpretation abilities was the participant's L1 metaphor interpretation abilities. It is therefore reasonable to hypothesise that L1 and L2 metaphoric competence are very probably related. However, neither of these studies looked at performance in the tests by the same population using their L1 and their L2. Moreover, they did not investigate different types of metaphoric competence. Given the range of functions performed by metaphor, it is important that language learners can *understand it*. They need to be able to *understand it relatively rapidly* as an inability to do so may hold up the conversation and may thus be a source of frustration on both sides. It is also important that they can *interpret it* at different levels as one of the main advantages of using metaphor is that it can say several different things at once. Of less immediate necessity, but still relatively important is the ability to *produce metaphor* in the target language.

The aim of the study described in this article was to explore metaphor comprehension and production skills in the L1 and the L2 in order to determine whether there is a relationship between them, and whether participants are generally better at dealing with metaphor in their L1 or their L2. Four dimensions of metaphoric competence were explored. These were: tendency to find meaning in metaphor, speed in finding meaning in metaphor, ability to identify multiple interpretations for a given metaphor, and novel metaphor production.

4. The study

The study was designed to investigate two research questions:

1. If a student displays a high level of metaphoric competence in the L1, will they also display a high level of metaphoric competence in the L2?
2. Are levels of L1 metaphoric competence higher or lower than levels of L2 metaphoric competence?

In order to answer these questions, four tests were devised. These were: a test of one's tendency to find meaning in metaphor, a test of speed in finding meaning in metaphor, a test of one's ability to identify multiple interpretations for a given metaphor, and a test of novel metaphor production. The participants were 82 upper-intermediate French-speaking university students of English studying at a university in Belgium. All of the participants were in their second year of a degree in Languages and Linguistics. Seventy-five of the participants were female and seven were male. Before participating in the tests the participants were informed by their class teachers that they would be asked to participate in a piece of research into language learning. They were told that participation was optional but that the tests would take place during class time. Out of a total possible of eighty five students, eighty two agreed to complete the tests and signed consent forms.

The tests were administered in two sessions, a two-hour session was based in a language classroom and language laboratory, and one one-and-a-half-hour session was based in a computer suite. In each session there were approximately ten students. All students completed the classroom-based session first and the computer-based session second, with a two-week period separating the two sessions. The participants asked to perform the tests in English and in French. The format of the test was identical in both languages, although the items varied. The tests were as follows:

4.1 Testing the participants' tendency to find meaning in metaphor and speed in finding meaning in metaphor

In order to measure one's tendency to find meaning in metaphor, a test was created which was partly inspired by the test devised by Pollio and Smith (1979). In Pollio and Smith's test participants were given a list of sentences of the form "The _____ is a _____" and they were asked to classify each sentence as one of the following five types:

For example:
A typical type 1 (synthetic) sentence was "The dog is a poodle."
A typical type 2 (analytic) sentence was "The tulip is a flower."
A typical type 3 (contradictory) sentence was "The dog is a cat."
A typical type 4 (anomalous) sentence) was "The mountain is a frog."
A typical type 5 (metaphoric) sentence was "The woman is a rose."

As the focus of interest in this study was on students' tendencies to differentiate between metaphor and anomaly, it was decided that only the fourth and fifth categories in Pollio

and Smiths' experiment would be used (metaphor/anomaly). Pre-piloting showed that participants found it easier to make their choice from a continuum of responses rather than to have to make a yes/no decision.

In conversational discourse containing metaphors it is important that the metaphors are understood quickly if the communication is to remain natural. This is particularly so in the case of foreign language discourse. If a learner is able to produce and to comprehend metaphors at a natural speed then this is likely to make a considerable contribution to his/her communicative competence. It was also suggested above that the speed with which people find meaning in metaphors reflects their *loose analogical reasoning* skills. The reasoning behind this assumption is that the fast interpretation of a metaphor depends on the speed with which a participant is able to draw analogies between its topic and vehicle.

A decision was therefore taken to measure not only the students' tendency to find meaning in metaphors, but also the time required for them to make a positive response whenever they did so. It was decided that a computer-based version of the test would permit accurate measurement of the students' response times. This is not the first time that a computer-based test has been used to measure the amount of time required to process metaphors. Gregory and Mergler (1990) investigated the amount of processing time required to decide whether or not metaphors were meaningful, and the processing time required to decide whether or not metaphorical sentences made sense. They found that participants took significantly longer to decide whether or not metaphors were meaningful than they did to decide whether or not they made sense, thus highlighting the importance of the instructional set. In light of Gregory and Merglers' findings, extra attention was paid to the clarity of the instructional set in this test.

Piloting the tests

In four pilot sessions a total number of thirty eight participants matched to the target population completed the following computer-based metaphor test which was created using Hypercard:[2]

First the participants were shown a rubric (see Appendix 1) explaining what is meant by the terms "metaphor" and "anomaly". This rubric also told them that they were going to be shown a series of metaphors and that their task was going to be to rate, on a scale of one to five, the extent to which they thought each metaphor made sense. In this rubric it was made very clear that their task was not to say whether or not they agreed with the metaphor. When they had read the rubric, they were told to click on a button in the bottom right hand corner of the screen and at this point the test began.

Throughout the test a French version of the following scale was displayed on the right hand side of the screen:

(5) "It's obviously a metaphor. The relationship between the two elements is clear"
(4) "The metaphor is less convincing. One can see that there is a relation but it's not immediately obvious"
(3) "This is the middle of the scale. You're really not sure if it's a metaphor or not"
(2) "There could be a metaphorical meaning but you can't see it"
(1) "It's obviously an anomaly. It is not possible to find a relationship between the two elements"
? There are words I do not understand in this sentence

Thirty three English metaphors and thirty four French metaphors were then displayed, one after the other at the top of the screen, and the participants were asked to make their decisions and to click on the appropriate part of the scale for each response. Half of the participants were shown the French metaphors first, and half of the participants were shown the English metaphors first.

The metaphors had been pre-selected from a study by Katz et al. (1988) in which four hundred and sixty four metaphors were normatively rated on ten scales by six hundred and thirty four raters. According to Katz et al., the aim of this study was to help researchers interested in systematic investigation of metaphoric processes control for different metaphor-related variables. The ten scales along which the metaphors were measured were: comprehensibility; ease of interpretation; degree of metaphoricity; metaphor goodness; metaphor imagery; participant imagery; predicate imagery; felt familiarity; semantic relatedness; number of alternative interpretation. For this test, metaphors which scored highly on comprehensibility were chosen. There were two reasons for the decision. Firstly, the highly comprehensible metaphors contained vocabulary that the students were more likely to understand. Secondly, the chosen metaphors were relatively concrete, and concrete metaphors might be expected to provoke more imagery than abstract metaphors in the interpreting process. Six "nonsense" metaphors were also included in this piloting session to control for random response. The French metaphors were translations of metaphors taken from the Katz et al. study. One possible weakness of the study is that the Katz et al. ratings may not have held for the French translations.

Statistical analysis revealed no random response. Furthermore the reliability of the test was found to be high (Cronbach's alpha = 0.88 for the English version and 0.84 for the French version). It was therefore decided that in the final version of the test twenty five English metaphors and twenty five French metaphors would suffice.

The Katz et al. scales were useful for deciding whether or not the participants actually carried out the procedure suggested in the rubric. If they had understood the rubric as it was meant to be understood their replies would correlate most closely either with Katz et als.' "comprehensibility" or "semantic relatedness". However, correlational analysis revealed that the participants' results correlated most closely with Katz et als.' "metaphor goodness" (correlation coefficient = 0.76; p = 0.00). The correlations with comprehensibility and semantic relatedness were only (0.62) and (0.65) respectively which means that the participants in this pilot study had judged the metaphors not for comprehensibility or for semantic relatedness, but for whether or not they thought that they were good metaphors. However there was a high correlation both between Katz et al's measures of metaphor goodness and comprehensibility (correlation coefficient = 0.84; p = 0.00) and semantic relatedness (correlation coefficient = 0.90; p = 0.00) indicating that the three are very closely related.

The final version of the test

The final version of the test was identical to the pilot version except that the students were asked to assess the meaningfulness of only fifty metaphors (twenty-five in English and twenty-five in French) (see Appendix 2), on a scale from one to five. Also, only two nonsense metaphors were left in as control items. The scale of responses was the same as above, except that this time the following message flashed on the screen:

> ATTENTION!
>
> Il ne s'agit pas de juger de la qualité du lien entre les deux éléments mais de se limiter à décider s'il y en a un ou pas!
>
> (Remember that your task is simply to decide whether there is a relationship between the two elements of the sentence, not to judge the quality of that relationship!)

This was to discourage the students from making judgements of "metaphoric goodness" and to encourage them to focus on comprehensibility and semantic relatedness. Both the types of response and the reaction times were measured for each student for each metaphor (= test of tendency to accept the metaphors as meaningful, and speed of interpretation). Two scores were calculated for each student: a score indicating the average response given (on the scale from 1 to 5). This was designed to show the general tendency of each student to find meaning in metaphor; and a score indicating the average amount of time taken to decide on a positive response (4 or 5).This was designed to give an indication of the speed with which a student was able to find meaning in metaphor.

Correlational analysis revealed that students were still making judgements based primarily on "metaphor goodness" on the French part of the test (correlation coefficient = 0.9; $p < 0.01$), but that their judgements in the English part of the test were more related to comprehensibility (correlation coefficient = 0.83; $p < 0.01$). This is perhaps to be expected as one's first focus, when dealing with another language is on whether or not one can understand it; judgements of quality are something of a second-order decision in that they presuppose comprehension. The fact that the students seemed to be doing slightly different things in the different language versions of the test is a potential weakness of the study. The implications of this are discussed in Section 6 below. The control items showed that, again, there was no random response factor. Furthermore, the reliability of the final version of the test was high (Cronbach's alpha = 0.85 for the English version and 0.83 for the French version).

4.2 A written "metaphoric fluency" test

Metaphoric fluency is defined as the mean number of interpretations that a participant can find for a number of metaphors. It was hypothesized above that metaphoric fluency would involve the psychological process of *associative fluency.* The reasoning behind this assumption was that the more associations a participant is able to find between the two components of a metaphor, the more interpretations he/she will be able to make of that metaphor.

In the pilot test, when the students had completed the computer-based test described above they were presented with the five metaphors that they had accepted most readily in each language. They were asked to write down as many interpretations as possible for each metaphor. The reason why they were given the ten metaphors that they had most readily accepted was because it was thought that this way no student would be asked to interpret a metaphor which they had previously described as "uninterpretable". Their interpretations were then counted and the student was duly attributed a score for "metaphoric fluency". This was simply a measure of the average number of interpretations offered by a participant per metaphor.

It was found, however, that giving each student ten different metaphors to interpret on the basis of their results in the previous test yielded data that were difficult to compare statistically. It was therefore decided that in the main study each student would receive the same ten metaphors to interpret (see Appendix 3). These items were selected from the piloting of the computerised metaphoric preference test described above. The metaphors chosen for inclusion in this part of the test were the five English metaphors and the five French metaphors that had been most readily accepted by the highest number of students. When the students had completed the computer-based test, they were then asked to write down as many interpretations as they could think of for these ten metaphors.

The students' metaphor interpretations were then counted by two independent judges in order to assess the students' metaphoric fluency. A problem arose with one of the English metaphors:

> "A dog is a walking stick."

Despite the fact that in the pilot session this was one of the five most readily accepted English metaphors and participants had no problem understanding it, analysis of data from the main study revealed the presence of a vocabulary problem. A majority of the students had not heard of a "walking stick" and thus understood this sentence to mean that a dog is a stick that walks. Naturally this led to problems for the interpretation of the metaphor and the item had to be eliminated from the analysis. There were therefore only four useable items in the English version of the metaphoric fluency test. This explains why, even in the final version of the test, although the reliability of the French version of this test was reasonably high: 0.65, the reliability of the English version was only 0.31.

4.3 A "metaphor production" test

A metaphoric production test was included as it was felt that it would give more free reign to the students' creativity than the other tests. The "metaphor production" test was adapted from a technique originally used by Gardner et al. (1974) to test children's capacity to create and to appreciate novel metaphors, and which was subsequently used by Trosborg (1985) to investigate the metaphoric skills of foreign language students. Trosborg's experiment was divided into two parts. In the first part of the experiment, participants were asked to think up appropriate metaphorical endings for a list of twelve very short scenarios. They were told that credit would be given for novelty. In the second part of the experiment, they were told to choose what they considered to be the "best" ending from a list of four endings containing a novel metaphor, a literal comparison, a conventional metaphor and an inappropriate metaphor. It was assumed that those students who had a strong appreciation of metaphor would elect a novel response. The students were therefore given a version of this test.

In a piloting session, it was found that the reliability of the second part of Trosborg's test was very low (Cronbach's alpha = 0.32), whereas the first part yielded relatively high reliabilities. It was therefore decided to administer only the first part of the original test. Furthermore, it was decided that reliable results could be obtained by using only eight

items (instead of the original twelve). In order to select the eight sentences to be completed, twelve sentences were piloted on a matched group of language students. Only those items that caused no confusion, accounted for statistical variation and were statistically reliable were accepted for inclusion in the final eight items. The students were therefore given eight uncompleted sentences in English and eight in French (half the students were given the English sentences first and half the students were given the French sentences first) and were told to complete them as creatively as possible (see Appendix 4).

The test was adapted slightly for another reason. In the Trosborg study the sentences to be completed all took the form of similes ("for example "it was as quiet as..."). Researchers have suggested that similes probably require slightly different thought processes from metaphors. Ortony (1979 and 1980) suggests two ways in which similes might be seen to differ from metaphors. In his 1979 article he claims that the "nonliteral similarity" contained in metaphors is different from the "literal similarity" contained in similes. In his 1980 article he goes on to propose that metaphors violate what he refers to as "the sincerity postulate" (ibid.: 76). This is a concept derived from one of Grice's (1975) maxims and which basically means "try to mean (literally) what you say and imply". Similes do not violate this postulate and therefore might, according to Ortony, be regarded as literal rather than metaphorical language. Therefore there is a case for not treating them in the same way as we treat metaphors. These differences led Kogan (1983: 660) to claim that "there is clearly no basis for asserting that the ability to reason by analogy is a precursor of metaphoric thinking". As the aim of this study was to examine metaphoric processes, the test was adapted to provoke metaphors rather than similes. After having seen two examples, one in English and one in French, the students were given sixteen very short expressions (eight in French and eight in English) and were asked to produce endings they liked and found right for these expressions. They were asked to complete the sentences as creatively as possible. For example the students were asked to complete the sentence:

> The lake was a shining ..... at the bottom of the valley.
>
> Responses were scored as follows:
>
> 1 = inappropriate ending
>
> 2 = literal ending
>
> 3 = conventional metaphor ending
>
> 4 = novel metaphor ending

Endings were considered to be "novel" only if they met at least one of the following criteria (adapted from Gardner et al., 1974):

a. The topic was projected onto a sensory domain where it was not literally applicable, and the resulting metaphor was not a familiar English or French saying. (For example: Dr. Livingstone had been walking across the Sahara for five days without any water. His throat was beginning to feel as dry as... a sheet of paper in Moses' bible; item E4, participant three)

b. The topic which was typically associated with the physical world was projected on to a psychological state or the reverse, and the resulting metaphor was not a familiar English or French saying. (For example: We could tell by the look on the teacher's face that his anger was... like a rocket searching its target; item E4, participant forty)
c. The topic remained in its customary domain (sense modality or physical reference) but a radical shift in perspective was required, and the resulting metaphor was not a familiar English or French saying. (For example: When I was a child, I was frightened of my grandma's teeth soaking in the glass in the bathroom. They made me think of... an old wreck forgotten in the sea; item E8, participant twenty seven).

So long as the underlying rationale for the metaphor was an established or familiar one, inclusion of a novel word or phrase did not qualify it as "novel". Endings were scored as "conventional metaphors" if the response was a familiar English saying. Endings were scored as "literal" if the adjective remained in its customary domain. Endings were scored as "inappropriate" if neither of the two judges could find a meaning. Examples of items that received each of these scores can be found in Appendix 5.

Scoring of the English part of the test was carried out by two independent native English speakers, and scoring of the French part of the test was carried out by two independent native French speakers and one English speaker. All of these judges had been instructed carefully in the scoring criteria. The English speaking judges achieved 98% agreement in their assessment of responses, and the French speaking judges achieved 95% agreement. In cases of disagreement, negotiations took place until agreement was reached. Most of the replies given were conventional metaphors. Students also provided novel metaphors and literal responses. There were very few inappropriate responses. These findings match those of both Trosborg (op. cit.) and Gardner et al. (op. cit.).

The sum of the scores was calculated for each student. A high score was interpreted as meaning that the student had a preference towards metaphorical production, and a low score was interpreted as meaning that a student had a preference towards literal production. The reliability of final version of the English version of this test (as measured by Cronbach's alpha) was 0.53 and the reliability of the French version was 0.58. Although both of these reliabilities are fairly low, it was decided that, as no students would be affected by any decisions made on the basis of their results in these tests, the tests would be acceptable for research purposes.

5. Results

It will be remembered that the research questions investigated in this study were as follows:

1. If a student displays a high level of metaphoric competence in the L1, will they also display a high level of metaphoric competence in the L2?
2. Are levels of L1 metaphoric competence higher or lower than levels of L2 metaphoric competence?

5.1 Findings for Research Question 1

The first question was whether the scores on the metaphor tests in the L1 would be related to the scores in the L2. This was found to be the case for all four tests. Scores on all four tests in the L1 correlated with scores on the equivalent test in the L2. The results are presented below in Table 1:

Table 1. Correlations between English and French versions of the metaphor tests

	Novel metaphor production (French)	**French metaphoric fluency**	**French metaphor interpretation**	**Speed of French metaphor interpretation**
Novel metaphor production (English)	**0.33** **$p < 0.01$***	−0.00 NS	0.12 NS	0.23 NS
English metaphoric fluency	0.18 NS	**0.35** **$p < 0.01$***	0.22 NS	0.12 NS
English metaphor interpretation	0.19 NS	−0.04 NS	**0.57** **$p < 0.01$***	**0.30** **$p < 0.01$***
Speed of English metaphor interpretation	0.26 NS	0.21 NS	0.11 NS	**0.59** **$p < 0.01$***

This means that students who performed well on the tests in their first language were significantly more likely to perform well on the equivalent tests in their second language, which indicates that a cognitive component is involved in metaphoric competence, as well as a linguistic one. These results also tell us that the English and French versions of the metaphor tests display good convergent validity. The fact that only one other relationship in the Table is significant at the 99% level indicates that the discriminant validity is good too.

Within the languages themselves there were more relationships between the different types of metaphoric competence. The correlations between the different tests in English are shown in Table 2 below:

Table 2. Correlations between the different tests in English

	Novel metaphor production (English)	**English metaphoric fluency**	**English metaphor interpretation**	**Speed of English metaphor comprehension**
Novel metaphor production (English)		0.18 NS	**0.62** **$p < 0.01$**	**−0.26** **$p < 0.05$**
English metaphoric fluency			0.04 NS	−0.16 NS
English metaphor interpretation				−0.19 NS
Speed of English metaphor interpretation				

The correlations between the different tests in French are shown in Table 3 below:

Table 3. Correlations between the different tests in English

	Novel metaphor production (French)	French metaphoric fluency	French metaphor interpretation	Speed of French metaphor interpretation
Novel metaphor production (French)		−0.19 NS	**0.75** **p < 0.01**	**−0.34** **p < 0.01**
French metaphoric fluency			−0.12 NS	−0.09 NS
French metaphor interpretation				**−0.31** **p < 0.01**
Speed of French metaphor interpretation				

In both English and French, it seems that novel metaphor production has most in common with the other types of metaphoric competence. Because of these findings we cannot say that the different types of metaphoric competence are independent of one another. This is perhaps to be expected as to some extent they reflect the same cognitive processes (see Littlemore and Low, 2006b, chapter 3).

5.2 Findings for Research Question 2

The fact that the English and French versions of the metaphor tests correlated with each other does not mean, however, that they were identical. The second research question was interested in the relative levels of performance on the tests in the two languages. As we can see in Table 4, the participants scored more highly in their first language than in their second language on the test of metaphoric fluency, but in all the three other tests they scored more highly in their second language than in their first language:

Table 4. L1 and L2 comparisons of scores on all four tests

Language	Novel metaphor production	Metaphoric fluency	Metaphor interpretation	Speed of metaphor interpretation (a low score indicates a better result)
First language (French)	2.79	1.74	2.91	6.595
Second language (English)	2.83	1.36	3.03	6.478

In order to assess the significance of these differences, t-tests were carried out on all four tests. These tests revealed that for some of the tests there were significant differences in level of performance between the two languages.

No significant difference was identified in the level of performance on the English and French versions of the test of novel metaphor production $t(81) = 1.11$, $p = 0.27$). The level of linguistic competence thus appears to have no effect on this type of metaphoric competence.

There was a significant difference in performance between the English and French versions of the test of metaphoric fluency $t(81) = –5.65$, $p = 0.00$). The results suggest that

the participants displayed higher levels of metaphoric fluency in the L1 than in the L2. This finding is probably due to shortcomings in the students' L2 vocabulary.

A significant difference was found between the English and French versions of the test of metaphor interpretation $t(81) = 2.34$, $p = 0.02$). The results suggest that participants are more likely to find meaning in L2 metaphors than in L1 metaphors. This finding is difficult to explain. It could be that in their L1 students have enough confidence to claim that a given metaphor makes no sense, whereas in their L2 students are used to having to guess meanings and are therefore reluctant to dismiss expressions as unmeaningful.

There was no significant difference between time taken to find meaning in English metaphors and time taken to do so in French metaphors $t(81) = -0.30$, $p = 0.77$). Thus the only two significant findings were that: students were significantly more likely to find meaning in metaphor in their second language, and that they exhibited significantly higher levels of metaphoric fluency in their first language.

The finding that students performed worse in their L1 than in their L2 on the tendency to find meaning in metaphor test, although unusual, is in keeping with Bromberek-Dyzman and Ewert's finding (this volume) that Polish learners of English found it easier to understand metaphor-based implicatures in English than in Polish. It is easier to understand the finding that students performed better in their L1 than in their L2 on the metaphoric fluency test. One would expect students to have greater access to the relevant encyclopaedic knowledge in the L1 to help them perform this kind of test.

