Linguistik Aktuell Linguistics Today 144

Towards a Derivational Syntax

Survive-minimalism

Edited by

Michael T. Putnam

John Benjamins Publishing Company

Towards a Derivational Syntax

Linguistik Aktuell/Linguistics Today (LA)

Linguistik Aktuell/Linguistics Today (LA) provides a platform for original monograph studies into synchronic and diachronic linguistics. Studies in LA confront empirical and theoretical problems as these are currently discussed in syntax, semantics, morphology, phonology, and systematic pragmatics with the aim to establish robust empirical generalizations within a universalistic perspective.

General Editors

Werner Abraham
University of Vienna / Rijksuniversiteit Groningen

Elly van Gelderen
Arizona State University

Advisory Editorial Board

Cedric Boeckx
ICREA/Universitat Autònoma de Barcelona

Guglielmo Cinque
University of Venice

Günther Grewendorf
J.W. Goethe-University, Frankfurt

Liliane Haegeman
University of Lille, France

Hubert Haider
University of Salzburg

Terje Lohndal
University of Maryland

Christer Platzack
University of Lund

Ian Roberts
Cambridge University

Lisa deMena Travis
McGill University

Sten Vikner
University of Aarhus

C. Jan-Wouter Zwart
University of Groningen

Volume 144

Towards a Derivational Syntax. Survive-minimalism
Edited by Michael T. Putnam

Towards a Derivational Syntax

Survive-minimalism

Edited by

Michael T. Putnam
Carson-Newman College

John Benjamins Publishing Company
Amsterdam / Philadelphia

∞™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Towards a derivational syntax : Survive-minimalism / edited by Michael T. Putnam.
p. cm. (Linguistik Aktuell/Linguistics Today, ISSN 0166-0829 ; v. 144)
Includes bibliographical references and index.
1. Grammar, Comparative and general--Syntax. 2. Generative grammar. I. Putnam, Michael T.
P295.P885 2009
415--dc22 2009010636
ISBN 978 90 272 5527 3 (HB; alk. paper)
ISBN 978 90 272 8941 4 (EB)

© 2009 – John Benjamins B.V.
No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

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

List of contributors

Gema Chocano
Departamento de Filología Inglesa
Universidad Autónoma de Madrid
Francisco Tomás y Valiente, 1
28049 Madrid
Spain
gema.chocano@uam.es

Kristin M. Eide
Department of Scandinavian Languages
NTNU, Dragvoll
7491 Trondheim
Norway
kristin.eide@hf.ntnu.no

Greg Kobele
Institut für deutsche Sprache und Linguistik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Germany
kobele@rz.hu-berlin.de

Winfried Lechner
University of Athens
Department of German Language and Literature
Panepistimioupoli, GR-15784 Athens, Greece
wlechner@otenet.gr

Omer Preminger
Department of Linguistics and Philosophy
The Massachusetts Institute of Technology
32 Vassar St., bldg. 32-D808
Cambridge, MA 02139
USA
omerp@mit.edu

Maria del Carmen Parafita Couto
ESRC Centre for Research on Bilingualism in Theory and Practice
Bangor University College Road Bangor, Gwynedd
LL57 2DG
Wales, UK
c.parafita@bangor.ac.uk

Michael T. Putnam
Carson-Newman College
Department of Foreign Languages & Linguistics
341 Henderson Humanities Building
1646 Russell Ave.
Jefferson City, TN 37760, USA
mputnam@cn.edu

Thomas Stroik
University of Missouri-Kansas City
Department of English
Cockefair Hall 115
Kansas City, MO 64110
USA
stroikt@umkc.edu

John te Velde
Department of Foreign Languages and Literatures
Oklahoma State University
Gundersen 309
Stillwater, OK 74078-1054, USA
john.te_velde@okstate.edu

Elly van Gelderen
Arizona State University
Department of English
Tempe, AZ 85287-0302, USA
ellyvangelderen@asu.edu

Preface

The contributions contained in this volume have originated from two sources. The first of them being the *Exploring Crash-Proof Grammars* conference held at Carson-Newman College in Jefferson City, TN on February 29-March 2, 2008. I am grateful for the financial support of the School of Humanities at Carson-Newman College for making this conference possible. I am also thankful to my friends and colleagues here at Carson-Newman who helped me organize and conduct this conference; namely, Mary Baldridge, Maria Clark, Richard Gray, Zachary Feltner-Reichert and Kip Wheeler. Throughout the weekend, many conflicting definitions were introduced and debated on how to interpret what a 'crash' is in Minimalist syntax and whether or not pursuing the idea of a crash-proof model of syntactic theory is a worthwhile venture. Since the Minimalist Program refers to a family of theoretical approaches that share core assumptions and guiding intuitions rather than a rigid research program that does not allow much internal variation, one of the ideas that gained much attention and traction was the notion of a purely derivational syntactic theory. The papers in this volume investigate the strengths and weaknesses of one particular instantiation of a purely derivational theory, Survive-minimalism. Survive-minimalism inherits its name from Stroik, whose formulation of the Survive Principle calls upon a more active role of the Numeration than previously assumed in other flavors of Minimalism. As demonstrated by the arguments put forward in this volume, this research program, although in an infant stage, shows great promise in delivering new and exciting perspectives to problems that have haunted the standard theory for decades. The second way in which participants for this volume were found was through conference presentations and papers that I either heard in person or stumbled upon, or, conversely, by scholars who had come into contact with work that Tom Stroik and I had written on or presented pertaining to Survive-minimalism. Regardless of how everyone became involved with the project, it was a great pleasure working with everyone through the various stages of production for this volume. These papers have enriched my knowledge of syntactic theory and expanded the horizons of the current research program of phenomena and topics handled to date in Survive-minimalism. It was an honor to work with these brilliant scholars