One reason for the differences in performance patterns on these two tests may relate to the fact that the tendency to find meaning in metaphor test only required the students to identify one meaning for the items concerned, whereas the metaphoric fluency test required them to find multiple meanings. It could be that students are used to thinking metaphorically in order to find meaning in the L2 (Picken, 2007; Underwood et al., 2004), whereas in their L1 so much of the language is processed formulaically that they no longer need to practise this skill (Wray, 1999). On the other hand, metaphoric fluency involves deeper processing and the activation of a wider area of encyclopaedic knowledge within the mental lexicon. As such, it is understandable that students will be better placed to do this in their L1 than in their L2.

6. Discussion

The findings made in the first part of this study suggest that the ability to understand and produce metaphor in the L1 is related to the ability in the L2. In other words, students who are able to do this in the L1 are also likely to do so in their L2. This suggests that metaphoric competence is a relatively stable individual difference variable, which may partially account for differences in student behaviour and success rates in the foreign language classroom (see also Littlemore, 2001). Further research in this area could usefully investigate the relationship between different types of metaphoric competence and cognitive flexibility. Successful second language acquisition must involve a degree of cognitive flexibility and openness to new ways of seeing things. Cognitive flexibility is also likely to be related to what Grigorenko et al. (2000) refer to as 'Cognitive Ability for Novelty

in Acquisition of Language'. Basically, this involves the ability to spot new patterns in the language input and to use one's existing knowledge selectively, along with analogical reasoning, to work out the new form-meaning pairing.

These findings underline the fact that metaphoric competence must not be seen as a homogeneous trait. Rather, it is, to some extent, a multifaceted entity, and a student can, for example, be good at finding the meaning in metaphor fairly quickly, but not good at producing multiple interpretations. Student performances on these tests were uncorrelated in both English (see Table 2) and French (Table 3). In order to truly understand the findings made in this part of the study, more work is needed investigating the nature of the L1 and L2 mental lexicons and the ways in which they overlap, and searches are carried out within them. A second promising approach would be to explore the ways in which the acquisition of a second or third language extends and enriches the number of possible ways of perceiving, describing and structuring our realities. This corresponds to Cook's (2002) notion of multicompetence, whereby linguistic knowledge is restructured in the mind of a bilingual leading to an integrated system which combines elements from both the L1 and the L2 to produce something new. This integrated system is likely to affect the ways in which learners understand and interpret metaphor.

The second part of the study indicated that students exhibited higher levels of metaphoric fluency in their L1 than in the L2 but that they were more likely to find meaning in L2 metaphors than L1 metaphors. There were no differences in terms of performance on the L1 and L2 versions of the tests of novel metaphor production and speed in finding meaning. The first of these findings is perhaps to be expected. As we discussed above, this test is likely to reflect vocabulary knowledge in the two languages. The other findings are more surprising. The finding that students are more likely to find meaning in L2 metaphors than L1 metaphors appears, at first sight, to be somewhat counter-intuitive. One possible explanation for this finding however is the fact that, as we saw above, the ratings given by the students for the individual items correlated with ease of comprehension on the English version of the test, whereas they correlated with goodness of metaphor on the French version of the test. This indicates that they were looking for slightly different things in these different parts of the test, which could explain the discrepancies in the findings. Items that are being assessed for ease of comprehension will be more likely to receive higher scores than items that are being assessed for 'goodness'. This needs to be acknowledged as a limitation of the study.

The study has a number of other limitations that could be remedied in future research. For example, it has not really unpacked the issue of metaphoric *competence* versus metaphoric *preference*. Although the rubrics encouraged the participants to try their hardest with the metaphor activities it could be that some students simply felt more motivated than others to do well in the tasks and this may explain the significant correlations across the two languages. On the other hand, even if this were the case, the discovery of individual differences in terms of metaphoric preference is also of interest as a substantial amount of second language learning involves thinking metaphorically. If learners vary in terms of how much they like to do so, then this could perhaps underlie some of the differential success rates that teachers observe in their students. Motivation to learn a second language is inevitably a complex issue (Dornyei, 2001) but at some level it is likely

to involve a love of language, and a willingness to 'play' with the target language (Cook, 2000). It would be very interesting to investigate how metaphoric competence relates to a propensity for language play.

Finally, this study has focussed on the comprehension and production of decontextualised metaphors in highly artificial test sessions. This of course is very different from the sort of thing that learners will need to do in the real world. Future studies should look more at naturally-occurring discourse, or compare the metaphor use of learners engaged in more naturalistic activities in their first and second language.

Notes

1. Care must be taken when interpreting these findings as there is likely to be variation in the children's tendencies to report on all the similarities they observe. However, the change from literal to metaphoric understanding has been widely observed across a number of different types of metaphor tests. Some of these tests (for example the Metaphoric Triads Test – see Kogan, 1983) contain a highly reduced reporting element.

2. Prior to this test, the participants were given training in how to use the mouse. This training consisted of a scale on the screen from one to five (see Appendix 2). A number appeared in the top left hand corner of the screen and the participant had to use the mouse to click on the corresponding number in the scale. If a number appeared that was not in the scale the participants were instructed to click on a box labelled "other". As soon as the participant had clicked in the appropriate place, another number appeared in the top left hand corner of the screen and the participant had to repeat the procedure. This procedure continued until the average reaction time of the participant was below three seconds or until thirty trials had taken place, whichever happened first.

References

Billow, Richard (1975). A cognitive developmental study of metaphor comprehension. *Developmental Psychology 11*, 415–423.

Bottini G., Corcoran R., Sterzi R., Schenone P., Scarpa, P., Frackowiak, R. S. J., & Frith, C. D. (1994). The role of the right hemisphere in the interpretation of Figurative Aspects of Language: A Positron Emission Tomography Activation Study. *Brain 117* (2), 1241–125.

Bromberek-Dyzman, Katarzyna & Ewert, Anna. (This volume). Figurative competence is better developed in L1 than in L2, or is it? Understanding conversational implicatures in L1 and L2.

Cameron, Lynne (2003). *Metaphor in educational discourse*. London: Continuum.

Cometa, Michael, & Eson, Morris (1978). Logical operations and metaphor interpretation: A Piagetian interpretation. *Developmental Psychology 49*, 649–659.

Cook, Guy (2000). *Language play, language learning*. Oxford: Oxford University Press.

Cook, Vivian (2002). Background to the L2 user perspective. In V. J. Cook (Ed.), *Portraits of the L2 User* (1–32). Clevedon: Multilingual Matters.

Danesi, Marcel (2008). Conceptual errors in second language learning. In Sabine De Knop & Teun De Rycker (Eds.), *Cognitive approaches to pedagogical grammar* (231–257). Berlin/New York: Mouton de Gruyter.

Dörnyei, Zoltan (2001). *Teaching and researching motivation*. Harlow: Longman.

Elbers, Lockie (1988). New names from old words: Related aspects of children's metaphors and word compounds. *Journal of Child Language 15*, 591–617.

Evans, Vyvyan & Green, Melanie (2006). *Cognitive linguistics*. An introduction. Edinburgh: Edinburgh University Press.

Fauconnier, Gilles & Turner, Mark (1998). Conceptual integration networks. *Cognitive Science 22* (2), 133–187.

Gardner, Howard, Kircher, Mary, Winner, Ellen & Perkins, David (1974). Children's metaphoric productions and preferences. *Journal of Child Language 2,* 125–141.

Gibbs, Raymond (1994). *The poetics of mind*. Cambridge: Cambridge University Press.

Giora, Rachel (2003). *On our mind. Salience, context, and figurative language*. Oxford: Oxford University Press.

Glucksberg, Sam, Newsome, M. R. & Goldvarg, Y. (2001). Inhibition of the literal: Filtering metaphor-irrelevant information during metaphor comprehension. *Metaphor and Symbol 16* (3 & 4), 277–293.

Gregory, Monica, & Mergler, Nancy (1990). Metaphor comprehension: In search of literal truth, possible sense, and metaphoricity. *Metaphor and Symbolic Activity 5/3*, 151–173.

Grice, Howard (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics 3*. New York: Academic Press.

Grigorenko, E., Sternberg, R. & Ehrman, M. (2000). A theory-based approach to the measurement of foreign language learning ability: the canal-f theory and test. *The Modern Language Journal 84* (3), 390–405.

Holme, Randal (2004). *Mind, metaphor and language teaching*. Basingstoke: Palgrave Macmillan.

Johnson, Janice (1989). Factors related to cross-language transfer and metaphor interpretation in bilingual children. *Applied Psycholinguistics 10*, 157–177.

Johnson, Janice (1991). Developmental versus language-based factors in metaphor interpretation. *Journal of Educational Psychology 83/4*, 470–483.

Johnson, Janice (1996). Metaphor interpretations by second language learners: Children and adults. *The Canadian Modern Language review 53/1*, 219–241.

Johnson, Janice & Rosano, Teresa (1993). Relation of cognitive style to metaphor interpretation and second language proficiency. *Applied Psycholinguistics 14*, 159–175.

Katz, Albert, Paivio, Alan, Marschark, Mark & Clark, James (1988). Norms for 204 Literary and 260 non-literary metaphors on 10 psychological dimensions. *Metaphor and Symbolic Activity 3/4*, 191–214.

Kogan, Nathan (1983). Stylistic variation in childhood and adolescence: Creativity, metaphor, and cognitive styles. In J. H. Flavell & E. M. Markman (Eds.), *A handbook of child psychology 3/4* (695–706). New York: John Wiley and Sons.

Lakoff, George & Johnson, Mark (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Li, Thomas (2002). *The acquisition of metaphorical expressions, idioms, and proverbs by Chinese learners of English: A conceptual metaphor and image schema-based approach*. PhD thesis (Chinese University of Hong Kong).

Littlemore, Jeannette (2001). Metaphoric competence: A possible language learning strength of students with a holistic cognitive style? *TESOL Quarterly 35* (3), 459–491.

Littlemore, Jeannette, & Low, Graham (2006a). Metaphoric competence and communicative language ability. *Applied Linguistics 27* (2), 268–294.

Littlemore, Jeannette, & Low, Graham (2006b). *Figurative thinking and foreign language learning*. Basingstoke: Palgrave Macmillan.

Ortony, Andrew (1979). Beyond literal similarity. *Psychological Review 86*, 161–180.

Ortony, Andrew (1980). Some psycholinguistic aspects of metaphor. In R. P. Honeck and R. P. Hoffman (Eds.), *Cognition and figurative language* (69–83). New Jersey: Erlbaum Press.

Paivio, Alan & Walsh, M. (1993). Psychological orocesses in metaphor comprehension and memory, In A. Ortony (Ed.), *Metaphor and thought* (307–308). 2nd ed. Cambridge: Cambridge University Press.

Pascual-Leone, J., & Burtis, J. (1975). Manual for FIT: Figural intersections test. Unpublished manuscript, York University Department of Psychology, Toronto. Cited in Johnson J. (1991). Developmental versus language-based factors in metaphor interpretation. *Journal of Educational Psychology 83/4*, 470–483.

Picken, Jonathan (2007). *Literature, metaphor and the foreign language learner*. Basingstoke: Palgrave MacMillan.

Pollio, Howard & Smith, Michael (1979). Sense and nonsense in thinking about anomaly and metaphor. *Bulletin of the Psychonomic Society 13* (5), 323–326.

Trosborg, Anna (1985). Metaphoric productions and preferences in second language learners. In William Paprotté and René Dirven (Eds.), *The ubiquity of metaphor* (525–557). Amsterdam: John Benjamins Publishing Company.

Underwood, Geoffrey, Schmitt, Norbert & Galpin, Adam (2004). The Eyes Have It: An eye-movement study into the processing of formulaic Sequences. In N. Schmitt (Ed.), *Formulaic sequences.* (Amsterdam/Philadelphia: John Benjamins Publishing Company).

Waggoner, James & Palermo, Diane (1989). Betty is a bouncing bubble: Children's comprehension of emotion-descriptive metaphors. *Developmental Psychology 25/1*, 152–163.

Woodcock, Richard (1980). *Woodcock language proficiency battery, English form*. Hingham Mass.: Teaching Resources.

Wray, Alison (1999). Formulaic language in learners and native Speakers. *Language Teaching 32*, 213–231

Appendix 1

English translation of the rubric used in the computer-based metaphor test

A metaphor is a statement which is not literally correct, but which establishes a relationship between two parts of a sentence. The ease with which this relationship can be interpreted can vary.

For example, the statement "snow is a Winter coat" is an obvious metaphor – snow is not a coat, but the idea of a coat provides relevant information about snow – it covers everything, it keeps you warm, it's thick...

On the other hand, the sentence "the piano is a spoon" cannot really be considered to be a metaphor: it is difficult to see what kind of information a spoon can give about a piano. This kind of expression is known as an anomaly.

You are going to be presented with two sets of sentences, in French and in English.

Your task is to decide to what extent each sentence can be said to be a metaphor, in other words, to what extent one element can be said to provide information about the other.

You are not being asked to judge the quality of the metaphors, but simply to decide whether you think they are metaphors or not.

Please indicate your response as soon as you have decided whether or not there is a relationship between the two elements.

Appendix 2

Metaphors used in the computer-based metaphor comprehension test

A dog is a walking stick

A dog's stomach is his master's alarm clock

A dream is a solar eclipse

A photograph is a one-sided skin of truth

A smile is a knife

A soft-boiled egg is a guillotined aristocrat

A sugar-cube melting in coffee is the fading of a ghost

Aeroplanes are angry birds

Beaches are grills

Creativity is a toaster

Death is a fat fly

Death is the cruel singing of deathless mosquitoes

Evolution is a lottery

History is a sponge

Memory is a snake

Mimes are wooden statues

Music is death

Nature is a vast laboratory

Night is a castle

Smiles are the channels of future tears

The sky is a parliament

The wind is a cat

Time is a bird

Winter is the warm south

Wisdom is a weatherman

L'esprit createur est une bouilloire sur le feu

L'esprit est une eponge

L'histoire est un sport d'hiver

L'hiver est un oiseau avec des dents

L'homme est un collier qui cherche un chien

L'humiliation est un rideau

La liberté est un deuxième soleil

La lune est un chat dansant

La mort est un joueur de tambour

La parole est la semence de la misère

La sagesse est un météorologiste

La terre est une bouche parfumée

Le ciel est un ordinateur

Le clair de la lune est de l'eau de javel

Le subconscient est une arène

Les étoiles errantes sont des enfants qui ne connaîssent pas leur arithmétique

Les nuages sont des mondes en laine

Les panneaux d'affichage sont des verrues sur le paysage

Un éléphant est une petite serre

Un chirurgien est un décorateur d'intérieurs

Un désir est un arc-en-ciel

Une île est un bouchon

Une armée est une mer vivante

Une dynastie est une pièce de théâtre

Une station d'essence est une oasis

Appendix 3

Answer sheet for metaphoric fluency test

METAPHOR INTERPRETATION SHEET

NAME:

Use this booklet to write down all the relationships which you can see between the two elements of each metaphor. Use English for metaphors 1 to 5 and French for metaphors 6 to 10.

Remember, the metaphors are:

1. A dog is a walking stick.
2. Nature is a vast laboratory.
3. Smiles are the channels of future tears.
4. A dream is a solar eclipse.

5. Evolution is a lottery.
6. L'esprit créateur est une bouilloire sur le feu.
7. La liberté est un deuxième soleil
8. Les panneaux d'affichage sont des verrues sur le paysage
9. Une armée est une mer vivante
10. Une station d'essence est une oasis

Please ask a teacher if you do not know the meaning of any of these words.

Appendix 4

Sentence starters used in the test of original metaphor production

English sentence starters

1. In Winter, the weather in Scotland is extremely cold. As soon as you go out of the house your face starts to feel.....
2. Tom hasn't cleaned his room for ages and it's starting to smell. The smell reminds me of.....
3. We could tell by the look on the teacher's face that his anger was.....
4. The lake was a shining ..... at the bottom of the valley.
5. Peter's violin playing isn't marvellous, but compared to that of Alf it sounds like.....
6. Dr. Livingstone had been walking across the Sahara for five days without any water. His throat was beginning to feel as dry as.....
7. What a beautiful day! The clear sky reminds me of.....
8. When I was a child, I was always frightened of my grandma's teeth soaking in the glass in the bathroom. They made me think of.....

French sentence starters

1. Pierre était dans la salle. Il était évident qu'il avait vu quelque chose d'effrayant. Il tremblait comme.....
2. Agnès renverse tout le temps quelque chose. On dirait.....
3. Les deux garçons entraient dans la hutte. Ils avaient très peur car il y faisait noir.....
4. Après sa maladie, mon père est devenu sourd.....
5. La police l'avait presque rattrapé. Il se sentait.....
6. La sorcière était très âgée ; elle semblait presque avoir la peau.....
7. Il faut faire attention si tu sors avec ce mec là. J'ai entendu dire qu'il boît.....
8. C'est vrai que maintenant elle est vieille et laide, mais quand elle était jeune, elle avait une peau.....

Appendix 5

Examples showing how the items were scored in the metaphor production test

English Examples

The following were scored as novel metaphors:

We could tell by the look on the teacher's face that his anger was... as terrible as the exam questions he gave us last year. (item E3, participant 30)

The lake was a shining... and warm bath for the most terrible leviathan I had ever seen..., at the bottom of the valley. (item E4, participant 20)

The following were scored as conventional metaphors:

Dr Livingstone had been walking across the Sahara for five days without any water. His throat was beginning to feel as dry as... the sand. (item E6, participant 54)

The lake was a shining... diamond ... at the bottom of the valley. (item E4, participant 58)

The following were scored as literal responses:

In winter the weather in Scotland is extremely cold. As soon as you go out of the house your face starts to feel the cold winds from the North. (item E1, participant 1)

What a beautiful day! The clear sky reminds me of... my home country. (item E7, participant 33)

The following were scored as inappropriate responses:

When I was a child, I was always frightened of my grandma's teeth soaking in the glass in the bathroom. They made me think of... old chewing gums. (item E8, participant 29)

The lake was a shining... river in fact... at the bottom of the valley. item E4, participant 62)

French Examples

The following were scored as novel metaphors:

Agnès renverse tout le temps quelque chôse. On dirait... qu'elle a été une boule de bowling dans une autre vie. (item F2, participant 12)

(Agnes is always knocking things over. You might say... that in a previous life she was a ten pin bowling ball.)

La police l'avait presque rattrapé. Il se sentait coincé comme une praline Léonidas dans une boîte d'allumettes. (item F5, participant 12)

(The police had almost caught him. He felt... as squashed as a Leonidas chocolate* inside a matchbox.)

* Leonidas chocolates are a particularly large kind of Belgian chocolates.

The following were scored as conventional metaphors:

C'est vrai que maintenant elle est vieille et laide, mais quand elle était jeune, elle avait une peau ... de pêche*. (item F8, participant 19)

(It's true that now she's old and ugly, but when she was young she had skin... like a peach.*)

Il faut faire attention si tu sors avec ce mec là. J'ai entendu dire qu'il boît... comme un trou.* (item F7, participant 35)

(You'd better watch out if you go out with that bloke. I've heard it said that he drinks... like a hole.*)

*Conventional French idioms.

The following were scored as literal responses:

Agnès renverse tout le temps quelque chôse. On dirait...qu'elle fait exprès. (item F2, participant 1)

(Agnes is always knocking things over. You might say that... she does it on purpose.)

Les deux garçons entraient dans la hutte. Ils avaient très peur car il y faisait noir... et celle-ci se trouvait au fond de la forêt. (item F3, participant 24)

(The two boys went into the hut. They were very frightened because it was dark... and the hut was in the depths of the forest.)

The following were scored as inappropriate responses:

Agnès renverse tout le temps quelque chôse. On dirait... qu'elle attire les objets. (item F2, participant 32)

(Agnes is always knocking things over. You might say... that she attracts objects.)

Agnès renverse tout le temps quelque chôse. On dirait... que c'est un peu pour elle. (item F2, participant 74)

(Agnes is always knocking things over. You might say... that it's a bit for her.)

CHAPTER 16

Figurative competence is better developed in L1 than in L2, or is it?

Understanding conversational implicatures in L1 and L2

Katarzyna Bromberek-Dyzman and Anna Ewert
Adam Mickiewicz University, Poznań

1. Introduction

In the present chapter we are concerned with how people who know more than one language comprehend and interpret figurative meanings. Addressing the issue of figurative competence, we are particularly interested in the interpretation of conversational implicatures as defined by Grice (1975), as this approach permits a wider discussion of a broad scope of different figurative meanings.

Our view of language is that of multicompetence (Cook 1992, 2002), in which the two or more languages of an L2 user form an integrated whole. Moreover, we consider language to be inextricably connected with general cognitive processing. For this reason, individuals' linguistic and communicative skills depend on the totality of experiences they have had, not only with each of the languages in their repertoire separately, but also with language and communication in general.

A large body of research evidence indicates that even proficient language learners have difficulty interpreting figurative meanings in the L2. The reasons for this are wide ranging. Since the two languages of an L2 user are interconnected in multicompetence, in our study we focus on the interpretation of figurative meanings in both the L1 and the L2. The aim of the study is to show how L2 figurative competence interacts with L1 figurative competence and whether L2 figurative competence of proficient L2 users is similar to or different from their L1 figurative competence.

2. Multicompetence

Multicompetence is defined simply as the knowledge of two or more languages in the same mind (Cook, 1991, 1992, 2002). The different languages are interconnected in the

L2 users' mind, forming one overall system (Cook, 1991: 115). Cook's view of multicompetence is consistent with the dynamic approach to multilingualism: "all the knowledge of different languages and varieties that an individual knows are part of one dynamic system" (de Bot, Lowie and Verspoor, 2005: 6). Multicompetence has been reconceptualized from a social-interactional, emergentist position by Hall, Cheng and Carlson (2006).

The two languages of an L2 user are integrated at the lexical (e.g. Costa et al., 2005; Colome, 2001; van Hell and Dijkstra, 2002), semantic (Ameel et al., 2005), conceptual (Athanasopoulos, 2006), pragmalinguistic (e.g. Blum-Kulka, 1991; Cenoz, 2003) and cultural level (Sicola, 2003). Interactions between the two languages are bidirectional (e.g. Pavlenko and Jarvis, 2002), i.e. transfer takes place both from the L1 to the L2 and from the L2 to the L1. The impact of the fact of knowing two or more languages extends beyond transfer into the cognitive domain, which Herdina and Jessner (2002) refer to as intermodularity of transfer. Bilinguals usually have different metalinguistic awareness (see Bialystok, 2001 for a review) and show different degree of sensitivity to the communicative needs of the interlocutor (Genesee et al., 1975). Bialystok (2001) ascribes these differences in metalinguistic knowledge between bilinguals and monolinguals to different cognitive underpinnings of bilingualism.

Apart from cognitive factors, another factor that contributes to differences between L2 users and monolinguals are aspects of language use in different social contexts of interaction and in different pragmatic pursuits. L2 users, by virtue of belonging to different communities of practice, have different experiences of language use. Differences between monolinguals' and L2 users' knowledge of a language arise because of a difference in the sum of contexts in which language is used for meaningful communication by these speakers (Hall, Cheng and Carlson, 2006).

An L2 user is defined for the purpose of the present study as someone who uses the L2 for real life purposes, such as receiving education through the medium of the L2, as opposed to an L2 learner, who does not use the L2 outside of a classroom (Cook, 2002).

3. Figurative competence

Figurative competence is defined by Levorato (1993: 104) as "the ability to deal with figurative language". Levorato distinguishes three features of figurative language: the discrepancy between speaker's meaning and utterance meaning, conventionality, and contextual dependence. The first and most outstanding characteristic is the discrepancy between the speaker's words and their communicative intended value. As Steinmann (1973: 224) put it: "speaking figuratively consists of saying what you mean (intended meaning) by not meaning what you say (literal meaning)". This discrepancy is very much a matter of degree. For instance, in irony the intended meaning is very often conveyed by the exact opposite of the literal meaning of the words uttered, while in metaphor, simile or metonymy the intended meaning is more a matter of substitution of meaning (source to target domain mapping). Conventionality is related to the frequency of usage of a particular figurative expression and its salience in the speakers' minds, which is relative to a community of usage. What is conventional and salient for native speakers, may not be salient at all, or to the same extent,

to those who do not belong to the community of usage. This is why second or foreign language users who know their L2 at an advanced level have consistent problems with figurative language comprehension and production (e.g. Littlemore, 2008; Littlemore and Low, 2006; Cieślicka, 2004). The third feature of figurative language is its contextual dependence. Context is the necessary background against which the speaker's intended meaning may be arrived at by the hearer, through various interpretation heuristics. Opinions on the exact role of the context in figurative meaning interpretation vary. In models assuming a single comprehension mechanism, sensitive to both linguistic and nonlinguistic information, the context is considered to be of primary importance in figurative language comprehension (Gibbs, 1986, 1994, 2001, 2002; Sperber and Wilson, 1986/1995). In two-stage models (Grice, 1975; Searle, 1979; Giora, 1995, 1997, 2002; Giora, Fein and Schwartz, 1998), the literal meaning is obligatorily accessed first, irrespective of the context, and the figurative meaning is attempted optionally, via implicature or other inferential processing, only after literal interpretation failed because of lack of contextual fit.

There seem to be different cognitive underpinnings and different onsets of the ability to make sense of different tropes, especially metaphor, idiom, irony and scalar implicature, in the L1 (see Winner, 1988, 1995 for a review). Systematic comparative studies between metaphor and other types of figurative language have been undertaken. De Groot et al. (1995) study the similarities and differences between the understanding of metaphor and irony. Happe (1995) explores the cognitive underpinnings of metaphor and irony in the framework of relevance theory (Sperber and Wilson, 1986/1995) and the theory of mind, examining the child's ability to attribute mental states to the speaker and to understand utterances in which what the speaker says is different from what the speaker means. Happe demonstrates that the understanding of metaphors requires an understanding of first-order intentions (what the speakers say). Understanding irony is more complex and demanding and cannot occur unless one is able to understand second-order intentions (what the speakers intend as opposed to what they say). The comprehension of verbal irony is related to second-order mental state inferencing ability and poses different cognitive demands than the comprehension of metaphor (Colston and Gibbs, 2002; Happe, 1993, 1995; Pexman et al., 2006; Winner et al., 1998). Numerous studies show that there are multiple cues to ironic intent: intonation, contextual incongruity, facial expression (Colston, 2002; Ivanko and Pexman, 2003; Pexman et al., 2006; Utsumi, 2000). Although irony is context sensitive, it may also be comprehended in impoverished contextual settings. As demonstrated by multiple empirical studies, ironic intents can be interpreted as long as context and utterance incongruity holds (e.g. Colston and O'Brien, 2000; Colston, 2002; Ivanko and Pexman, 2003).

Figurative language has been noted to cause special difficulties to L2 learners in culture and context related areas of pragmalinguistic and sociopragmatic acquisition, comprehension and production. Littlemore and Low (2006: 6) distinguish three possible causes of foreign language learners' difficulties with figurative language. Foreign language users may not know the conventions governing usage of figurative language. They may lack cultural connotations that need to be activated in order to grasp the figurative meaning. Foreign language learners may approach figurative chunks trying to understand each word separately. Lazar (1996) observes that understanding figurative language involves

a process of inference. Glucksberg (2001) emphasizes the importance of culture-specific knowledge involved in figurative competence since figurative meanings are deeply embedded in culture – cultural and social values are tacitly governing perception and interpretation of discourse. For Glucksberg, this culture-bound dependence of communicative competence constitutes a serious problem for foreign and second language learners, since they must not only become bilingual but also bicultural to become fully proficient in figurative language.

The comprehension of idioms and metaphors is the most researched aspect of L2 figurative competence. Cieślicka (this volume), interpreting the results of an experimental study on how second language learners process idioms, emphasizes the special status of literal meaning in the course of L2 processing. As her study shows, even highly advanced foreign language learners tend to rely overly on literal meaning, which very often leads to misinterpretation of idiomatic meaning. Littlemore and Low (2006: 49) see the major difficulty in foreign language learners' comprehension of metaphors and arriving at the speaker's intended meaning in the potential incompatibility of shared speaker-hearer knowledge. Littlemore and Low (2006: 54), as well as Giora (2003: 7ff), emphasize that in interpreting metaphors a foreign language user may pay attention to different features of the context than native speakers of the language do. Moreover, the native and non-native networks of contextually salient features may not overlap, leading to different, often discrepant, interpretations of metaphors by native and foreign language users. Littlemore (2003), discussing the influence of cultural background on metaphor interpretation, emphasizes the role of culture-specific assumptions about the target domains that affect the student's recognition of the target language speakers' attitudes. Littlemore (2001), examining the concept of figurative competence and its connection to L2 learning, demonstrates the complexity of this competence. She emphasizes the role individual differences, such as cognitive style or L2 communicative ability, play in the process of acquiring L2 figurative competence. Littlemore (2008), discussing L2 learners' conceptual and linguistic competence in comprehending and interpreting metaphor and metonymy in specific discourse communities, notices that merely teaching figurative expressions may not be sufficient. In order to equip foreign language learners with native-like figurative competence, it seems necessary to help them develop useful interpretation strategies – by and large these will involve making use of contextual cues and, what she calls, 'figurative thinking'. Littlemore's focus on figurative thinking and interpretation strategies is very much in line with the inferential approach espoused in this paper.

Individual differences in figurative language comprehension are also emphasized by Sperber and Wilson (1986: 16) who observe that beyond the species-specific cognitive abilities and a number of culture-specific experiences, views and attitudes, individuals tend to differ in terms of features of cognition and perception, individual experiences and memories, which leads to differences in individual cognitive frameworks of world representation and communication interpretation. Since the comprehension of figurative language is generally subject to individual variation, there should be differences between L2 users and monolinguals, as the former have not only the experience of different cultures but also the experience of communicating in different contexts and different individual experiences of linguistic communication in general.

4. Conversational implicature

Grice's theory of non-natural meaning (1968) and conversational implicature theory (1975) have promoted an inference-based approach to the interpretation of conversational meaning of utterances, and especially the nonliteral meaning. Grice's greatest contribution to the linguistic theories of meaning was the recognition of the primary role conversational inferencing plays in communication. According to Grice all human communication is governed by the search for the speaker-intended meaning. Grice's preoccupation with intentions and the necessity to arrive at the speaker-intended meaning led him to believe that attribution of the speaker's intention is possible not only because of features of the linguistic code but, most of all, because of the inferential abilities that humans ordinarily use. This intentionality-driven inferential approach to communication resulted from Grice's conviction that all nonliteral communication becomes accessible in the intended by the speaker sense only and always through inferencing. This inference-based approach to communication in general, and nonliteral communication specifically, offers a single comprehensive inference-based research perspective on different tropes.

After over three decades of research, initiated by Grice (1975), it is generally agreed that comprehending verbal communication people do not respond to the literal content of a figurative utterance, but rather to its intended, implied import. In order to derive meaningful information in communicative encounters, the hearer needs to inferentially enrich the linguistic structures contained in an utterance. Sperber and Wilson observe that in verbal communication we are interested in the meaning of the sentence uttered only in so far as it provides evidence to what the speaker means: "Communication is successful not when hearers recognize the linguistic meaning of the utterance, but when they infer the speaker's meaning from it" (1986: 23). Grice (1975) was the first to claim that meaning in communication is not derived solely from decoding linguistic meaning, but the intended meaning has to be inferred on the basis of linguistic meaning. The meaning inferred, i.e. derived through inferencing, not through decoding, is what Grice dubbed conversational implicature. Conversational implicature is for Grice a special type of non-natural, i.e. intended, meaning, as well as a means of deriving this intended meaning.

The Gricean approach to communication has been extensively modified by Sperber and Wilson (1986/1995) and Sperber (1995, 2000) in terms of cognitive conditions of comprehension. According to Sperber and Wilson (1986/1995), in order to comprehend the implied meaning, a discourse participant needs to know the language of communication, the conversational context and inferencing rules. Additionally, assumptions the speaker and the hearer hold and contribute to a conversational exchange play a significant role either facilitating or hindering the working out of the intended meaning. Assumptions are defined by Sperber and Wilson (1986: 85) as structured sets of concepts, which are part of our representation of the world. A representation of the world may then be regarded as a stock of assumptions that we all hold and apply to construct or reconstruct our most current representation of the world.

Sperber and Wilson (1986: 15) define the context as "a psychological construct, a subset of hearer's assumptions about the world, that affect the interpretation of the utterance." According to Sperber and Wilson (1986: 16), "members of the same linguistic community

converge on the same language, plausible that they converge on the same inferential abilities, but the same is not true of their assumptions about the world." As they observe, inferential abilities stabilize after a learning period and remain unchanged from one utterance or inference to the next. However, each new experience expands the range of assumptions that one adds to construct the context of interpretation. While individuals tend to be highly idiosyncratic in their personal interpretations of the world, members of the same cultural group share a wide range of assumptions, experiences and views that might predispose them to certain unanimity in arriving at inferential meanings conversationally implied in discourse. On the other hand, cognitive characteristics, such as memory, perception or attention, vary from one individual to another, so inferential abilities have to vary as well to a certain degree among members of the same cultural group and, all the more, for people who speak different languages, have mastered different concepts, and construct different representations of the world. Hence, hearers from different cultural groups than the speaker are very likely to differ in their communicative inferences from hearers from the same cultural group.

Grice's intentionality-driven approach (1968, 1975, 1989) to inference-based conversation allows for some further distinctions with regard to the literal and figurative meanings. In his inferential approach to communication driven by intention recognition, Grice provided some further criteria for distinguishing between literal and figurative meanings and introduced his own taxonomy to discuss figurative meaning. For Grice, all utterance meaning that is intended and implied, but not literally or explicitly expressed, must be inferentially derived; all the meaning that is meant but not said is an implicature. Grice's criteria for deriving figurative, intended senses include:

1. cancellability (defeasibility) – speakers may cancel a figurative meaning attributed to them by adding some additional premises or denying the proposition, saying that they didn't mean it;
2. non-detachability – figurative senses are attached to the semantic content of what is said and not to the linguistic form;
3. calculability – language users are capable of calculating what the speaker intended to convey;
4. non-conventionality – implicatures are non-conventional, i.e. they are not a part of the conventional meaning of linguistic expressions.

Levinson (1983: 120) extended Grice's original implicature test-list to include also:

5. reinforceability, which means that implicatures can be conjoined with an overt statement of their content without a sense of anomalous redundancy;
6. universality, which results from the expectation of cooperation and means that in every language, language users tend to interpret figurative utterances looking for the implied meanings.

Grice's taxonomy, motivated by the comprehension of figurative language in communication, is explicitly dedicated to the intended-meaning-driven paradigm. His approach to figurative language is through the speaker's recognition of intention-behind-a-figurative-expression, rather than through recognition of a figurative genre.

Grice's (1975/1989: 26) typology of conversational implicatures hinges on the cooperative principle ("Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange"), accompanied by four conversational maxims (quality; quantity; relevance; manner) and ensuing submaxims, which together guide the interlocutors to understanding not only what the communicator said, but also what they meant. These maxims specify how communicators should converse to be efficient and cooperative: they should speak sincerely, relevantly and clearly, while providing sufficient information. According to Grice, whether interlocutors follow the maxims or violate them, intended meaning is sought through the assumption of cooperation at some deeper level. The inferences that arise to preserve the assumption of cooperation are implicatures. Speakers in everyday communication often chose not to follow the maxims for some communicative purposes. Such exploitations of the maxims give rise to many of the traditional 'figures of speech', or tropes. Metaphors, idioms and ironies in Grice's theory of conversational implicature result from the exploitation of the maxim of quality (say the truth), since their literal meanings are blatantly false. The use of quantifiers (e.g. *some*, *(not) all*), forming a scale of informativeness, indicates that the speaker chose a relatively weak form (*some*) not to articulate a more informative term from the same scale (*all*). Given the prominent role of scales in everyday discourse, where scales range from less to more informative, this kind of inference has been dubbed scalar implicature. Scalar implicatures and understated criticism are cases of exploitation of the maxim of quantity (be as informative as is required). Flouting the maxim of relation (be relevant) leads to conversational indirectness, forcing the hearer to make bridging inferences to preserve the assumption that what the speaker says is relevant. The maxim of manner (be perspicuous) is flouted when a speaker chooses to be prolix instead of brief to piggyback some additional meaning on top of the literal one.

Pragmatic comprehension in SLA with respect to implied meanings has been researched by Carrell (1981, 1984), Kasper (1984), Bouton (1988, 1992, 1994, 1996), Takahashi and Roitblat (1994), Kubota (1995), Lee (2002) and Taguchi (2005).

Bouton (1988, 1992, 1994, 1996) was the first, in SLA paradigm, to demonstrate in a longitudinal study the ubiquity and importance of conversational implicature as a communicative tool which, when neglected in foreign/second language pedagogy, leads to communicative failure. Bouton conducted a series of studies in which he tested whether foreign students, non-native speakers of English, derive the same meaning from conversational implicatures as native speakers of English do. In his cross-cultural study (1988), he found that the ability of NNSs to interpret implicatures in English varied. In his longitudinal study, Bouton (1994) observed that NNSs can become as proficient as NSs in their ability to interpret conversational implicatures when they have had ample communicative experience in the target language country. Bouton (1996) found that the more explicit the instruction language learners get, the better the results they achieve in comprehending and interpreting indirectly conveyed figurative meanings. In his 1996 study, Bouton singled out a pool of "The Tough 10" implicatures, most troublesome for NNSs, that remain systematically difficult in terms of comprehension of the intended figurative meaning. The items marked off as especially and consistently difficult covered implicatures based on violation of the four conversational maxims (Grice, 1975) that resulted in

indirectness, understated criticism, metaphor, critical irony (sarcasm) and praising irony (banter), and other.

Lee (2002), attempting to find out to what extent Korean NNSs of English are able to interpret conversational implicatures in the manner NSs of English do, found that L2 learners find it more difficult to interpret implicatures that are sensitive to cultural context and/or rely on suprasegmental features (e.g. intonation, tone of voice). She interprets the results dividing implicatures (after Grice, 1975) into generalized, i.e. less context dependent, and particularized, i.e. more context dependent. Within the set of particularized conversational implicatures she found statistically significant differences in the response patterns of native and non-native speakers, while no difference was found between native and non-native speakers in their interpretations of generalized implicatures. Lee's results are consistent with Sperber and Wilson's (1986) observations concerning the role of individual, idiosyncratic assumptions about the world and communication, and assumptions shared by members of one culture, or community of usage.

Taguchi (2005) investigated whether L2 proficiency affects the ability to comprehend implied meaning in spoken dialogs in terms of accuracy and speed of comprehension. Taguchi's results support claims about conventionality in language comprehension as playing a substantial role and suggest differential comprehension mechanisms for the two types of implied meaning (more conventionalized versus less conventionalized) in EFL learning.

5. The present study

5.1 The aim of the study

Littlemore and Low (2006: xvi) observe that figurative language research in foreign language learners is minimal as compared with that for other areas of foreign language acquisition. They also note that no second language acquisition research involves figures other than metaphor. This has motivated us to investigate into a wider spectrum of figurative language comprehension in L2 users. Moreover, no study has so far examined the L1 figurative competence of L2 users. Therefore, in this study we want to investigate the comprehension of figurative language in multicompetence, that is both the development of L2 figurative competence in advanced L2 users and the impact of second language knowledge on L1 figurative competence.

In order to investigate the ability to comprehend figurative language in a wide range of its communicative applications and diversity of patterns and modes of figurative language presence in everyday discourse, we chose to study conversational implicature, as it allows us to cover this diversity.

The following research hypotheses can be formulated with regard to comprehension of conversational implicatures by proficient L2 users on the basis of relevant previous research:

1. Both monolingual and bilingual NSs of Polish will tend to derive the intended meaning aiming at the conversationally implied rather than literally expressed meanings.

2. L1 Polish figurative competence of L2 users will diverge from the L1 Polish figurative competence of monolingual language users in terms of conversational implicature response patterns.
3. L2 figurative competence of L2 English users will lag behind their L1 Polish figurative competence.

5.2 The subjects

A total of 75 subjects participated in the study. The respondents were divided into two groups: L2 users (bilinguals) and L2 learners (monolinguals). The L2 users ($N = 28$; 6 male, 22 female; mean age: 22.14) were students at Adam Mickiewicz University in Poznań, specializing in English studies. Since a vast majority of the courses they take are taught in English, they use English for a real-life purpose every day throughout the academic year, reading, attending classes, preparing homework assignments, and taking examinations in English. All the data were collected towards the end of their final third year at the university, by which time they were supposed to reach a level of English comparable to that required to pass the Cambridge Proficiency Examination. The L2 learners ($N = 47$; 10 male, 37 female; mean age: 21.91) were students at Poznań University of Economics, specializing in commodity science. All the subjects in this group had completed a course in English as a second language at the university and their level was judged by their instructors as pre-intermediate. Polish is the language of instruction in all their courses.

Additionally, data from 9 native speakers of English (5 British, 3 American, 1 Australian), who were at the time employed as teachers at the School of English at Adam Mickiewicz University, were collected for comparison.

5.3 The instrument and procedure

The research tool in this study was an implicature test – an off-line, multiple choice instrument for examining the comprehension of conversational implicatures. The implicature test used in this study was a moderated version of an instrument originally designed by Bouton (1988, revised in 1994). The test consisted of 16 items. Each item on the test was composed of a brief description of a situation with a short dialog (2 to 4 lines, which means 1 or 2 exchanges), in which one of the utterances involved the use of implicature. Care was taken to feature canonical discourse situations, strongly constraining the type of relationship between interlocutors and their status (e.g. teacher – student; friend – friend), as well as the goal of the communicative exchange and the expectations of the interactants. The dialog was followed by four possible interpretations of the implicature and a blank to fill in an alternative interpretation if the respondent did not agree with any of the choices provided. The four possible options offered as implicature interpretation were constructed in such a way as to provide only one correct (intended meaning) interpretation of the implicature, and the three remaining options of implicature interpretation were either literal interpretations, or false, because not licensed by the implicature trigger. They were either under or overstatements of the target sentence carrying implicature. An example of

implicature test item featuring an exchange between two professors and possible interpretations of the final, target statement below:

Two teachers are talking about a student's term paper.
Professor Davis: Have you read Cobb's term paper on cross-cultural influence?
Professor Dyke: Yes, I read it last night.
Professor Davis: What did you think of it?
Professor Dyke: *I thought it was well typed.*

How did professor Dyke like Cobb's term paper?
a. He liked the paper. He thought it was good.
b. He thought it was certainly well typed.
c. He didn't like it.
d. He admired the form of the paper.
e. other ...

The test was prepared in two language versions – Polish and English. The situations, dialogues, and multiple choice interpretative options were translated as faithfully as possible to ensure comparability and to make sure that in both language versions the inference triggering mechanism was the same. The order in which the multiple choice options were presented was different in the two language versions.

The 16 items representing 7 different types of implicature were modelled on Bouton's (1996) pool of "The Tough 10" implicatures most troublesome for NNSs that remain systematically difficult in terms of comprehension of the intended figurative meaning. The items covered implicatures based on violation of the conversational maxims that resulted in indirectness, understated criticism, metaphor, critical irony (sarcasm) and praising irony (banter), and other (for a complete list of items, see Table 1).

Table 1. List of items and types of implicature

	Maxim	Type of implicature		Maxim	Type of implicature
1	quality	idiom	9	quality	praising irony (banter)
2	quality	metaphor	10	relevance	indirectness
3	quantity	understated criticism	11	quality	critical irony (sarcasm)
4	quantity	scalar (quantifier scale)	12	manner	be brief
5	quantity	understated criticism	13	relevance	indirectness
6	quality	critical irony (sarcasm)	14	quantity	scalar (numerical scale)
7	quality	idiom	15	quality	metaphor
8	relevance	indirectness	16	quality	praising irony (banter)

The multiple-choice test included three tokens of indirectness, two tokens of set idiomatic expression, two tokens of metaphor, two tokens of understated criticism, two tokens of scalar implicature, one token of implicature of manner, and four tokens of irony (sarcasm and banter) embedded in dialogic exchanges. Each dialogue contained a target sentence, at the end, which triggered an inference. The multiple choice options that followed were aimed at probing whether a respondent managed to get the figurative, intended meaning, or got stuck with the literal meaning. In two cases (situation 2 – metaphor, and 8 – relevance) the respondents could choose between two types of non-literal meanings: generalized and particularized.

The participants were asked to choose the interpretation that most closely approximated what the speaker meant by what they said. The L2 users completed both a Polish and an English version of the implicature test. The completion of each language version was separated by a break. The L2 learners completed only the Polish version. Each participant was given approximately 20 minutes to complete the test.

5.4 Results

Table 2 shows how Polish monolinguals and bilinguals interpret conversational implicatures in L1. *T*-tests for independent samples were carried out to compare the responses of the two groups on each test item.

Table 2. Mean number of instances of correct identification of the intended meaning in L1 Polish by L2 learners and L2 users

	Implicature type	L2 learners	L2 users		
1	Quality – idiom	.83	.82	n.s.	
2	Quality – metaphor (generalized)	.72	.79	n.s.	
	Quality – metaphor (particularized)	.19	.14	n.s.	
3	Quantity – understated criticism	.74	.96	$t = 2.32$	$p < .05$
4	Quantity – scalar	.21	.00	$t = 2.71$	$p < .01$
5	Quantity – understated criticism	.51	.79	$t = 2.24$	$p < .05$
6	Quality – irony (sarcasm)	.77	.89	n.s.	
7	Quality – idiom	.98	1.00	n.s.	
8	Relevance (generalized)	.11	.25	n.s.	
	Relevance (particularized)	.74	.54	n.s.	
9	Quality – irony (banter)	.77	.54	n.s.	
10	Relevance – indirectness	.96	.93	n.s.	
11	Quality – irony (sarcasm)	.74	.54	n.s.	
12	Manner	1.00	.93	n.s.	
13	Relevance – indirectness	.57	.57	n.s.	
14	Quantity – scalar	.74	.68	n.s.	
15	Quality – metaphor	.81	.82	n.s.	
16	Quality – irony (banter)	.79	.79	n.s.	

The results show that the two groups interpret the intended meaning in a similar way. However, both the bilingual and monolingual native speakers of Polish are rarely unanimous in their interpretations of the intended meanings, rarely reaching 100% consistency in their judgements of the speaker's intentions (situation 12 in the monolingual group, and situation 7 in the bilingual group).

No inter-group differences have been observed in interpreting idiom and metaphor. In the case of idiom, over 80% of respondents in both groups pointed to the figurative meaning as the intended one in both situations (1 and 7). In the case of metaphor (situations 2 and 15), over 70% of all respondents pointed to the figurative meaning.

Intended irony (situations 6, 9, 11, 16) was somewhat less well understood, with only half of the L2 users decoding the utterances as ironic in two situations out of four (situation 9 and 11), but the inter-group differences are not statistically significant.

Similarly, no differences have been observed between the groups in the interpretation of indirectness resulting from the exploitation of the maxim of relation (be relevant) and the maxim of manner (be brief) in situations 8, 10, 12, 13.

Statistically significant differences between the groups appeared in the interpretation of understated criticism and scalar implicatures. In both situations (3 and 5) containing understated criticism the bilinguals have shown more sensitivity towards the speaker's intention than the monolinguals.

Scalar implicatures interpretations provided unexpected results. In case of numerical scale (cardinal numbers) no statistical differences were observed for both groups of respondents (situation 14). The implicature employing a quantifier scale (*all/some/not all*) (situation 4) proved more difficult to interpret. In case of quantifier scale, *not all* (*nie wszyscy*) is considered to be the pragmatic meaning of *some* (*niektórzy*) after Noveck (2001). The respondents interpreted the utterance containing *some* in L1 in three ways: meaning *hardly any* (*mało kto*) (43% of the monolinguals, 50% of the bilinguals), *not many* (*niewielu*) (28% of the monolinguals, 43% of the bilinguals), or *not all* (*nie wszyscy*) (21% of the monolinguals, 0% of the bilinguals). The majority of the monolinguals and almost all the bilinguals interpreted the utterance containing *some* as an understatement, actually overinterpreting the intended meaning. As it is highly unlikely that the subjects would not know the meaning of *some* in their L1, it is very probable that such an interpretation was induced by the context of the situation, in which an instructor informs students that some of them have passed an exam.

Table 3 shows how L1 Polish L2 English bilinguals interpret conversational implicatures in L1 and L2. *T*-tests for dependent samples were carried out to compare the L2 users' responses in the two languages.

The results show numerous differences in the interpretation of implicatures in L1 and L2. The subjects interpret figurative language differently in L1 and in L2 in 8 out of the 16 test items. However, the differences not always go in the expected direction. In three situations (5 – understated criticism, 9 – irony, 14 – scalar implicature), the L2 users show weaker comprehension of figurative language in L2 than in L1, as expected. In four situations (1 – idiom, 4 – scalar implicature, 13 – indirectness, 15 – metaphor), their comprehension of figurative language, in terms of deriving the intended meaning, is actually better in L2 than in L1.

The analysis of the results by type of implicature does not show any clear preference for any particular kind of implicature in either language. Idioms are understood equally well in both languages (situation 7), or better in L2 than in L1 (situation 1). Metaphors are either understood better in L2 than in L1 (situation 15), or simply differently in different languages (situation 2). Irony is interpreted similarly in both languages (situation 6, 11, 16), or better in L1 than in L2 (situation 9). Understated criticism shows no difference between the languages (situation 3), or is less well understood in L2 (situation 5). Indirectness is interpreted similarly in both languages (situation 8, 10), or better in L2 than in L1 (situation 13). Scalar implicatures are understood better either in L1 (situation 14) or in L2 (situation 4).

Table 3. Mean number of instances of correct identification of the intended meaning in L1 Polish and L2 English by L2 users

	Implicature type	L1	L2		
1	Quality – idiom	.82	.96	$t = 2.12$	$p < .05$
2	Quality – metaphor (generalized)	.79	.21	$t = 6.00$	$p < .000005$
	Quality – metaphor (particularized)	.14	.71	$t = 6.00$	$p < .000005$
3	Quantity – understated criticism	.96	.86	n.s.	
4	Quantity – scalar	.00	.21	$t = 2.71$	$p < .05$
5	Quantity – understated criticism	.79	.39	$t = 4.18$	$p < .0005$
6	Quality – irony (sarcasm)	.89	.89	n.s.	
7	Quality – idiom	1.00	1.00	n.s.	
8	Relevance (generalized)	.25	.29	n.s.	
	Relevance (particularized)	.54	.43	n.s.	
9	Quality – irony (banter)	.54	.32	$t = 2.71$	$p < .05$
10	Relevance – indirectness	.93	.96	n.s.	
11	Quality – irony (sarcasm)	.54	.61	n.s.	
12	Manner	.93	.89	n.s.	
13	Relevance – indirectness	.57	.79	$t = 2.71$	$p < .05$
14	Quantity – scalar	.68	.36	$t = 3.58$	$p < .005$
15	Quality – metaphor	.82	1.00	$t = 2.42$	$p < .05$
16	Quality – irony (banter)	.79	.82	n.s.	

As in the previous analysis, scalar implicature employing quantifier scale (*all/some/not all*) (situation 4) showed an untypical pattern of responses. The bilingual subjects interpreted the implicature differently in their two languages. In L1 no one pointed to the canonical meaning of *some* (*niektórzy*), that is *not all* (*nie wszyscy*), 43% chose *not many* (*niewielu*), 50% chose *hardly any* (*mało kto*). In L2, 21% pointed to *some*, 43% chose *not many*, 29% pointed to *hardly any*. While the majority of the L2 users interpreted the utterance as an understatement in both languages, they seem to be less prone to overinterpret the speaker's intention in L2 than in L1.

6. Discussion and conclusion

The data confirm the first research hypothesis: both monolinguals and bilinguals derive the intended meanings focusing on the conversationally implied rather than the literally expressed.

The second research hypothesis, concerning the L1 of L2 users, is partially confirmed. Both monolinguals and bilinguals show high rates of reliance on the figurative meanings in interpreting the speaker's intentions. Differences between L2 users and monolinguals in this respect are rare. In some situations, the bilinguals tend to be more sensitive towards the implied against the literal. The baseline data from native speakers of English are insufficient to state whether the differences between the L2 users and the other native speakers

of Polish are due to the influence of the L2. It seems that at least in one case (situation 3) L2 influence is likely, since the L2 users point to the figurative meaning as the intended one more often than the other subjects and the figurative meaning has been pointed out as the intended one in this particular situation by all the native speakers of English.

The third research hypothesis, maintaining that L2 figurative competence of L2 English users will lag behind their L1 Polish figurative competence, at least with regard to interpreting conversational implicatures, is not borne out by the data. In half of the cases in this study the L2 users interpreted conversational implicatures differently in their two languages. While the understanding of implicatures expressed in the L2 was sometimes weaker than in L1 (in 3 of 16 situations), this was not always the case. In 4 of the 16 situations the L2 users show better understanding of the implied meaning in L2 than in L1. The data permit to conclude that the understanding of the implied meanings is different in the second language of fully proficient L2 users, which does not mean that it needs to be worse than in L1. The differences in interpreting an implicature as generalized or particularized in situation 2 and different degrees to which the subjects overinterpret the scalar implicature in situation 4, depending on the language in which the implicatures are expressed, suggest that figurative language might be simply interpreted differently in the respective languages of an L2 user from the way it is interpreted by monolinguals.

The differences found in this study suggest that in some cases the L2 users make different assumptions about the speaker's intentions than other native speakers of their L1. They also make different assumptions about the speaker's intentions in the same situational context in their two languages. Further research is needed to see to what extent these assumptions are culturally specific and to what degree they are transferable between languages. The pattern of responses obtained in situation 4 (scalar implicature) suggests that sometimes the L2 users might also overdifferentiate between their two languages.

The results show that, if figurative competence is defined as "the ability to deal with figurative language" (Levorato, 1993: 104), we need to be very careful about operationalizing this ability before any definite statements are made about the figurative competence of advanced L2 users. The results of the present study show that some aspects of figurative competence might be better developed in L2 than in L1, while some others might indeed lag behind L1 figurative competence. Better understanding of L2 users' figurative competence will be possible if we come to understand better the linguistic and the cognitive mechanisms responsible for the processing of figurative language. If the figurative is defined as non-literal, then inferencing mechanisms are part and parcel of figurative language processing. Although inferencing mechanisms are deeply interrelated with linguistic abilities and skills, they do not depend on them entirely – "Linguistic comprehension is an inferential task using decoded material as evidence" (Sperber, 2000: 122).

The understanding of implied, non-literal, meanings depends on the situational context. Figurative competence of L2 users is built by different experiences of interaction in different cultural and situational contexts. Different patterns of social interaction develop in different situational contexts, both crossculturally and within the same culture. While inferencing mechanisms, involved in understanding conversational implicatures, belong to universal cognitive processes (Sperber, 1995), the social and situational context of an interaction provides a frame within which certain inferences become legitimate and

others do not. Figurative competence of L2 users differs from the figurative competence of monolinguals because the L2 user has a qualitatively different experience of communication in general. The comprehension of conversational implicatures involves both decoding the literal meaning of an utterance, which depends on the knowledge of the linguistic code, and inferring the implied, which relies on universal inferencing processes operating within the limits of a given sociocultural context. When L2 users know the linguistic code very well, the inferences they make will still differ as their experience of the sociocultural contexts and the totality of experience at communicating with other people is different. This view of bilingual figurative competence is consistent with Cook's (1992, 2002; see also Ewert, forthcoming) view of multicompetence as two or more systems integrated at the linguistic and cognitive level and, in particular, with the reconceptualized multicompetence proposal of Hall, Cheng and Carlson (2006), which emphasizes the role of L2 users' pragmatic action in different sociocultural contexts.

The results of our study show different levels of comprehension in the L2 users' two languages for different types of figurative language. However, since our focus has been on figurativeness in general, no definite conclusion should be made here as to any regularities concerning the interpretation of different tropes in any particular language. The reason for this is that, endeavouring to encompass a full spectrum of figurative cases, we had to limit the number of stimuli for each examined category. The study shows that bilingual figurative competence is a complex entity and more research is needed to explore its various aspects. In particular, future research should focus on the interpretation of different kinds of figurative language in both languages of the L2 user.

Acknowledgement

The research reported in this chapter was supported in part by grant N104 040 32/3713 from the Polish Ministry of Science and Higher Education to Anna Ewert.

References

Ameel, Eef, Gert Storms, Barbara C. Malt & Steven A. Sloman (2005). How bilinguals solve the naming problem. *Journal of Memory and Language, 53*, 60–80.

Athanasopoulos, Panos (2006). Effects of the grammatical representation of number on cognition in bilinguals. *Bilingualism: Language and Cognition 9*, 89–96.

Bialystok, Ellen (2001). *Bilingualism in development. Language, literacy, and cognition.* Cambridge: Cambridge University Press.

Blum-Kulka, Shoshana (1991). Interlanguage pragmatics: The case of requests. In R. Philipson, E. Kellerman, L. Selinker, M. Sharwood Smith, & M. Swain (Eds.), *Foreign/second language pedagogy research* (255–272). Clevedon: Multilingual Matters.

Bouton, Lawrence F. (1988). A cross-cultural study of ability to interpret implicatures in English. *World Englishes, 7,* 183–196.

Bouton, Lawrence F. (1992). *The interpretation of implicature in English by NNS: Does it come automatically – Without being explicitly taught?* [Pragmatics and Language Learning Monograph Series, 3]. Urbana-Champaign, IL: University of Illinois.

Bouton, Lawrence F. (1994). Conversational implicature in a second language: Learned slowly when not deliberately taught. *Journal of Pragmatics, 22,* 157–167.

Bouton, Lawrence F. (1996). *Pragmatics and language learning.* [Pragmatics and Language Learning Monograph Series, 7]. Urbana-Champaign, IL: University of Illinois.

Carrell, Patricia L. (1984). Inferencing in ESL: Presuppositions and implications of factive and implicative predicates. *Language Learning, 34*, 1–21.

Carrell, Patricia L. & Beverly H. Konneker (1981). Politeness: Comparing native and non-native judgements. *Language Learning, 31,* 17–31.

Cenoz, Jasone (2003). The Intercultural Style Hypothesis: L1 and L2 interaction in requesting behavior. In V. Cook (Ed.), *Portraits of the L2 user* (62–80). Clevedon: Multilingual Matters.

Cieślicka, Anna (2004). *On processing figurative language: Towards a model of idiom comprehension in foreign language learners.* Poznań: Motivex.

Cieślicka, Anna (this volume). Formulaic language in L2 – Storage, retrieval, and production of idioms by second language learners.

Colome, Angels (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language, 45,* 721–736.

Colston, Herbert, L. (2002). Contrast and assimilation in verbal irony. *Journal of Pragmatics, 34,* 111–142.

Colston, Herbert, L. & Raymond W. Gibbs (2002). Are irony and metaphor understood differently? *Metaphor and Symbol, 17,* 57–80.

Colston, Herbert, L. & Jennifer O'Brien (2000). Contrast of kind versus contrast of magnitude: The pragmatic accomplishments of irony and hyperbole. *Discourse Processes, 30* (2), 179–199.

Cook, Vivian J. (1991). The poverty-of-the-stimulus argument and multicompetence. *Second Language Research, 7,* 103–117.

Cook, Vivian J. (1992). Evidence for multicompetence. *Language Learning 42,* 557–591.

Cook, Vivian (2002). Background to the L2 user. In V. Cook (Ed.), *Portraits of the L2 user* (1–28). Clevedon: Multilingual Matters.

Costa, Albert, Mikel Santesteban & Agnes Cano (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language, 94,* 94–103.

De Bot, Kees, Wander Lowie & Marjolijn Verspoor (2005). *Second language acquisition. An advanced resource book.* London: Routledge.

De Groot, Annette, Joan Kaplan, Elizabeth Rosenblatt, Shelly Dews & Ellen Winner (1995). Understanding versus discriminating non-literal utterances: Evidence for a dissociation. *Metaphor and Symbolic Activity, 10,* 255–273.

Ewert, Anna (forthcoming). *L2 users' L1.* Bristol: Multilingual Matters.

Genesee, Fred, G. Richard Tucker & Wallace E. Lambert (1975). Communication skills of bilingual children. *Child Development, 46,* 1010–1014.

Gibbs, Raymond (1986). On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General, 115* (1), 3–15.

Gibbs, Raymond (1994). *The poetics of mind: Figurative thought, language, and understanding.* Cambridge: Cambridge University Press.

Gibbs, Raymond (2001). Evaluating contemporary models of figurative language understanding. *Metaphor and Symbol, 16,* 317–333.

Gibbs, Raymond (2002). A new look at literal meaning in understanding what is said and implicated. *Journal of Pragmatics, 34,* 457–486.

Giora, Rachel (1995). On irony and negation. *Discourse Processes*, 19, 239–264.

Giora, Rachel (1997). Understanding figurative and literal language: The graded salience hypothesis. *Humour, 12,* 425–436.

Giora, Rachel (2002). Literal vs. figurative language: Different or equal? *Journal of Pragmatics, 34,* 487–506.

Giora, Rachel (2003). *On our mind: Salience, context, and figurative language.* Oxford: Oxford University Press.

Giora, Rachel, Ofer Fein & Tamir Schwartz (1998). Irony: Graded salience and indirect negation. *Metaphor and Symbol, 13,* 83–101.

Glucksberg, Sam (2001). *Understanding figurative language: From metaphors to idioms*. New York: Oxford University Press.

Grice, Paul (1968). Utterer's meaning and intentions. *Philosophical Review, 78,* 147–177.

Grice, Paul (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (41–58). New York: Academic Press.

Grice, Paul (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Hall, Joan Kelly, An Cheng & Matthew T. Carlson (2006). Reconceptualizing multicompetence as a theory of language knowledge. *Applied Linguistics, 27,* 220–240.

Happe, Francesca G. E. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition, 48,* 101–119.

Happe, Francesca, G. E. (1995). Understanding minds and metaphors: Insights from the study of figurative language in autism. *Metaphor and Symbolic Activity, 10,* 275–295.

Herdina, Philip & Ulrike Jessner (2002). *A dynamic model of Multilingualism. Perspectives of change in psycholinguistics.* Clevedon: Multilingual Matters.

Ivanko, Stacey, L. & Penny M. Pexman (2003). Context incongruity and irony processing. *Discourse Processes, 35,* 241–279.

Kasper, Gabrielle. (1984). Pragmatic comprehension in learner-native speaker discourse. *Language learning, 34,* 1–20.

Kubota, Mikio (1995). Teachability of conversational implicatures to Japanese EFL learners. *IRLT, 5,* 35–67.

Lazar, George (1996). Using figurative language to expand students' vocabulary. *ELT Journal, 50,* 43–51.

Lee, Jin Sook (2002). Interpreting conversational implicatures: A study of Korean learners of English. *The Korea TESOL Journal, 5,* 1–25.

Levinson, Stephen, C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Levorato, Maria, C. (1993). The acquisition of idioms and the development of figurative competence. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (101–128). Hillsdale, NJ: Lawrence Erlbaum Associates.

Littlemore, Jeannette (2001). Metaphoric competence: A language learning strength of students with a holistic cognitive style. *TESOL Quarterly, 35,* 459–491.

Littlemore, Jeannette (2003). The effect of cultural background on metaphor interpretation. *Metaphor and Symbol, 18,* 273–288.

Littlemore, Jeannette (2008). 'He's got a bit of a loose nappy': Figurative thinking and foreign language learning: Metaphor and metonymy in institutional discourse. Plenary paper presented at the 33rd International LAUD Symposium, Landau/Pfalz, 10–13 March 2008.

Littlemore, Jeanette & Graham Low (2006). *Figurative thinking and foreign language learning*. New York: Palgrave Macmillan

Noveck, Ira (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition, 78,* 165–188.

Pavlenko, Aneta & Scott Jarvis (2002). Bidirectional transfer. *Applied Linguistics, 23,* 190–214.

Pexman, Penny, Melanie Glenwright, Suzanne Hala, Stacey Kowbel & Sara Jungen (2006). Children's use of trait information in understanding verbal irony. *Metaphor and Symbol, 21,* 39–60.

Searle, John (1979). Metaphor. In A. Ortony (Ed.), *Metaphor and thought*. (92–123). Cambridge: Cambridge University Press.

Sicola, Laura (2003). Communicative lingerings: An exploratory study of the emergence of 'foreign' communicative features in the interactions of American expatriates after re-entry. *Working Papers in Educational Linguistics, 18,* 27–53.

Sperber, Dan (1995). How do we communicate? In J. Brockman, K. Matson (Eds.), *How things are: A science toolkit for the mind* (191–199). New York: Morrow.

Sperber, Dan (2000). Metarepresentations in an evolutionary perspective. In D. Sperber (Ed.), *Metarepresentations. A multidisciplinary perspective* (117–138). Oxford: Oxford University Press.

Sperber, Dan & Deirdre Wilson (1986/1995). *Relevance: Communication and cognition*. Oxford: Blackwell.

Steinmann, Martin (1973). Figurative language and the two-code hypothesis. In R. W. Fasold & R. W. Shuy (Eds.), *Analyzing variation in language* (219–239). Washington, DC: Georgetown University Press.

Taguchi, Naoko (2005). Comprehending implied meanings in English as a foreign language. *Modern Language Journal*, *89,* 543–562.

Takahashi, Satomi & Herbert, L. Roitblat (1994). Comprehension process of second language indirect requests. *Applied Psycholinguistics, 15,* 475–506.

Utsumi, Akira (2000). Verbal irony as implicit display of ironic environment: Distinguishing ironic utterances from nonirony. *Journal of Pragmatics, 32,* 1777–1806.

Van Hell, Janet G. & Ton Dijkstra (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin and Review, 9*, 780–789.

Winner, Ellen (1988). *The point of words: Children's understanding of metaphor and irony*. Cambridge, MA: Harvard University Press.

Winner, Ellen (1995). Introduction: 10th anniversary special issue: Developmental perspectives on metaphor. *Metaphor and Symbolic Activity, 10,* 247–253.

Winner, Ellen, Hiram Brownell, Francesca Happe, Ari Blum & Donna Pincus (1998). Distinguishing lies from jokes: Theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain and Language, 62,* 89–106.

CHAPTER 17

Attention to phonological form

Modified output in task-based negotiated interaction*

Laura Sicola
University of Pennsylvania

1. Introduction

As Joan Morley succinctly put it, "intelligible pronunciation is an essential component of communicative competence" (Morley, 1991: 488). So how can "intelligible pronunciation" – or at least *more* intelligible pronunciation – be facilitated? Broadly, L2 phonological development involves a complex array of "[a]rticulatory, interactional, and cognitive processes [which entail] cognitive, psychomotor, linguistic, and interactive factors" (Pennington & Richards, 1986: 212). In particular, cognitive processes such as attention and noticing are as potentially influential in acquisition of L2 phonology as for other areas of language development, as evidenced by the fact that learners attend to pronunciation problems as much as – if not more than – lexical and morphosyntactic errors when they cause a breakdown in communication (e.g. Mackey, Gass & McDonough, 2000). Subsumed by these processing categories and incorporating all of the aforementioned factors are the issues of how learners attend to and modify phonological output. Although modified L2 lexical and morphosyntactic output has been a frequently discussed topic among researchers and practitioners alike, the inclusion of pronunciation in these discussions has been peripheral at best. Yet many researchers have called for further exploration of the cognitive processes involved in L2 phonological development (Ellis, 1999a; Hansen, 2004; Marinis, 2003; Morley, 1991; Pennington, 1998; Pennington & Richards, 1986; Williams, 2005). Fraser (forthcoming) also observes that cognitive linguistics contributes new, pedagogically applicable metalinguistic descriptions, but which are typically used when analyzing lexical and metaphorical reference, grammatical representation, etc. She recognizes an opportunity – and I argue, a need – 'to similarly apply them to L2 pronunciation.' As such, this study brings conditions and processes of the modification of L2 phonological output into focus through the lens of cognitive linguistics.

2. L2 phonological development

Can adults acquire a new phonological system? Among the many factors that influence the answer to this contentious question, "just one factor, *sensitivity to input*, is the most likely explanation for fossilization" (Long, 2003: 516, emphasis in original). Although decreased sensitivity over time is inevitable for adult non-native speakers (NNSs), it is not necessarily irrevocable. Markham asserts that

> due to habitual use of attention and production strategies over time for the L1, the speaker '*de-tunes*' the response characteristics of the perceptual and motor systems to new information or commands (e.g. Flege, 1995a; Pisoni, Lively & Logan, 1994). The systems can be '*re-tuned*' under certain circumstances (1997: 28; emphasis mine).

As such, helping NNSs resensitize to new phonetic contrasts will be essential, as they cannot be expected to independently "re-tune" their entire range of perceptual and motor skills, since "to detect and to produce the subphonemic details of L2 categories [during automatic processing] requires the inhibition of well established routines so that new ones can be established" (Schmidt, 2001: 30). A starting point might be to examine disparities in *intra*learner variation, as it provides evidence that learners have *demonstrated* the ability to produce certain L2 phonological forms more accurately than they *typically* produce. For example, there is usually a discrepancy between target form pronunciation on formal word-list read-aloud tasks as compared to pronunciation of the same forms when participating in informal interviews, due to the relative amount of attention allotted to language on each respective task, which directly influences form accuracy (Hansen, 2004; Major, 1987; Markham, 1997). In such a case, "confounding psychomotor components unique to pronunciation do not function as potentially impregnable physical barriers to the acquisition of that form"(Sicola, 2009: 16). Additionally, since learners use internal auditory and proprioceptive feedback mechanisms to self-monitor when speaking (Markham, 1997), learners who have even infrequently produced a form with targetlike accuracy should be more easily sensitized to the physical contrasts between targetlike and nontargetlike productions of the same form.

3. Attention, output and pedagogic tasks

Cognition-based theories and approaches to the study of SLA commonly address the role of NNS attention to form and how attention, with or without conscious noticing, can facilitate further processing of the form as a necessary step towards acquisition. As Schmidt affirmed, "SLA is largely driven by what learners pay attention to and notice in the target language input and what they understand the significance of noticed input to be" (2001: 3–4), with noticing being the key element for converting input to intake, thereby making it available for additional processing in working memory.

Interaction is a context in which the likelihood of such noticing and attention is increased. Through interaction, learners encounter a variety of opportunities to negotiate form and meaning. Such negotiations provide learners with multiple vehicles for noticing,

particularly the opportunity to receive modified comprehensible input, corrective feedback on their own utterances, and the opportunity to modify their own production (Pica, 1994). For the purposes of this study, the most important is the chance to modify their own production, i.e. to produce modified output. Ideally, output functions as a means to "force the learner to move from semantic processing to syntactic processing" (Swain, 1985: 249). Swain further explains the value of this cognitive shift within the context of interaction as follows:

> Negotiating meaning needs to incorporate the notion of being pushed toward delivery of a message that is not only conveyed, but that is conveyed precisely, coherently, and appropriately. Being 'pushed' in output…is a concept parallel to that of the i+1 of comprehensible input. Indeed one might call this the 'comprehensible output' hypothesis. (Swain, 1985: 249)

Ultimately, through output NNSs can test language-related hypotheses, notice holes in their interlanguage, and take advantage of opportunities to restructure their interlanguage by modifying nontargetlike utterances (Swain, 1985, 1998). Learners can choose to modify their output after receiving corrective feedback from an interlocutor (i.e. other-initiated modification), or after independently realizing one's own production errors (i.e. self-initiated modification), both of which are common learner strategies visible across tasks and interlocutors (Shehadeh, 1999). Self-correction without receiving corrective feedback is facilitated particularly for pronunciation through both auditory and proprioceptive feedback mechanisms (Markham, 1997). These modifications can assist learners in internalizing new forms as well as controlling already internalized forms (Swain & Lapkin, 1995).

As part of these negotiations, modified output can provide both the speaker and the hearer with the link to connect input, attention and output (Long, 1996). Interaction increases these opportunities for all areas of L2 development, including phonology, as R. Ellis (1999a: 235) explains regarding the revised interaction hypothesis (IH):

> The revised IH, with its emphasis on the role that interaction plays in stimulating the learners' attention to form, draws heavily on the current models of acquisition that view intake as a primary stage in the acquisitional process. The IH can be used to address all levels of language – pronunciation, lexis, grammar, and discourse.

In order to promote negotiated output modifications, pedagogic tasks have become commonplace in research and practice for lexical and morphosyntactic development, and can be equally facilitative in L2 phonological acquisition. First, information gap tasks (IGTs) should incorporate developmentally appropriate language forms which are complex and/or nonsalient and thus challenging for learners to acquire through more naturalistic means (Pica, Kang, & Sauro, 2006). In particular, IGTs that require a two-way exchange of uniquely-held, essential information to reach a single correct possible outcome have been argued to be the most likely task type to facilitate "modified interaction among participants and orient their attention to form, function and meaning" (Pica et al., 2006:301), and thus help learners notice target language forms.

It is not only important to notice target language forms in the interlocutor's input, but also nontargetlike forms in one' own production errors. As Shehadeh (1999) explained, "…when successful completion of the task depends primarily on the NNS

supplying accurate and comprehensible output…a higher proportion of extended negotiation routines is more likely to occur" (p. 659). The essentialness of a target form to successful task completion is at least partially related to the efficacy of a given task in stimulating this kind of negotiation (Loschky & Bley-Vroman, 1993; Spada, Lightbown, & White, 2005). When these negotiation routines result from the need to produce a more targetlike form – whether lexical, morphosyntactic or phonological – in order to achieve more accurate and comprehensible output, the learner is temporarily attending to the form in question. It is believed that noticing processes are stimulated under these conditions (e.g., Doughty & Williams, 1998; Ellis, 1999c; Long, 1996; Long & Robinson, 1998; Williams, 2005).

Although to date targeted, task-based interaction studies have focused primarily on morphosyntactic or lexical negotiations, such methods and findings should be equally applicable to the study of L2 phonological development. The acquisition of phonological components of a language – despite the neuromuscular element uniquely influencing pronunciation in contrast to other areas of language development – should also be heavily dependent upon whether or not they are noticed by learners, and whether or not the learners perceive the communicative value they hold in a given context. IGTs can serve as a tool to promote this type of noticing and attention, and ultimately resensitization.

As such, this study addressed the following research questions:

1. When working together on communicative pronunciation tasks, do NNSs draw each other's attention to a targeted phonological form in ways generally understood to facilitate SLA? I.e., do they modify their production of the target form?
2. If NNSs do modify their target form production, do the modifications result in more targetlike pronunciation?

4. Methodology

Participants

Three intermediate-level, elective pronunciation classes ($n^a = 16$; $n^b = 16$; $n^c = 15$) participated from a university level English as a second language (ESL) program. These classes were ideal participants for various reasons. From an affective perspective, it was believed that learners' motivation would be relatively high as all had chosen to enroll in the elective course. More importantly, it could also be presumed that there would be a random mix of pronunciation abilities within and across class groups since the university program does not assess oral production as a component in its placement testing, and there is no known correlation between L2 phonological accuracy and grammatical competence, fluency or overall proficiency. Table 1 below provides the L1 and gender distribution of each class.

The main priority in dyad selection was ensuring mixed-L1 combinations whenever possible, as research has indicated that this condition is one of the most ideal for SLA in learner-learner interaction. (Bitchener, 2004; Griggs, 2005; Jenkins, 2000; Long, 1996; Long & Porter, 1985). This factor has been identified as being more important than gender, which would have been impossible to balance given the gender mix within the

Table 1. L1 distribution in ESL pronunciation classes

L1	Class A, n = 16 (5 male, 11 female)	Class B, n = 16 (7 male, 9 female)	Class C, n = 15 (5 male, 10 female)
Japanese	2	9	5
Korean	9	4	2
Chinese (Taiwan)	2	3	5
Spanish	3	0	2
Russian & Chinese*	0	0	1

* This student was actually born in Turkey, but relocated internationally several times in childhood, adolescence and adulthood. Thus although Turkish was technically his "first" language, he considered his "primary" and dominant languages to be Russian and Chinese.

classes, regardless of intent. Once controlled for mixed-L1 dyads, other important factors such as level of fluency and overall phonological accuracy would likely be randomized, as indicated above.

Target form selection

Theta (/θ/) was chosen as the target form for several reasons. First, /θ/ is not found in the L1s of the participants, and all expressed an interest in learning it. Typically it is replaced in interlanguage with /s/, /t/ or /f/ (Archibald, 1998; Avery & Ehrlich, 1992; Swan & Smith, 1987). Although /θ/ is part of the phonemic inventory of standard Castilian Spanish varieties throughout most of Spain, speakers of those varieties associate it with the orthographic symbol 'z' or 'c' (if followed by 'i' or 'e'), and Spanish speakers in general typically pronounce the orthographic combination of 'th' in English words as [t] instead of [θ] (when 'th' is unvoiced). It is not aurally salient, and although all participants were able to produce it accurately as [θ] when asked to do so out of context and during simple minimal pair exercises in class, none incorporated it regularly in obligatory contexts in free conversation. As suggested above, this contrast between learners' typical accuracy at the most automatized/least conscious level and their accuracy when allocating more attentional resources to the target form in decontextualized (i.e. least automatized/most conscious) production opportunities showed a potential range for resensitization and learning to occur.

Map task

The map created for this two-way information gap task depicts a fictitious city, New Philadelphia (see Appendix A), and successful task completion requires a two-way request and provision of uniquely held information, which includes the target form, /θ/, and/or one of its contrast forms, /f/, /t/ and /s/, as described above (Sicola, 2009). The street layout is such that pairs of parallel streets have minimal-pair names and each intersection is

numbered. For example, Truth Alley and Truce Alley are next to each other and run in the same direction. At the top of the map are two lists of city landmarks, four each for Partner A and Partner B. Each partner must secretly assign random intersection numbers to each of the landmarks on her list. One person would give his interlocutor a random starting point, identified by cross-streets, and request directions to the first place on the interlocutor's list from that point. E.g., the receiver might begin by asking, "How can I get to City Hall from the intersection of Truce (or Truth) Alley and Faithful (or Fateful) Way?" The speaker would confirm the starting location as necessary (which was predictably common as there are four intersections that sound almost the same, phonetically, as indicated by the previous example request,) and then select an appropriate route along which to direct the listener. Upon arriving at the appropriate intersection, the speaker would ask the listener to confirm the number of the intersection at which he had arrived.

The map task also includes an instruction sheet (see Appendix B) which provides examples of permissible and non-permissible phrases for giving directions. This served two purposes. First, in case direction-giving in English was a new experience for some participants who were unfamiliar with typical language routines for this function, these examples would prevent the task from being too grammatically challenging (Gorsuch, 2001) and distract the learner from the phonological purpose of the task. Second, the examples demonstrate explicitly how to give directions without resorting to providing numbers, either referring to intersection numbers or number of blocks to go, etc. as destination confirmation was the only permissible (and required) use of numbers since such numerical references along the way would negate the need to say the names of the streets and thus enable students to avoid use of the minimal pair street names. These requirements are due to the fact that "in order to make structural accuracy essential to accurate communication of meaning, the designer of the task must exert a great deal of *control* over the discourse in the task" (Loschky & Bley-Vroman, 1993: 140). Additionally, multiple task demands and greater task complexity would be expected to increase the quantity and quality of negotiation (Robinson, 2001), but with most negotiated breakdowns centering around the *theta* and its contrast forms. Participants were told that they could ask each other for clarification as necessary.

Procedures

The task was incorporated as the activity used in the weekly class meeting at the language laboratory for each of the three participating classes. Interlocutors were paired with partners of different L1s whenever possible, and communicated with each other through headphones, but were physically unable to see each other. This forced them to pay strict attention to the accuracy of their aural perception and oral production without supplemental paralinguistic cues or interference such as gestures or lip-reading (Shehadeh, 1999). Target phonemes and contrast phonemes were clearly distinguishable through the headphones used.

Students first received the map and instruction sheet, and were told to assign random intersection numbers to the locations on their respective lists. Then the teacher or

researcher read each street name aloud to model targetlike pronunciation and allow students to repeat the name for practice and take any notes regarding pronunciation as they liked. This was intended to make students aware of any non-intuitive pronunciations particularly due to orthography and unfamiliar words which would negate the need for the target form contrasts to be distinguishable. For example, if the "w" were assigned sound in the word "Wreath" in "Coral Wreath Lane," where there is no "w" in the minimal pair "Reef," it would not matter if the word-final /θ/ were pronounced accurately for their listener to know which street was intended.

The teacher or researcher then read the instruction sheet aloud, clarifying each point as necessary, while the participants followed along and made notes on their own sheets. The teacher and the researcher (or teacher and a volunteer student, in the case when the researcher was the teacher in a class) modeled one complete example of how to request and provide directions from a starting point to one of the locations on the list. Once participants indicated comprehension of the task, the dyads recorded their performance on cassettes using the headsets and recording equipment at the language laboratory stations, and interactions were recorded and transcribed afterwards.

Data analysis

Negotiations pertaining to the target form and its contrasts in minimal-pair street names were coded for modified pronunciation, both self-initiated and following corrective feedback from an interlocutor. Specifically, modifications for *target-like use in obligatory contexts* were analyzed, including avoidance of oversupplying the target form, *theta*, in non-obligatory contexts (Pica, 1988). Only modifications in utterances in which the intended minimal pair option was contextually verifiable were coded (see Sicola, 2009, for an explanation of verification strategies), since, for example, it is unclear at first if a participant who says "[tru:s] alley" actually means "Truce Alley," as it appears, or if it is a nontargetlike production of "Truth Alley." Once the intended street name was verified, the modification could be classified as more (+) or less (–) targetlike.

5. Results

Data demonstrate substantial evidence of learners' ability to attend to and modify phonological form, helping themselves and each other develop greater productive and receptive accuracy in the process of L2 learning. The total compilation of all more-targetlike (+) and less-targetlike (–) modifications for verifiable street names, is in Table 2 below:

Most notably, the total proportion of more- to less-targetlike modifications is nearly 2:1, indicating that learners were almost twice as likely to modify their utterances with greater targetlike accuracy in obligatory contexts as they were to make nontargetlike modifications in such contexts. These instances of modified output seemed to be identifiable by interlocutors, transcribers and coders alike. On one extreme, dyad 26 had the greatest number of overall modifications at 46, with nearly equal numbers of more- and

Table 2. Verifiable modifications: More targetlike (+) and Less targetlike (–)

DYAD	(+)	(–)	Dyad Total
1	4	1	**5**
2	5	1	**6**
3	3	1	**4**
5	7	2	**9**
6	1	1	**2**
7	0	0	**0**
8	12	7	**19**
9	8	3	**11**
11	1	0	**1**
12	3	0	**3**
13	3	1	**4**
17	9	2	**11**
18	4	1	**5**
19	3	1	**4**
20	4	1	**5**
21	6	3	**9**
22	3	0	**3**
23	3	5	**8**
24	11	9	**20**
25	7	4	**11**
26	24	22	**46**
Total	**121**	**65**	**186**
%	**65.1%**	**34.9%**	**100.0%**

less-targetlike verifiable changes, in contrast to dyad 7, which was the only dyad with no verifiable modifications. It should be noted, however, that both dyads – as well as the rest – made many additional modifications in their output in contexts that were not verifiable, and thus could not be counted.

6. Discussion

Prior to interpreting the actual numbers, it is important to address the influence of NNS-NNS interactional dynamics on the map task performance. Primarily, it is challenging to determine if a breakdown in communication occurs because the speaker's utterance is nontargetlike (hereafter referred to as a speaker-based error), because the hearer misperceived the utterance (hereafter a hearer-based error), or perhaps both. To illustrate, the following excerpt illustrates both error types:

(1) 258. Dong: (laughs) Uhl on da intersection, Sink, Sink Street?
259. Li-Xin: Think? ([siŋk]? *unclear*)
260. Dong: Yeah Sink. *Think.* Think Street?
261. Li-Xin: Wait a minute (3)

262. Dong: It's a east. (2)
263. Li-Xin: Ah. Thing:::
264. Dong: Yeah, Think.
265. Li-Xin: *Sink.*
266. Dong: Yeah.
267. Li-Xin: Or Think?
268. Dong: *Think* (2)
269. Li-Xin: Sink
270. Dong: Yeah.
271. Li-Xin: Ahah.
272. Dong: T-H. ([ti::eitš]) (*spelling the letters*)
273. Li-Xin: Ah! Think. Think.
[ ]
274. Dong: Hah.
275. Dong: Think. Think. Think
[ ]
276. Li-Xin: Think Street.
277. Dong: A::n Sree Town Avenue
278. Li-Xin: Three Town, okay.
279. Dong: Yeah.

The dual nature of the miscommunication (hearer- and speaker-based) is visible in line 272 where Dong finally resorts to spelling to confirm his intent. This also is verification that his original utterance in line 258 was non-targetlike, and thus his modified output in his next turn, line 260, is a more-targetlike modification (+) and evidence of his attending to the form, having been pushed for clarification by his interlocutor in line 259. He also manages to maintain this targetlike accuracy without backsliding to "Sink" or making other less-targetlike modifications in the rest of the exchange. LiXin's vacillation back and forth between *think* and *sink* demonstrates her physical capability of producing both forms, /θ/ and /s/, but has less ability to perceive the distinction clearly. She confirms this by producing the target form accurately in line 276, "Think Street," (+). Her initial utterance in line 259 was unclear, and her uncertainty in what she was hearing ultimately resulted in two less-targetlike (–) and three more-targetlike (+) modifications. Yet by the end of the exchange, both interlocutors had converged on a targetlike form, rather than accepting and appropriating each other's original nontargetlike efforts. This phenomenon supports Ellis' observation that "[i]n exchanges between learners, one learner's modified output is another learner's comprehensible input" (1999b: 14).

Moreover, her increasingly targetlike control in obligatory contexts is further visible in line 278, "Three Town Avenue," despite Dong's nontargetlike production thereof in line 277, articulating "three" as [sri:]. Since the minimal pair alternative was *Free* Town Avenue, both interlocutors could have substituted [s] for /θ/ without resulting in misunderstanding.

Self-initiated output modification was also quite common, as seen in the above example. It was evidence that NNSs can identify their own errors through auditory or proprioceptive feedback and have the ability to modify their production without receiving external feedback (Markham, 1997; Swain, 1985).

Most interlocutors demonstrated targetlike production of the *theta* and its contrasting forms in obligatory contexts, specifically the street names, but had more problems distinguishing the forms aurally in the same contexts. The design of the task likely posed a strong influence on this finding, since there were always two street options, either one of which could have been correct at any point. As there were no other contextual clues that would have helped the listener decipher the intended street name, the essentialness of target form accuracy (Loschky & Bley-Vroman, 1993) was thereby maximized, which increased the listener's need to be absolutely certain about which option had been said.

Participants did learn to identify which form was obligatory in a given context, the target form *theta* or one of its contrasting forms. For example, Dyad 8 demonstrated the following:

(2) 82. JunHee: From:::, Steel Symbol Court, and, Tem, De, Tem, Temeh? Tema Park Drive?
83. SoYoung: Uh, you mean Steel Symbol Court and Tem, *Theme* Park Drive, right [ ]
84. JunHee: Tema? Thim? Tim?
85. JunHee: Uh, Symbol
86. SoYoung: Steel Symbol Court and Theme Park Drive right?
87. JunHee: Theme yeah not Team
88. SoYoung: Yes. Um::: (2) Yes, so:::.

Several more- and less-targetlike modifications are visible in this exchange. Line 82 contains JunHee's initial nontargetlike output as she engages in hypothesis testing to determine how she believes the street name should be pronounced. In line 83, SoYoung provides corrective feedback to JunHee and also self-corrects (+), modifying her own production from "Tem" to "Theme." In line 84, JunHee produces (+) and (−) modifications of the target form while continuing with her hypothesis testing. SoYoung provides additional feedback by recasting the two streets at the intersection in question, confirming with a tag question at the end, "right?" JunHee's final modified, pushed output is ultimately targetlike, "Theme," emphasized by explicitly juxtaposing it with the alternative, not-desired street name, "not Team." In this case, the self- and other-initiated modifications resulting in more targetlike output are particularly noteworthy since due to the L1 composition of this class, both interlocutors here were Korean women. Thus, the example demonstrates that even two interlocutors who share an L1 can still push each other to attend to phonological form and modify their output to be more targetlike.

However, a few other unexpected findings arose. Although all participants demonstrated the physical ability to articulate targetlike /θ/, /s/, /t/ and /f/ distinctly under at least some conditions (in particular when juxtaposing two minimal pair words for more salient contrast), six participants demonstrated non-standard /s/ production that ranged from a slightly less-strident production in some environments to a clearly interdental "lisp" of sorts. It should be noted that none of these participants were Spanish, and thus none of this nonstandard /s/ articulation could be seen as L1 influence. The other L1s represented by these learners do not typically have any interdental phonemes in their standard inventories. Thus, whereas the expected problem was overgeneralization of [s] where /θ/ was obligatory, the actual challenge was when the [θ] was over generalized in /s/-obligatory contexts when the minimal pair contrasted /s/ and /θ/ specifically. For most of these learners, the negotiation routines forced them to distinguish the phonemes more carefully and

were surprisingly able to modify their output and produce a more strident /s/ when contextually necessary. Only one participant was unable to perceptibly articulate a difference between the two from beginning to end, although her discourse seemed to indicate that she believed she was producing two completely unique sounds and that all breakdowns in communication (related to that contrast) were due to her interlocutor's weak listening skills. Several native speakers, and a native speaker of her L1, who listened to excerpts of her task performance were all unable to confirm what she (Tsin-Hua) claimed to be saying, as in the following exchange:

(3) 261. Tsin-Hua: I'm in da:::… uh:::… inter::section? of… chu, chru::ce? *(/θ/?/s/? unclear)* Alley.
262. HaeRyu: Truce?
263. Tsin-Hua: Truth.
264. HaeRyu: Could you tell me again?
265. Tsin-Hua: Truth? Th, th, th, (*laughs*)
266. HaeRyu: (*laughs*)
267. Tsin-Hua: ***Truth*** Alley.
268. HaeRyu: ***Tru-thə***? (*[truθə]; stressed the final [θə] syllable for added salience*)
269. Tsin-Hua: No, ***Truth***. *(Thinks she's saying Truce? Non-strident, interdental quality to final phoneme; sounds like /θ/.*)
270. HaeRyu: Truce?
271. Tsin-Hua: Yeah.
272. HaeRyu: Okay.

Interestingly, Tsin-Hua seemed quite capable of discerning which form her interlocutor was articulating, but was otherwise unaware that her production of those forms was not equally clear to others. Nevertheless, the IGT design (specifically balancing form-essentialness and communication requirements, as mentioned above) still enabled the two interlocutors to participate fully, providing each other with corrective feedback and thus modified output opportunities, which did help HaeRyu improve her production clarity of the target form contrasts and help both interlocutors hone their listening accuracy. Potentially, if an instructor were to explain to Tsin-Hua the difference between the /s/ and /θ/, and make her realize that she was not actually producing a targetlike /s/, substituting /θ/ instead, (perhaps by recording her voice and the teacher's for auditory contrast, or using speech analysis software such as Praat™ that provides spectrographic images of one's voice and will visually show such strident differences, as I have successfully done with other students) an IGT such as this map task could be a useful tool to help her learn to automatize this new knowledge as well.

One additional issue regarding more- and less-targetlike modifications warrants clarification. When comparing total numbers of each modification type, there is no numerical distinction between three potential degrees of consciousness in these modifications: (a) a general lack of control in both directions and lack of awareness of this problem, (b) what I call "backsliding," which is a subconscious non-targetlike modification away from a previously corrected form; and (c) deliberate, conscious modification towards a more targetlike form. The first was rare and difficult to verify, but there were times when it appeared that a speaker was simply inconsistent in his production of a particular target form or contrast

in a given exchange and unaware of his inconsistency, so a series of seemingly random more- and less-targetlike modifications in a very short time frame did not appear to be due to attention-related choices. However, usually corrective feedback (e.g. clarification requests and confirmation checks) from the interlocutor would soon force the speaker to be more careful and attend to the form, which lessened the problem in future exchanges. Secondly, "backsliding" is a situation in which the speaker began with a nontargetlike utterance, and after modifying it to be more targetlike (+), soon thereafter seemed to lose focus on the form and relax her vigilance in articulatory accuracy, subsequently reverting to the original non-targetlike utterance (–). This is demonstrated in the following excerpt in which Li-Xin produces a non-obligatory [θ] when trying to direct her partner along Norse Man's Alley but immediately self-corrects to a more targetlike (+) form, "Norse" in line 197, but then "backslides" in line 203 with "North" (–), again catching herself and immediately modifying (+) to the obligatory [s] for Norse Man's Alley:

(4) 197. Li-Xin: Ahah. Okay, Sor, s, serv, serv, way an North, Norse Man's
198. Dong: Yeah
199. Li-Xin: Alley.
200. Dong: Yeah.
201. Li-Xin: Okay (2) Uh, you can keep, keep going to the south?
202. Dong: Yeah?
203. Li-Xin: On North, Norse Man Alley.

As Li-Xin's /s/ production was less strident than standard, closer to /θ/, and her /θ/ articulation was not interdental enough to be targetlike, resulting in a sound that was virtually identical to her version of /s/, this also caused confusion for her interlocutor. Yet through the negotiations, having received corrective feedback from her interlocutor multiple times before this exchange, she attended to the forms and managed to distinctly produce them more accurately. Though her backsliding seemed subconscious, her own audio and proprioceptive feedback brought it to a more conscious level, which then prompted her to deliberately endeavor to modify it to be more targetlike (+). Thus in this exchange, regardless of the degree of conscious attention to the modifications, there were three instances of modified output in total: two that were more-targetlike (+) and only one that was nontargetlike. I argue that this excerpt and others like it are evidence of *resensitization,* as mentioned above, as a necessary step towards adult L2 phonological development.

As this was specifically a process-oriented study, looking at how adult learners attend to phonological forms through modified output, there was no pre- or post-task assessment on accuracy of target-form production or perception. Although such measurements would be instructive for teachers and researchers alike in the future, the goal was to identify online cognitive processes involved in phonological acquisition. The evidence appears clear that NNSs do indeed attend to phonological forms and help each other become more targetlike in their perception and production thereof in much the same way they might attend to lexical or morphosyntactic forms that cause a breakdown in communication, particularly if the task conditions facilitate this by balancing the communicative demands and requirements of the task with essentialness of target form accuracy.

Note

* This paper presents a portion of the findings from my doctoral dissertation, *"No, they won't 'just sound like each other'": NNS-NNS Negotiated Interaction and Attention to Phonological Form on Targeted L2 Pronunciation Tasks*, from the University of Pennsylvania, Philadelphia, USA (Sicola, 2009).

References

Archibald, J. (1998). *Second Language Phonology* (Vol. 17). Philadelphia: John Benjamins.

Avery, P., & Ehrlich, S. (1992). *Teaching American English Pronunciation*. Oxford: Oxford University Press.

Bitchener, J. (2004). The relationship between the negotiation of meaning and language learning: A longitudinal study. *Language Awareness, 13*(2), 81–95.

Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.

Ellis, R. (1999a). Retrospect and prospect. In R. Ellis (Ed.), *Learning a second language through interaction* (233–257). Amsterdam: John Benjamins.

Ellis, R. (1999b). Theoretical perspectives on interaction and language learning. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 3–31). Amsterdam: John Benjamins.

Ellis, R. (Ed.). (1999c). *Learning a Second Language through Interaction* (Vol. 17). Amsterdam: John Benjamins.

Fraser, H. (forthcoming). Cognitive theory as a tool for teaching pronunciation. In S. de Knop, A. de Ryker and F. Boers (Eds.). Berlin: Mouton de Gruyter.

Gorsuch, G. J. (2001). Testing textbook theories and tests: The case of suprasegmentals in a pronunciation textbook. *System, 29*(1), 119–136.

Griggs, P. (2005). Assessment of the role of communication tasks in the development of second language oral production skills. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (407–432). New York: Mouton de Gruyter.

Hansen, J. G. (2004). Developmental sequences in the acquisition of English L2 syllable codas: A preliminary study. *Studies in Second Language Acquisition, 26*, 85–124.

Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford: Oxford University Press.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (413–468). San Diego, CA: Academic Press, Inc.

Long, M. H. (2003). Stabilization and fossilization in interlanguage development. In C. Doughty & M. H. Long (Eds.), *The Handbook of second language acquisition* (487–535). Malden, MA: Blackwell Publishing.

Long, M. H., & Porter, P. A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly, 19*(2), 207–228.

Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (15–41). Cambridge: Cambridge University Press.

Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (123–167). Clevedon: Multilingual Matters.

Major, R. C. (1987). A model for interlanguage phonology. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (101–124). Cambridge: Newbury House Publishers.

Marinis, T. (2003). Psycholinguistic techniques in second language acquisition research. *Second Language Research, 19*(2), 144–161.

Markham, D. (1997). *Phonetic Imitation, Accent, and the Learner*. Lund: Lund University Press.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25*(3), 481–520.

Pennington, M. C. (1998). The teachability of phonology in adulthood: A re-examination. *IRAL, 36*(4), 323–341.

Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly, 20*(2), 207–225.

Pica, T. (1988). Morpheme data analysis in second language acquisition research: Renewing an old debate and raising new issues. *ITL, Review of Applied Linguistics,* 79–80, 77–112.

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning, 44*(3), 493–527.

Pica, T., Kang, H.-S., & Sauro, S. (2006). Information gap tasks: Their multiple roles and contributions to interaction research methodology. *Studies in Second Language Acquisition, 28*(2), 301–338.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Psycholinguistics, 21*(1): 27–57.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (3–32). Cambridge: Cambridge University Press.

Shehadeh, A. (1999). Non-native speakers' production of modified comprehensible output and second language learning. *Language Learning, 49*(4), 627–675.

Sicola, L. (2009). *No, They won't 'Just Sound like Each Other': NNS-NNS Negotiated Interaction and Attention to Phonological Form on Targeted L2 Pronunciation Tasks* (Vol. 72). Frankfurt/M.: Peter Lang.

Sicola, L. (in press). Learner-learner corrective feedback in dyadic pronunciation tasks. In A. Shehadeh & C. Coombes (Eds.), *Task-based learning*. Arlington: TESOL.

Spada, N., Lightbown, P. M., & White, J. L. (2005). The importance of form/meaning mappings in explicit form-focused instruction. In A. Housen & M. Pierrard (Eds.), *Investigations in Instructed second language acquisition* (199–234). New York: Mouton de Gruyter.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (235–253). Rowley, MA: Newbury House Publishers.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (64–81). Cambridge: Cambridge University Press.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*(3), 371–391.

Swan, M., & Smith, B. (1987). *Learner English: A Teacher's Guide to Interference and other Problems*. Cambridge: Cambridge University Press.

Williams, J. (2005). Form-focused instruction. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (671–691). Mahwah, NJ: Lawrence Erlbaum Associates.

Appendix A

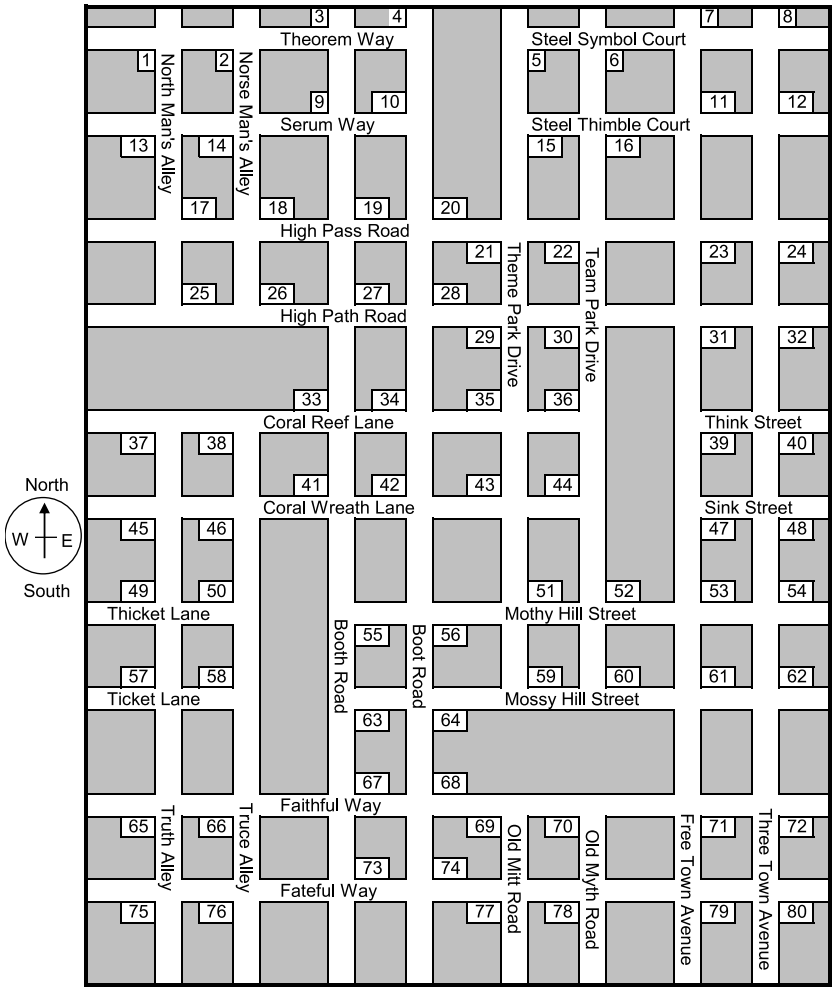
Sicola
University of Pennsylvania GSE

2006

Where in New Philadelphia is…?

Partner A:	Corner #	Location
	____	Constitution Center
	____	Reading Terminal
	____	Franklin's House
	____	Penn's College

Partner B:	Corner #	Location
	____	Liberty Bell
	____	Art Museum
	____	City Hall
	____	Italian Market



Appendix B

"Where in New Philadelphia is…?" Map Task Instructions:

(*Note: You may not look at your partner's picture, and you may not make eye-contact with your partner for this activity.)

1. On the next page, you have a map of "New Philadelphia." You will be Partner A, and your partner will be Partner B. Above the map, there is a list of popular tourist locations in New Philadelphia. You will need to give your partner directions to get to the places on your list, and your partner will give you directions to get to the locations on the Partner B List. One corner of each intersection is numbered.

2. For each place on List A, choose a street corner for its location. Write the number of the street corner on the line next to the place name on your list. (Example: If you want the Constitution Center to be at the intersection of Truce Alley and Fateful Way, you can write the number 76 next to the words "Constitution Center" on your list.) Choose a different number for each place.

3. When you are ready to "go sight seeing," ask your partner for directions to the places on List B. These are the rules for giving and receiving directions:

- First, the receiver must tell the other person where she/he will start walking. (Example: "How do I get to City Hall from the intersection of High Pass Road and Booth Road?")
- You can NOT tell your partner how many blocks to go. (Example: You can NOT say "Go three blocks south on Booth Road, and then…") L
- You can ONLY give street names and compass directions such as north, south, east, west, right or left. (Example: "From High Pass Road, walk east until Theme Park Drive. Turn south on Theme Park Drive and go to Mossy Hill Street. Then…") J
- You can NOT mention numbers on street corners when you are giving or receiving directions.

4. Finally, the person giving directions will lead the receiver to the intersection of streets where the location is, and ask what number is on the corner. (Example: "And finally, go south on Truth Alley until you reach Ticket Lane. The Liberty Bell is on the north-west corner. What number is it?")

5. The receiver should then tell what number is on that intersection. If it is correct, then you can go on to the next location on the other person's list, and switch roles. If it is not correct, go back to the starting point and try again.

CHAPTER 18

Quality and type of corrective feedback, noticing, and learner uptake in synchronous computer-mediated text-based and voice chats

Susana Sotillo
Montclair State University

1. Introduction

Extensive research on second language acquisition (SLA) and interaction in face-to-face contexts has established that learners need opportunities to interact with native speakers of the language and other interlocutors in order to negotiate meaning and modify their output (Long, 1991, 1996; Gass & Varonis, 1994; Pica, 1994). Following Swain's (1985) Output Hypothesis, researchers have been able to show that negotiated interaction often leads to repair moves that bring about modifications of learner output (Pica, Lincoln-Porter, Paninos, & Linnell, 1996). Thus, it is during negotiation work, that language learners are able to notice linguistic gaps in their knowledge as they shift attention from meaning to form (Gass, Mackey, & Pica, 1998; Pica, 1994). The cognitive process of noticing has been the focus of recent SLA investigations, and researchers are currently examining psycholinguistic factors such as attention and awareness, as well as previously overlooked learner-initiated repair moves (Shehadeh, 2001).

While most SLA research has taken place in classroom contexts and language laboratories, advances in technology offer new opportunities for the study of language in computer-mediated settings. For example, researchers have studied the quantity and quality of language produced by learners in negotiation work while engaged in jigsaw, decision-making, and free-discussion tasks in asynchronous and synchronous CMC (Beauvois, 1998; Smith, 2003, 2005). CMC studies have also replicated earlier work on focus on form and negative or corrective feedback. Findings based on these recent studies indicate that opportunities provided for learner interaction and negotiation work in online environments appear to facilitate language learning (Basturkmen, Loewen, & Ellis, 2002; Ellis, Basturkmen, & Loewen, 2001; Williams, 1999). Partly motivated by Warschauer's (1997) observations that the text-based aspect of synchronous computer-mediated communication (SCMC) amplifies students' attention to linguistic form, recent investigations have focused on the psycholinguistic process of noticing in online interaction (Shekary &

Tahririan, 2006). Other studies have examined the impact of different types of computer-mediated corrective feedback on language learning (Loewen & Erlam, 2006).

The present study reanalyzes 89% of the chat logs and transcribed voice chats from a previous investigation of exchanges between tutors and ESL learners as they collaborated on five learning tasks in an instant messaging environment *Yahoo! Instant Messenger* (YIM). It examines the instantiation of noticing as tutor- and learner-initiated language-related episodes (LREs), the type and quality of corrective feedback, learner uptake and successful uptake.

1.1 Focus-on-form studies and SLA

As it became apparent that language learners in communicative classrooms continued to experience linguistic problems, a group of SLA researchers refined and extended the focus on form (FonF) approach originally proposed by Long (1996), who argued that instruction in a communicative classroom that includes focus on linguistic form can increase the salience of positive evidence and provide essential negative evidence as negative feedback to learners. A more narrow definition was later proposed by Long and Robinson (1998), who stipulated that attention to form be drawn in the course of communicating meaning in language classrooms. Doughty and Williams (1998) examined the optimal degree of attention to form at length in their edited collection of FonF studies in second language classrooms. Successive FonF investigations have revealed that focusing on form in meaning-centered language classrooms is an instructional option that leads to instances of successful learner uptake (Ellis et al., 2001; Williams, 2001). However, this is not a new approach since it was first observed and reported by Canadian researchers investigating second language learning in immersion classrooms.

1.2 Language-related episodes, attention to form, and SLA

Swain and Lapkin (1995) were among the first to study how learners negotiated form during collaborative writing tasks. They labeled these negotiation-of-form activities Language-Related Episodes (LREs). The collaborative nature of the LREs was stressed by Swain (1995, 1997, 2000) because of its importance in allowing researchers a means of observing the cognitive process of noticing, viewed as a crucial component of language acquisition by Schmidt (1995). Swain (2000) claimed that LREs as mini-dialogues allow learners to focus on various aspects of language while collaborating on meaning-focused tasks.

Though Schmidt (1995) defined noticing as detection or registration plus controlled activation into the focus of conscious attention, Robinson (1997) maintains that noticing requires not only detection, but also the allocation of attentional resources and rehearsal in short-term memory. SLA researchers hold different positions with respect to the quality of attention and memory deemed important in language learning (Robinson, 1995). These positions draw heavily on seminal work on attention and memory storage by Cowan (1995, 1998).

With respect to the cognitive process of noticing, various SLA researchers have used LREs in their investigations. In most cases, new features have been added to the original definition. For example, Williams (1999) redefined LREs as learner-initiated episodes that focused on negotiation of either meaning or form. She found that the degree and type of learner-initiated attention to form was related to proficiency level and the nature of the learning activity. Subsequently, Ellis, Basturkmen, & Loewen (2001) used the term Learner-Initiated Focus on Form in their classroom investigations of learner language and attention to form. Whether explicit or implicit, negative or corrective feedback to learners encourages these processes which seem to facilitate language learning as evidenced by immediate and delayed post-test gains reported.

1.3 Explicit and implicit negative or corrective feedback

Some researchers point out that the provision of error correction or negative information, orally or in writing, should be referred to as negative feedback since intentionality, that is, the interlocutor's intention to correct, cannot be assumed. Explicit and implicit forms of negative feedback have been investigated in classroom studies and laboratory experiments, and various SLA researchers have pointed out that implicit negative feedback in the form of recasts is beneficial to learners' short-term development of the grammatical forms targeted (Iwashita, 2003). In the Doughty and Varela's (1998) quasi-experimental study, implicit negative feedback provided to learners as corrective recasts on oral and written measures contributed to significant gains in post-test scores. Others, however, have suggested that negotiation moves, such as clarification requests, are more beneficial to learners than recasts because they draw learners' attention to linguistic forms (Lyster & Ranta, 1997).

Also, recent studies that examine the provision of indirect or explicit negative feedback by teachers or interlocutors have been carried out in a variety of technology-mediated environments that allow for both asynchronous and synchronous computer-mediated communication.

1.4 Computer-mediated communication (CMC) and SLA

CMC, whether synchronous or asynchronous, oral or written, has been shown to offer learners a potentially rich environment for language learning. In the late 1990s, investigations of networked language learning reported that more equitable learner participation and more opportunities to use more complex language were among the major benefits found in this type of environment (Beauvois, 1992; Warschauer, 1996). CMC studies have also investigated the development of linguistic complexity in students' asynchronous and synchronous discussions (Sotillo, 2000). Other researchers have examined language learning outcomes and the differences between F2F and CMC interaction (Warschauer, 1996; Warschauer & Kern, 2000). From a sociocultural perspective, CMC has been investigated as a site where participants co-construct a dynamic and student-centered discourse community (Darhower, 2002).

The Varonis and Gass (1985) schema, which includes a trigger, indicator, response, and reaction, is frequently used for coding negotiation work in mixed dyads (Blake, 2000). The focus of most CMC negotiation work has often involved lexical items and not morpho-syntax (Blake, 2000; Morris, 2005), but others have utilized the standard jigsaw, decision-making, and free-discussion tasks used in SLA research (Smith, 2003, 2005).

1.5 Noticing in synchronous computer-mediated communication or SCMC

There are three recent studies that have investigated noticing (Shekary & Tahririan, 2006) and different types of corrective feedback in online negotiated interaction (Loewen & Erlam, 2006), but only a handful of studies have explored the advantages of video and audio in current synchronous computer-mediated communication (SCMC) and its impact on language learning. Additionally, a limited number of studies have explored these modalities during interaction work outside classroom settings and across national borders (see Belz & Thorne, 2006; Levy & Kennedy, 2004; Sotillo, 2005). Jepson (2005) was among the first to investigate NNS-NNS conversational repair moves in both text and voice synchronous chat rooms. He found that self-correction, which is considered evidence of noticing, was not found in SCMC. He claimed that the NNS-NNS SCMC context was not conducive to self-correction, but Smith (2008) challenged Jepson's findings since there is ample research that shows learners do indeed focus on form in SCMC contexts. Smith (2008) argues that SCMC interaction needs to be captured using video, audio, and screen capture tools in order to highlight its many nuances. According to Smith and Gorsuch (2004), relying on printed chatscripts is insufficient to document the important cognitive process of noticing for language learning.

Thus, these recent SCMC studies appear to indicate that L2 learners could benefit from computer-mediated corrective feedback and attention to form during negotiation work with tutors in both text-based and voice chat environments. It seems possible that L2 learners may notice the difference between their non-targetlike output and the correct/appropriate form provided as negative information or corrective feedback by tutors.

2. Study methodology

This study reanalyzes data from chat logs and transcribed voice chats from a previous investigation of exchanges among members of six dyads: tutors who were teachers-in-training and ESL learners as they collaborated on five learning tasks in an instant messaging environment *Yahoo! Instant Messenger* (YIM). The original study was motivated by focus-on-form (FonF) investigations of the communicative language classroom, specifically the work of Ellis, Basturkmen, & Loewen (2001).

The data analyzed for the current investigation focuses on voice and text-based chat data obtained from two NS-ESL dyads and two ANNS-ESL dyads. This reanalysis of data was motivated primarily by investigations of repair work in SCMC voice and text chats (Jepson, 2005) and recent work on noticing in SCMC (Shekary and Tahririan, 2006). Three research questions are addressed and the categories used in the methodology sec-

tion are adapted from previous work by Williams (1999), and Loewen (2004). In order to examine the cognitive process of noticing, the Language-Related Episode (LRE) is utilized. An LRE occurs when tutors and learners shift attention to linguistic forms during negotiation work while completing a task. LREs consist of implicit and explicit types of negative information or corrective feedback. Implicit corrective feedback moves include tutors' negotiation moves (comprehension checks and clarification and confirmation requests), repetitions, recasts (reformulation of all or part of a learner's utterance). Explicit corrective feedback includes metalinguistic feedback and the provision of information or explanations, and reaction to learner-initiated requests for linguistic information. In addition to learner-initiated explicit requests for linguistic information, learner self-correction moves are also included in LREs. Thus, given the number of learner-interlocutor exchanges and negotiation work documented in previous CMC studies (Blake, 2000; Jepson, 2005; Smith, 2003), the current study seeks to identify the frequency of all LREs in text-based and voice modalities, types and quality of corrective or negative feedback provided by tutors, and instances of optional learner uptake or response.

2.1 Research questions

Voice and text SCMC chats have been shown to have the potential to benefit language learners outside of classroom contexts and to possibly facilitate L2 learning processes. The examination of noticing in what Swain (1997, 2000) refers to as collaborative dialogs is an important first step in determining how learners indicate lack of knowledge of a specific linguistic item. Since in this study no tailor-made immediate or delayed post-tests were administered to participants following the completion of collaborative learning tasks, it was not possible to evaluate the effectiveness of incidental noticing in SCMC on subsequent language learning. One of the goals of this study is to identify if learners were able to notice or detect and bring into focus linguistic forms during naturally occurring negotiation work in an online environment using voice and text-based chat modalities. Noticing as used in this study is operationalized as the learner's act of detecting a problematic lexical or grammatical feature in his/her output that prompts a language-related episode (LRE) during negotiation work with a tutor. This LRE may be initiated by the learner as either a self-correction, a clarification request, a request for information concerning a linguistic item, or it may also constitute the learner's reaction to a clarification request, comprehension check, or confirmation check/repetition by the tutor. Attention is shifted from the content or meaning being negotiated to a linguistic item. Noticing in this study is operationalized as LREs. The following three research questions are addressed:

1. Which modality, text-based or voice chat, affords learners more opportunities for noticing their own non-targetlike output or inaccuracies during learner- and tutor-initiated LREs?
2. What type and quality of tutor-initiated negative information or negative feedback is provided to ESL learners during negotiation work in these two modalities?
3. Is there evidence of learner uptake and subsequent successful uptake in these online LREs?

2.2 Participants and tasks

Negotiation work in voice and text-based chats occurred naturally since participants were not monitored by the researcher or any ESL instructor. This was not a quasi-experiment taking place in a language laboratory. The participants were collaborating on various simple tasks from the comfort of their own homes while using *Yahoo!* Instant Messenger and a Webcam that had been provided to 12 volunteers. In the original study, two dyad members dropped out because of scheduling conflicts and technical problems.

The two ANNS tutors working with ESL learners spoke Spanish and Brazilian Portuguese, respectively, and the NS tutors were enrolled in teacher-training classes at a large public four-year university. The ESL learners recruited for the original pilot study had been referred to the researcher by their former ESL instructors, but they were not enrolled in classes. There were four females and one male whose ages ranged from 18 to 36. These volunteers were encouraged to participate in this nine-week study so that they could receive ESL tutoring via *YIM*. One of the ESL volunteers was a male student who lived in Sao Paolo, Brazil, and spoke Brazilian Portuguese as a first language. Other first languages spoken by the ESL learners included: Korean, Spanish, and Vietnamese. The NS and ANNS tutors ranged in age from 24 to 32. These teachers-in-training were familiar with SLA theory and research that addressed the relevance of both negative and positive feedback to L2 learners. However, they were not instructed by the researcher or staff to favor either the provision of direct or indirect negative feedback to learners. They were simply asked to encourage learners to use their L2 during interaction via *YIM*, and to use both voice and text-based chat whenever possible.

In exchange for their participation in the original pilot study, all volunteers were loaned *LogiTech* Webcams and the accompanying software for up to five months. All participants, tutors and ESL learners, first filled out a questionnaire developed by the Office of Information Technology (OIT) staff in order to determine whether they could use the software and Webcams. They were also shown how to use the equipment on a one-on-one basis. NS tutors were encouraged to contact the researcher via *YIM* in order to test the voice and visual modalities of this type of SCMC. They were taught how to initiate chat sessions, record their voice chats, and save the chat logs for subsequent analysis.

2.3 Tasks and materials

Data were collected from tutors and ESL learners as they completed the following 45-minute collaborative tasks: (a) jointly filling out a career goals questionnaire, (b) synthesizing information inferred from reading a newspaper or magazine article, (c) negotiating individual perceptions about the content of a movie or documentary each student had seen separately, and (d) discussing and evaluating the usefulness of instant messaging as a language learning tool. All dyad members interacted outside a laboratory or traditional classroom setting using similar equipment, resources, and tasks.

2.4 Coding procedures for current study

Text and voice chat data from four of the original five dyads of the pilot project were separately reanalyzed by a trained research assistant and the principal investigator. Seventy minutes of oral and voice chat data were excluded from this reanalysis. The transcribed voice chats were separated from the text-based chats logs in order to identify instances of noticing in negotiation work between tutors and learners. This allowed for more reliable identification and coding of the type and quality of negative information or corrective feedback provided to learners, and instances of both learner uptake and successful uptake.

Chat logs and transcribed voice chats were coded separately using categories from recent SCMC research such as a modified definition of LREs, which included both learner- and student-initiated LREs. Coding results were then compared and misunderstandings clarified in order to reach consensus. The Pearson correlation coefficient obtained for all the coded categories was .9454, following a Z transformation. MonoConc was used following the coding and tagging of categories in this 16,525 word corpus.

These LREs are examined within the context of negotiation work between ESL learners and their tutors (Swain & Lapkin, 1995; Williams, 1999). For example, in a text-based chat, a learner may express non-understanding (trigger) while completing a learning task or activity with his/her tutor or engage in self-repair by rephrasing or restructuring an utterance in voice chats. Following Shekary and Tahririan's (2006) reformulation, each LRE is defined as consisting of three basic moves: trigger (indicator of non-understanding), response or corrective feedback, and the optional uptake or learner response.

Excerpt 1 from a text-based chat illustrates an LRE which occurred in response to an error as the learner and his tutor were collaborating on a task that involved finding out about each other's background and career goals. Learner non-targetlike output is italicized, and type of corrective feedback and learner response (uptake) appear in bold.

LRE – Excerpt 1 – Text-based Chat

249Ana (ESL):	I have five more sisters, two in my country and 3 more here in the US. This is the house of my older sister *she had* always been very nice. When I came to US she was living in Montclair, thank Good, we just stayed there for a little *bet* over a week
250Ana (ELS):	*she has* always been (**self-correction**)
251Lupe (ANNS):	okay, she has always been (tutor **confirmation check**)
252Lupe (ANNS):	Great.
253Lupe (ANNS):	*a little bit* (**recast**)
254Ana (NNS):	yes, *little bit* (**uptake**)

2.5 Type and quality of corrective/negative feedback and timing of turns

Negative feedback is the term currently in use for the provision of error correction or negative information, orally or in writing. The provision of negative feedback can be done either indirectly through the use of recasts, clarification requests, confirmation checks,

Table 1. Language-Related Episodes (LREs) in Tutor-ESL Learner SCMC (Adapted from Loewen, 2004; Shekary & Tahririan, 2006)

Features/ Characteristics	Definition	Type and Examples
Language-Related Episodes (LREs)	A LRE is instantiated when attention is momentarily shifted from focusing on content or meaning to linguistic form while collaborating on a task (e.g., negotiating individual perceptions concerning a movie both had seen separately). It normally consists of a trigger, a response, and an optional learner acknowledgement or uptake. These are often classified as either reactive or preemptive in the SLA literature.	Learner-Initiated LERs: Attention shift to linguistic form is initiated by the ESL learner during negotiation work. This may be done through self-correction or by requesting information about a specific linguistic item. Other-Initiated LREs: A NS or ANNS tutor shifts attention to linguistic form during negotiation work in reaction to lexical, morphosyntactic or phonological difficulty.
Negative Information or Corrective Feedback	Negative information or corrective feedback is provided to L2 learners by tutors in response to a trigger or in reaction to a learner's request for linguistic information. Two types of feedback may be provided to L2 learners: direct or explicit, indirect or implicit. Feedback may also be of two types: Accurate Feedback that conforms to academic or Standard American English (SAE) lexical, morphosyntactic, and phonological conventions. Inaccurate: Feedback does not conform to SAE conventions.	Direct feedback includes metalinguistic feedback, definition of terms, or the provision of explicit information or explanations. Indirect feedback involves an implicit response to a learner's request or a tutor's reaction to non-targetlike learner output. This includes: recasts or reformulation of all or part of a learner's utterance/output, clarification requests (e.g., *What do you mean?*), repetitions (*horror, horror*), confirmation checks (e.g., *Heaven is where happiness is?*), or comprehension checks (e.g., *Does that make sense?*)
Intervening Turns following corrective feedback (CF)	Immediate and Delayed Immediate CF is provided following the trigger or learner request Delayed CF occurs after two or more turns	50Rob (ESL): I'm student in binesses and I not finish my underqraduate class 51Kay (ANNS): YOU should say, *Invite me to view your webcam*. (metalinguistic information) 52Rob (NNS): ok (uptake)
Learner Response	ESL learners' response to feedback or Uptake and no Uptake.	Uptake includes acknowledgment of corrective feedback provided by tutor. No Uptake occurs when corrective feedback is not acknowledged, topic shifts or learner continues with meaning-focused discussion.
Type of Learner Response	There are two types: successful uptake and unsuccessful uptake.	Successful Uptake is when the learner incorporates corrective feedback on a linguistic item into his/her output as evidenced by its targetlike use in subsequent turns. Unsuccessful Uptake consists of no evidence that learner has incorporated corrective feedback provided by tutor.

repetitions (of learner's exact utterance) and comprehension checks, or directly via metalinguistic explanations, definition of terms for clarifying lexical confusion, or by supplying explicit information concerning vocabulary or morphosyntax.

The timing of turns refers to the point at which negative feedback follows a trigger or the point at which a learner notices his/her non-targetlike output following a self-correction trigger. In immediate turns, the feedback occurs following the trigger, whereas in deferred turns, the feedback occurs two or three turns after the trigger.

In Excerpt 2, Tina (an intermediate-advanced ESL learner) and Dana, her NS partner, are discussing a newspaper article both have read. Dana politely provides indirect negative feedback in the form of a recast (i.e., repetition of all or part of a learner's utterance that corrects the error), which Tina acknowledges (uptake) but fails to incorporate into her subsequent response or in negotiation exchanges following this LRE. Dana, however, did not address Tina's pronunciation errors during their voice chats.

LRE Episode – Excerpt 2 – Voice Chat (transcribed oral exchange all in italics)

001Tina (NNS):	*I was thinking that this little girl was very brave too. I mean a lot of… it has a lot of courage*
002Dana (NS):	*She had a lot of courage (**recast**)*
003Tina (NNS):	*Uh huh (**uptake**)… because he was ... were not too rude, but too Impolite with her, and that remind__ me that sometimes parents are like that or aunts or your [dʒúwr] grandma because they want you [dʒúw] to learn stuff but sometimes it's just not the way..I mean I don't know*

In contrast, Kay, another ANNS tutor, addresses her partner's pronunciation errors in their voice chats, but on many occasions she provides inaccurate or non-targetlike feedback to Rob. For example, in Excerpt 3, she stresses the second syllable of "character."

LRE Episode – Excerpt 3 – Voice Chat (all italics) – Quality of Negative Feedback

199Rob (NNS):	*I like the film because that story is not only horror. The story of the movie shows a life [J]ess-an and a romance too.*
200Kay (ANNS):	*That's true. I also love[d] this movie, but I think it was a little scary to me. There were a lot of strong scenes, and I had to close my eyes. But what do you think about Bruce Willis character /kær 'ik tər/ (instead of /'kær ik tər/)*
201R (NNS):	*I don't believe that you closed your eyes! I think he's a very good actor. He looked all the time like he were alive.*

Often LREs initiated by this ANNS interlocutor provided non-targetlike feedback or incorrect grammatical input to the learner. Though the learner acknowledged his tutor's corrective feedback, he was unable to detect the inaccuracy in the feedback.

Finally, we also coded for instances of successful uptake or incorporation of targetlike word order, closed-class grammatical structures (e.g., articles, prepositions), and vocabulary provided as negative feedback to learners. Excerpt 4 illustrates how a learner incorporates corrective feedback into his evolving interlanguage. This learner continued

to use "invite me to view your Webcam" throughout the remaining tasks he completed in collaboration with his tutor.

LRE – Excerpt 4 – Successful Uptake in Negotiation Work – Text-based Chat
(learner inaccuracies or errors are italicized)

47Kay (ANNS): MY second question ...Are you an undergraduate or graduate student?
48Rob (ESL): ok
49Rob (ESL): I can't see you. *Please send an invitation for the webcam.*
50Rob (ESL): I'm student in *binesses and I not finish my underqraduate class*
51Kay (ANNS): YOU should say, Invite me to view your webcam. (**explicit correction**)
52Rob (ESL): ok (**uptake**)
53Rob (ESL): MY plans are to study *on post graduation*, after look for a good job and live in a big city or live in other country
54Rob (ESL): sorry other (**self-correction**)
55Rob (ESL): invite me to view your webcam? (**successful uptake**)

3. Results

Table 2 displays raw frequency counts for instances of noticing, operationalized as tutor- and learner-initiated LREs in text-based chats. It also shows the duration and total number of words. NS-ESL learner dyad pairs spent considerably less time interacting in an SCMC environment than ANNS-ESL learner dyad pairs (60 vs. 355 minutes, respectively). Learners initiated 54% of all LREs identified in these text-based chats, whereas tutors initiated 46% of all LREs. The word count for the ANNS-ESL learner dyads displayed in Table 2 is twice that obtained for the NS-ESL learner dyads (6327 vs. 3098, respectively).

Table 2. LREs in Text-Based Chats, Duration, and Word Count
(based on Long, 1983; Swain, 2000; Williams, 1999; Shekary & Tahririan, 2006)

Dyads	Minutes Text-based Chats	Language-Related Episodes (tutor- and learner-initiated – Explicit and Implicit corrective feedback and instances of self-correction)		# Words in Text-based Chats
		TI	LI	
NS-ESL Learners (2 dyads)	60	1	11	3098
ANNS-ESL learners (2 dyads)	355	27	22	6327
Totals	415	28 (46%)	33 (54%)	9425

Table 3 shows that learners initiated 65% of all LREs in voice chats, whereas tutors initiated only 35% of all LREs (LI vs. TI), and 52% of these were prompted by learner output

errors. Though learner-initiated LREs in voice chats were fewer than those found in text-based chats as shown in Table 2 (24 vs. 33, respectively), we need to keep in mind that only one pair of the two ANNS-ESL learner dyads tape recorded their voice exchanges via *YIM*. Voice chat data from the other ANNS-ESL learner dyad were not available. The results show that the ANNS-ESL learner dyad pair collaborating on tasks in voice chats spent more time negotiating both form and meaning than members of the two NS-ESL learner dyads (140 minutes vs. 70 minutes, respectively). The word count for the two NS-ESL learner dyads was slightly higher than that obtained for this ANNS-ESL learner dyad pair (4002 vs. 3098, respectively). Noticing, operationalized as paying attention to form and self-correcting during tutor- and student-initiated LREs, did occur and was more frequent in text-based chats than in voice chats (61 vs. 37 instances, respectively). These frequency counts are displayed in Tables 2 and 3. In other words, learners noticed their gaps in linguistic knowledge as evidenced by the number of LREs they initiated while completing learning tasks with their tutors. There were a total of 118 LREs identified in these two modalities of SCMC, which were initiated by both tutors and learners, but not all these LREs resulted in the provision of explicit or implicit negative information or corrective feedback to learners.

Table 3. LREs in Voice Chats, Duration, and Word Count
(based on Long, 1983; Swain, 2000; Williams, 1999; Shekary & Tahririan, 2006)

Dyads	**Minutes Voice Chats**	**Language-Related Episodes (tutor- and learner-initiated – Explicit and Implicit corrective feedback and instances of self-correction)**		**# Words in Voice Chats**
		TI	**LI**	
NS-ESL Learners (2 dyads)	70	4	7	3098
ANNS-ESL learners (1 dyad)	140	9	17	6327
Totals	210	13 (35%)	24 (65%)	9425

Table 4 displays frequency counts concerning the type of negative or corrective feedback provided by tutors in either reaction to learner errors or following learner requests during negotiation work. Direct feedback accounts for 77% of all negative feedback to learners in text-based chats, compared with 23% for indirect feedback. With respect to the timing of turns, 85% of text-based feedback occurs in the turn following the trigger. Few instances of feedback are deferred or occur more than two or three turns after the trigger. In voice chats, metalinguistic explanations, definitions, clarification of lexical confusion, and explicit information, all explicit or direct forms of feedback, account for 61% of all corrective feedback to learners. In contrast, 39% of all negative or corrective feedback is of an indirect type, primarily in the form of recasts, requests for clarification, confirmation checks, or repetitions. Concerning the timing of turns in voice chat LREs, raw frequencies and percentages show that 89% of this feedback is provided immediately following the trigger.

Table 4. Type of Negative Feedback to Learners and Timing of Turns in Text-based and Voice Chat

Dyad Type	Type of Negative Feedback		Timing of Turns	
	Indirect	Direct	Immediate	Deferred
Text-based Chats				
NS-ESL learner **(2 dyads)**	1	11	9	3
ANNS-ESL learner (2 dyads)	13	36	43	6
Totals	14 (23%)	47 (77%)	52 (85%)	9 (15%)
Voice Chats				
NS-ESL learner (2dyads)	7	3	9	1
ANNS-ESL learner (1 dyad)	7	19	23	3
Totals	14 (39%)	22 (61%)	32 (89%)	4 (11%)

Table 5 displays raw frequency counts for both accurate and inaccurate or inappropriate negative feedback provided to learners by native-speaker and advanced non-native speaker tutors. ANNS tutors provided inaccurate feedback to ESL learners 37 times out of a total of 40 instances of inaccurate or ungrammatical feedback provided by both NS and ANNS tutors in these SCMC environments. Thus, 47% of 97 instances of negative or corrective feedback to ESL learners was inaccurate or ungrammatical.

Table 5. Quality of Negative Feedback Provided by Tutors

Dyads	Quality of Feedback in both Text-Based and Voice Chats	
	Accurate	Inaccurate
NS-ESL learner (2 dyads)	19	3
ANNS-ESL learner (2 dyads)	38	37
Totals	57	40

Learner response or uptake characteristics and quality of uptake are shown in Table 6.

Table 6. Learner Response or Uptake and Quality of Uptake in Text-based and Voice Chats

Dyads	Learner Response		Type of Uptake	
	Uptake	No Uptake	Successful	Unsuccessful
Text-based chats				
NS-NNS	9	2	6	3
ANNS-NNS	27	22	21	6
Totals	36	24	27 (75%)	9 (25%)
Voice Chats				
NS- ESL learners	10	1	2	8
ANNS-ESL learners	21	5	12	9
Totals	31	6	14 (45%)	17 (55%)

Raw frequencies are shown for learner uptake or acknowledgement following the provision of negative feedback in text-based chats. There were 36 occurrences of learner uptake, and 24 instances of no uptake. Successful learner uptake, which is defined as instances

of learners' incorporation of targetlike linguistic forms following corrective feedback in subsequent sessions with their tutors, represents 75% of all uptake in text-based chats. Raw frequencies for learner uptake and successful uptake in voice chats are also displayed in Table 6: 31 instances of learner uptake and 6 of no uptake. Uptake that was successfully incorporated by learners represents 45% of all learner uptake. Frequency counts or instances of unsuccessful uptake in voice chats were higher than those found in text-based chats (55% vs. 25%, respectively).

3.1 Summary of findings

Research question 1 asked which modality, whether text-based or voice chat, afforded learners more opportunities for noticing their own non-targetlike output or inaccuracies during learner- and tutor-initiated LREs. Frequency counts obtained show that text-based chats appear to offer learners more opportunities for error noticing and awareness of linguistic forms.

With respect to Research question 2, which asked about the type and quality of negative information or corrective feedback provided by tutors during negotiation work in these SCMC modalities, the results show that in both modalities, text-based and voice chats, NS and ANNS tutors provided direct or explicit forms of corrective feedback to ESL learners. In addition, frequency counts show that inaccurate and ungrammatical negative information or feedback was provided primarily by ANNS tutors to ESL learners.

Research question 3 asked about the occurrence and type of learner response or uptake found in these SCMC sessions. There is evidence of learner response or uptake following the provision of corrective feedback and that over 60% of this uptake was successful as evidenced by learners' use of grammatical/accurate linguistic forms and structures in subsequent sessions with tutors.

4. Discussion

The raw frequency counts obtained for instances of LREs in both text-based and voice chats provide evidence that it is possible for L2 learners to notice their own non-targetlike utterances/output or engage in self-correction while completing learning tasks with tutors. In other words, LREs in both text-based and voice chats represent instances or opportunities where L2 learners consciously reflected on their own language use and paid attention to linguistic form. Not all these LREs resulted in the provision of negative information or corrective feedback by tutors. As Swain & Lapkin (1998) and Williams (1999) have shown, LREs instantiated as questions about the correct spelling, pronunciation, or well-formedness of a word, a tutor's usage of a structure, and instances of self-correction demonstrate that language learners are capable of the cognitive process of noticing in SCMC environments. The frequency counts obtained for LREs in the text-based chat data seem to support the findings of Shekary and Tahririan (2006) who demonstrated that L2 learners in their study, which included 16 Persian EFL learners, do initiate significant numbers of LREs in text-based SCMC. However, these researchers employed a variety of language tasks that had been used in previous studies of interaction: dictogloss, jigsaw, and free discussion tasks.

4.1 Amplifying attention to linguistic form

The text-based and voice chat data analyzed show that these SCMC modalities do indeed amplify learners' attention to linguistic form as noted by Warschauer (1997). They also problematize "current notions of meaning negotiation" in SCMC (Kern, Ware, & Warschauer, 2004: 246). Since adjacency is lacking in text-based chat, implicit forms of corrective feedback such as recasts and partial repetitions may not be noticed by the learner because of the number of intervening turns and types and length of these forms of corrective feedback. This means that the L2 learner will not be able to recognize the type of implicit corrective feedback provided by a tutor and would be unable to compare his/her output in the trigger with the targetlike reformulation provided by his/her interlocutor or tutor (Sauro, 2008).

Despite some of these disadvantages, SCMC investigations of negotiation work and noticing from a sociocultural perspective have shown that learners and their partners use scaffolding to build meaningful learning partnerships and facilitate the language learning enterprise. In this exploratory study, a qualitative analysis of voice and text-based chat data for NS-ESL learner dyad members reveals that NS tutors often spent an average of 15–20 minutes working on some of the free discussion tasks. They seemed to be in a rush to complete the tasks and often ended their sessions abruptly. In contrast, this was not the case with the ANNS tutors. The ANNS tutors spent between 45–60 minutes collaborating on each task with their ESL partners.

The following excerpt illustrates scaffolding as a tutoring strategy that makes it possible for a learner to notice her non-targetlike output, receive new information about the target language, and build on prior knowledge which will eventually allow her to internalize new information. Lupe, the tutor, and Ana, the learner are attentively focused on meaning.

LRE – Excerpt 5 – Explicit Negative Feedback in Free Discussion Task

17Lupe (ANNS):	Did you find any words that were hard for you?
18Ana (ESL):	I love hot souce
19Ana (ESL):	how do you write sauce..? (**self-correction**)
20Ana (ESL):	yes
21Lupe (ANNS):	yes sauce
22Ana (ESL):	amid
23Ana (ESL):	haste
24Lupe (ANNS):	amid means between
25Ana (ESL):	vexations
26Lupe (ANNS):	it sounds poetic doesn't it?
27Ana (ESL):	sometimes when I don't understand the words I just keep on reading to get a better understanding
28Lupe (ANNS):	vexations means (*molestias*) something that really bothers you. (instance of **code switching)**
29Ana (ESL):	here some other ones aridity perennial
30Ana (ESL):	I didn't have time to look them up

With respect to noticing, operationalized as tutor- and student-initiated LREs, frequency counts obtained for both the text-based and voice chat data appear to contradict findings reported by Jepson (2005), who observed the use of repair moves of 10 groups of anonymous NNSs in 5-minute, synchronous chat room sessions (5 text-chat sessions, 5 voice-chat sessions). He obtained significant differences between the higher number of total repair moves made in voice chats and the smaller number in text chats, but found no instances of self-correction. Jepson (2005) argues that repair work in voice chats is often pronunciation-related, and that "self-corrections may be rare because speakers do not notice their errors, and thus would not see the need to correct them." (Jepson, 2005: 89). He adds that self-correction depends on the social context, and that NNS electronic chats may not be conducive to self-correction. Qualitative analysis of the text and voice chat data in this study clearly show that L2 learners engaged in self-correction and initiated repair work.

Regarding the type and quality of corrective feedback or negative information ESL learners obtained from tutors in these synchronous electronic chats, a qualitative analysis of the data reveals that direct or explicit corrective feedback, specifically metalinguistic feedback, such as explanations, provision of information about a linguistic structure, and reformulation of the learner's utterance, was more frequent than indirect or implicit forms of corrective feedback (recasts, repetitions, and negotiations moves). However, this analysis also identified 40 instances of incorrect or inappropriate feedback to learners by ANNSs. The following excerpt illustrates how one of the ANNSs provides inaccurate corrective feedback to an ESL learner. Kay misspells 'all right', leaves out the definite article [the] in line 114, and the indefinite article [a] in 116:

LRE Episode – Excerpt 6 – Voice Chat/Text-Based Chat Combination –
(Tutor is using Talk button and typing at the same time) – Quality of Negative Feedback

114Kay (ANNS):	*alright*, can you see this sentence. Can you say that to me?
	Say after me… 'I have never been in [ ] United States'
115Rob (NNS):	I have never been in [the] United States
116Kay (ANNS):	yes, it sounds good, you just should practice more – you are
	by yourself in a loud voice. I did this when I didn't know either.
	It's easier to learn. When you read English, read in [ ] loud voice.
	o.k. I sent an e-mail to show you how to take your picture.

Learner response or uptake and instances of successful uptake were evident in these two modalities of SCMC. Learner uptake has been reported by SLA researchers investigating the ESL communicative classroom. It is normally linked to patterns of corrective feedback, L2 learners' proficiency levels, and instructional context (Ellis et al., 2001; Ellis et al., 2006; Loewen, 2004; Lyster & Ranta, 1997). In this study, a qualitative analysis of the data shows that when learners are encouraged to collaborate on tasks that include opportunities for them to focus on the pronunciation of specific words in voice chats, subsequent successful uptake of targetlike pronunciation is indeed possible. As to learner uptake in text-based chats, frequency counts show that learners did respond to negative information or corrective feedback, and that instances of successful uptake or learners' incorporation of targetlike linguistic forms were apparent in subsequent exchanges with their tutors.

The importance of successful uptake is stressed in Shekary and Tahririan's (2006) study that focused exclusively on text-based chats. They found that SCMC enhanced the process of noticing and that this was associated with subsequent L2 learning. Furthermore, the results of their statistical analysis showed that successful uptake in LREs was the strongest predictor of correct responses in immediate and delayed post-tests (Shekary & Tahririan, 2006: 570).

5. Conclusion and implications

Recent quasi-experimental studies have examined negotiation work and the provision of different types of corrective feedback in chatrooms by partially replicating previous face-to-face studies in classroom settings. The results have been mixed. For example, Loewen & Erlam (2006) found no statistically significant gains in response to two different types of corrective feedback in chatrooms, but they point out that this might be a productive environment for teaching and research. In their study, learners at a beginning level of L2 proficiency were able to perform communicative tasks in the electronic chatroom. A more recent study replicated Loewen & Erlam's (2006) study and examined two different types of corrective feedback, explicit and implicit, but the results showed no significant gains for either type of feedback on immediate or sustained gains about the specific language forms targeted (Sauro, 2008).

The results of this reanalysis of data from a small-scale descriptive study suggest that the cognitive process of noticing, operationalized as language-related episodes (LREs), is present in different modalities of SCMC, voice and text-based chats. A qualitative analysis of the data showed that learners were capable of momentarily shifting the focus of attention from meaning to linguistic form during completion of various tasks. Since post-tests were not designed to test learners' interlanguage restructuring or incorporation of new linguistic forms, including new vocabulary, we cannot claim that noticing leads to subsequent L2 learning. Despite the limitations of this exploratory study, the results show that noticing in SCMC is possible, as indicated by the number of LREs found in the text-based and voice modalities. L2 learners become aware of non-targetlike features in their output during LREs and are capable of self-correcting and responding to their tutors' corrective feedback. The synchronous online chats appear to maximize opportunities for task collaboration and negotiation of both form and meaning outside the traditional classroom or language laboratory.

We also became aware of some of the limitations of negotiation work between different types of tutors and ESL students. For example, qualitative analysis of the text-based and voice chat data revealed that there were instances of ungrammatical or inaccurate corrective feedback, which raises questions about the possible implications that the provision of this quality of corrective feedback may have on learners' interlanguage development. Learners run the risk of internalizing inaccurate morphosyntactic, lexical, or phonological negative information/corrective feedback provided by ANNS tutors. Thus, timely and appropriate intervention by instructors or other competent native speakers of the language is needed.

This is something that has not been successfully addressed in SLA studies of negotiated interaction and corrective feedback in classroom settings or language labs.

Corrective feedback or negative information was most often provided by both NS and ANNS tutors immediately after a learner's non-targetlike utterance or written output. Immediate turn sequences in these online chats may be beneficial to L2 learners. This appears to be in keeping with observations made by SLA researchers who have often emphasized the importance of providing immediate negative feedback to learners in order to minimize the possibility of their internalizing non-targetlike features into their evolving L2 grammars (Doughty & Long, 2003).

Online learning environments have evolved and there has been an increase in the popularity of voice chats given recent software improvements made by providers of free online chats (*AIM, MSN, NetMeeting, SKYPE, Windows Live, and Yahoo! IM)*. This may encourage more language-related studies in text-based, voice, and video chats. Therefore, this and other recent studies of SCMC have important pedagogical implications. Language practice is possible in these online environments, but language instructors and tutors need to use appropriate tasks, such as two-way information gap tasks, which Pica, Kang, & Sauro (2006: 301) have identified as encouraging attention to form, function, and meaning that often result in modified interaction among learners and their interlocutors. LREs in SCMC afford learners opportunities to notice non-targetlike features in their own output. However, learners' momentarily shifting attention to form while negotiating meaning during a task completion activity with their partners, is not enough because online chatting does not necessarily result in the incorporation or learning of a targeted linguistic form. Even if it is shown in the chatlogs or transcribed voice chats that a modified linguistic form is subsequently used in the completion of other tasks, this does not constitute evidence that a specific linguistic item has been successfully acquired/learned.

Chatlogs and transcribed voice chats allow learners and tutors to reflect on corrective feedback provided in response to an error in the learner's output or in reaction to a learner-initiated clarification request regarding lexical or morphosyntactic confusion. In general, text-based and transcribed voice chat data are also helpful to instructors and teachers-in-training because they provide evidence of a learner's developmental stage and evolving interlanguage grammar. Tutors need to be aware of learners' developmental errors and feedback needs. These data can be used in preparing lessons or developing new simple and complex online tasks.

Though these LREs were not divided into reactive or preemptive episodes in this study because of the small sample size, the results of learner uptake and successful uptake support Shekary and Tahririan's (2006) findings concerning the nature of reactive LREs in text-based chats which focus primarily on grammar. The authors argue that learners are more likely to repeat the same successful uptake of grammatical structures in subsequent turns and thus reinforce previous noticing and incorporation of the correct linguistic form provided by tutors reacting to a learner's error (Shekary & Tahririan, 2005: 569).

Some final observations are in order with respect to the use of ANNS tutors in SCMC. Face-to-face L2 research by Pica, Lincoln-Porter, Paninos, and Linnel (1996) has shown that NNS partners cannot meet the input or negative feedback needs of ESL learners as

well as native speaker partners can. However, Sicola's (this volume) recent study on NNS-NNS negotiated interaction and attention to phonological form shows that indeed "NNSs can provide each other with corrective feedback and push each other towards more targetlike production and perception of L2 phonological forms" (Sicola, this volume). As the results of this study show, NS tutors were unwilling to provide explicit negative feedback to learners, even though they were familiar with SLA research in classroom settings that highlights the value of the occasional focus on form. It is possible that NS tutors did not want to discourage learners from using their second language by providing explicit negative feedback. Additionally, they were familiar with NNS speech and were thus able to understand the intended meaning of the learners' utterances or text chat output. This allowed them to fill in gaps in learner output, a fact described by Gass and Varonis (1984) in a study of NS-NNS interaction.

It is possible that new developments in technology will facilitate the language learning venture and change modes of instructional delivery. Future research into various modalities of SCMC needs to examine not just language learning outcomes in and outside classroom contexts that focus on specific grammatical structures but also the acquisition of L2 pragmatics. Language learning is a global phenomenon and substantial benefits for learners might accrue as these new technologies evolve and expand, giving them choices and alternative ways of learning.

References

Basturkmen, Helen, Shawn Loewen & Rod Ellis (2002). Metalanguage in focus on form in the communicative classroom. *Language Awareness*, *11*, 1–13.

Beauvois, Margaret Healy (1992). Computer-assisted classroom discussion in the foreign language classroom: Conversation in slow motion. *Foreign Language Annals*, *25*, 455–464.

Belz, Julie A., & Steven L. Thorne (2006). Introduction: Internet-mediated intercultural foreign language education and the intercultural speaker. In J. A. Belz & S. L. Thorne. (Eds.), *Internet-mediated intercultural foreign language education* (iix–xxv). Boston, MA: Heinle & Heinle.

Blake, Robert (2000). Computer mediated communication: A window on L2 Spanish Interlanguage. *Language Learning & Technology*, *4*(1), 120–136.

Cowan, Nelson (1995). *Attention and memory: An integrated framework*. Oxford Psychology Series #26. New York: Oxford University Press.

Darhower, Mark (2002). Instructional features of synchronous computer-mediated communication in the L2 class: A sociocultural case study. *CALICO Journal*, *19*, 249–277.

Doughty, Catherine J., & Michael H. Long (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning & Technology, 7*(3), 50–80.

Doughty, Catherine, & Elizabeth Varela (1998). Communicative focus on form. In Doughty, Catherine & Jessica Williams (Eds.), *Focus on form in classroom second language acquisition* (114–138). New York: Cambridge University Press.

Doughty, Catherine, & Jessica Williams (1998). *Focus on Form in classroom second language acquisition*. Cambridge: Cambridge University Press.

Ellis, Rod, Helen Basturkmen & Shawn Loewen (2001). Learner uptake in communicative ESL lessons. *Language Learning*, *51*(2), 281–318.

Ellis, Rod, Helen Basturkmen & Shawn Loewen (2002). Doing focus on form. *System*, *30*, 419–432.

Gass, Susan M., & Evangeline Arlos Varonis (1984). The effect of familiarity on the comprehension of non-native speakers. *Studies in Second Language Acquisition*, *7*, 37–57.

Gass, Susan M., & Evangeline Varonis (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition*, *16*, 283–302.

Gass, Susan M., Alison Mackey & Teresa Pica (1998). The role of input and interaction in second language acquisition: Introduction to the special issue. *Modern Language Journal*, *82*, 299–307.

Iwashita, Noriko (2003). Negative feedback and positive evidence in task-based interaction: Differential effects on L2 development. *Studies in Second Language Acquisition*, *25*(1), 1–36.

Jepson, Kevin (2005). Conversations – and negotiated interaction – in text and voice chat rooms. *Language Learning & Technology, 9* (3), 79–98.

Kern, Richard, Paige Ware & Mark Warschauer (2004). Crossing frontiers: New directions in online pedagogy and research. *Annual Review of Applied Linguistics*, *24*, 243–260.

Levy, Mike, & Claire Kennedy (2004). A task-cycling pedagogy using stimulated reflection and audio-conferencing in foreign language learning. *Language Learning &Technology*, *8*(2), 50–68.

Loewen, Shawn (2004). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, *27*(3), 361–386.

Long, Michael H. (1991). Focus on Form: A design feature in language teaching methodology. In Kees de Bot, Ralph B. Ginsberg & Claire Kramsch (Eds.) *Foreign language research in cross-cultural perspectives* (39–52). Amsterdam: John Benjamins.

Long, Michael H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of research on second language acquisition* (413–469). New York: Academic Press.

Long, M. H., & P. Robinson (1998). Focus on form: Theory, research, and practice. In Doughty, Catherine, & J. Williams (Eds.), *Focus on form in classroom SLA* (15–41). Cambridge: Cambridge University Press.

Lyster, Roy, & Leila Ranta (1997). Corrective feedback and learner uptake. Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*, 37–66.

Mackey, Alison (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21,* 557–587.

Mackey, Alison, & J. Goo (2007). Interaction research: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (407–452). Oxford: Oxford University Press.

Morris, Frank (2005). Child-to-child interaction and corrective feedback in a computer mediated L2 Class. *Language Learning & Technology*, *9*(1), 29–45.

Pica, Teresa (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning*, *44*, 429–527.

Pica, Teresa, Felicia Lincoln-Porter, Diana Paninos & Julian Linnell (1996). Language learner interaction: How does it address the input, output, and feedback needs of second language learners? *TESOL Quarterly, 30,* 59–84.

Pica, Teresa, Hyun-Sook Kang & Shannon Sauro (2006). Information gap tasks: Their multiple roles and contributions to interaction research methodology. *Studies in Second Language Acquisition*, *28*(2), 301–338.

Robinson, Peter (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning*, *45*, 283–331.

Robinson, Peter (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning*, *47*(1), 46–99.

Sauro, Shannon (2008). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning & Technology*, *13*(1), 96–120.

Schmidt, Richard W. (Ed.) (1995). *Attention and awareness in foreign language learning.* Honolulu, HI: University of Hawai'i, Second Language Teaching & Curriculum Center (Technical Report No. 9).

Shehadeh, Ali (2001). Self- and other-initiated modified output during task-based interaction. *TESOL Quarterly, 35,* 433–454.

Shekary, M., & M. H. Tahririan (2006). Negotiation of meaning and noticing in text-based online chat. *The Modern Language Journal*, *90*, 557–573.

Smith, Bryan (2003). Computer-mediated negotiated interaction: An expanded model. *The Modern Language Journal*, *87*(1), 38–54.

Smith, Bryan (2005). The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *TESOL Quarterly, 39,* 33–58.

Smith, Bryan (2008). Methodological hurdles in capturing CMC data: The case of the missing self-repair. *Language Learning & Technology*, *12*(1), 85–103.

Smith, Bryan, & Greta J. Gorsuch (2004). Synchronous computer mediated communication captured by usability lab technologies: New interpretations. *System, 32*, 553–575.

Sotillo, Susana (2000). Discourse functions and syntactic complexity in synchronous and asynchronous communication. *Language Learning & Technology*, *4*(1), 82–119.

Sotillo, Susana (2005). Corrective feedback via instant messenger learning activities in NS-NNS and NNS-NNS dyads. *CALICO Journal*, *22*(3), 467–496.

Swain, Merrill (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In Susan Gass & Carolyn Madden (Eds.), *Input in second language acquisition* (235–253). Rowley, MA.: Newbury House.

Swain, Merrill (1995). Three functions of output in second language learning. In G. Cook, Guy & Barbara Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (125–144). Oxford: Oxford University Press.

Swain, Merrill (1997). Collaborative dialogue: Its contribution to second language learning. *Revista Canaria de Estudios Ingleses*, *34*.

Swain, Merrill (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (97–114). Oxford: Oxford University Press.

Swain, Merrill, & Lapkin, Sharon (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*, 371–391.

Varonis, Evangeline, & Susan M. Gass (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics*, *6*, 71–90.

Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal, 13,* 7–26.

Warschauer, M. (1997). Computer-mediated collaborative learning: Theory and practice. *Modern Language Journal*, *81*, 470–481.

Warschauer, M., & R. Kern (Eds.). (2000). *Network-based language teaching: Concepts and practice.* Cambridge: Cambridge University Press.

Williams, J. (1999). Learner-generated attention to form. *Language Learning*, *49*, 583–625.

Williams, J. (2001). The effectiveness of spontaneous attention to form. *System*, *29*, 325–340.

LREs

Index

In the series *Converging Evidence in Language and Communication Research* the following titles have been published thus far or are scheduled for publication:

14 **STEEN, Gerard J., Aletta G. DORST, Julia Berenike HERRMANN, Anna A. KAAL, Tina KRENNMAYR and Tryntje PASMA:** A Method for Linguistic Metaphor Identification. From MIP to MIPVU. *Expected June 2010*

13 **PÜTZ, Martin and Laura SICOLA (eds.):** Cognitive Processing in Second Language Acquisition. Inside the learner's mind. 2010. vi, 373 pp.

12 **ZLATEV, Jordan, Timothy P. RACINE, Chris SINHA and Esa ITKONEN (eds.):** The Shared Mind. Perspectives on intersubjectivity. 2008. xiii, 391 pp.

11 **LEWANDOWSKA-TOMASZCZYK, Barbara (ed.):** Asymmetric Events. 2008. xii, 287 pp.

10 **STEEN, Gerard J.:** Finding Metaphor in Grammar and Usage. A methodological analysis of theory and research. 2007. xvi, 430 pp.

9 **LASCARATOU, Chryssoula:** The Language of Pain. Expression or description? 2007. xii, 238 pp.

8 **PLÜMACHER, Martina and Peter HOLZ (eds.):** Speaking of Colors and Odors. 2007. vi, 244 pp.

7 **SHARIFIAN, Farzad and Gary B. PALMER (eds.):** Applied Cultural Linguistics. Implications for second language learning and intercultural communication. 2007. xiv, 170 pp.

6 **DEIGNAN, Alice:** Metaphor and Corpus Linguistics. 2005. x, 236 pp.

5 **JOHANSSON, Sverker:** Origins of Language. Constraints on hypotheses. 2005. xii, 346 pp.

4 **KERTÉSZ, András:** Cognitive Semantics and Scientific Knowledge. Case studies in the cognitive science of science. 2004. viii, 261 pp.

3 **LOUWERSE, Max M. and Willie van PEER (eds.):** Thematics. Interdisciplinary Studies. 2002. x, 448 pp.

2 **ALBERTAZZI, Liliana (ed.):** Meaning and Cognition. A multidisciplinary approach. 2000. vi, 270 pp.

1 **HORIE, Kaoru (ed.):** Complementation. Cognitive and functional perspectives. 2000. vi, 242 pp.