and I only wish to express my gratitude for their hard work, without which this volume would not be possible. Lastly, I would like to recognize Werner Abraham, Elly van Gelderen and Kees Vaes for their interest in this project as well as their insightful comments.

Michael T. Putnam
Jefferson City, TN
June, 2009

PART I

Introduction

Traveling without moving

The conceptual necessity of Survive-minimalism

Michael T. Putnam & Thomas Stroik
Carson-Newman College/University of Missouri-Kansas City

1. Introduction

According to Boeckx (to appear: 1), "the Minimalist Program (MP) refers to a family of theoretical approaches that share core assumptions and guiding intuitions. Even though there is, within the MP, a great deal of theoretical variation (mostly based on the articulation of core assumptions), if a particular theoretical approach embraces the core aspects of the MP, it can be classified as "minimalist" to the extent that it advocates the fundamental desiderata of this program." Below we sketch out these "core assumptions and guiding intuitions" found in the MP:[1]

– Language is a component of nature: Language is regarded as a "biological endowment." There is an innate language faculty in the human brain (or mind, in an abstract description) possessing an array of capacities specifically dedicated to language acquisition, language understanding and language use (cf. MP: 167).
– The language faculty (F_L) is comprised of (at least) two core components: (i) a *cognitive component* and (ii) a *performance component*.
– The *cognitive component* (also referred to as the syntactic component) consists of (i) a computational system and (ii) a lexicon. Based on experience with a particular language (Principle Linguistic Data — PLD), the cognitive component can develop into a mature form, I-language, which can generate an infinite set of syntactic descriptions (SDs).

1. Although many — if not most — of these ideas and concepts are widely held in the minimalist literature, this particular list is an adapted form of Brian Huen's "The Basis of Language in the Human Mind\Brain" (http://www.personal.kent.edu/~pbohanbr/Webpage/New/Huen/mind-brain.html).

– The *performance component* includes those parts of *articulatory-perceptual* (A-P) system and the *conceptual-intentional* (C-I) system that are relevant to language use. The performance systems of F_L interface with the cognitive component and are not affected by the idiosyncrasy of any particular language.
– The cognitive component feeds the performance component: SDs generated are presumably "instructions" for the A-P and C-I systems to perform their tasks. SDs generated by the *cognitive component* of F_L must be "interpreted" by these "external performance" systems. An SD must have two levels of representation (i.e., sound and meaning), due to the fact that F_L has the A-P and C-I systems, and only these (telepathy is not a viable means of communication for most humans).
– The design of Human Language should be *optimal* in that it follows from the conceptual needs of the performance component.

Any (derivational) syntactic theory that adheres to these core minimalist assumptions can be counted among those theoretical approaches that belong to the MP.

In his earliest minimalist analyses, Chomsky took as his primary challenge to develop a design for the computational system of human language (C_{HL}) that follows from conceptual necessity. For Chomsky, the F_L cannot exist without the following design features: some access to lexical items LIs, operations that construct larger syntactic objects SOs from LIs and other SOs, and interfaces that interpret the form and meaning of the SOs. This suggests that the Lexicon (and perhaps, the Numeration), External Merge, and the C-I and A-P interfaces are all conceptually necessary to derive and use Human Language. In addition, since languages exhibit displacement effects in which a single SO can participate, in highly restricted ways, in more than one syntactic relationship, Chomsky contends that F_L must have displacement operations as well as some constraints on these operations. To this end, he currently proposes that F_L must possess *Internal Merge* (a displacement operation) and some machinery to limit the displacement domain of Internal Merge — including ontological/representational units called (strong) phases, the Phase Impenetrability Condition (PIC) (cf. Chomsky 2001), and the operations of Transfer and (multiple) Spell-Out. Although all of this displacement machinery serves to "account for" non-local (i.e., non-sisterhood) relations, this explanation is an ***operational*** explanation, not a ***conceptual*** one; that is, it "accounts for" how non-local structural relations might work, but not why these relations *must* exist in Human Language. Explaining displacement effects in terms of displacement operations and constraints on these operations can at best be an empirical (operational) description of the effects, but it cannot conceptually explain the nature of ***displacement*** itself since it presupposes ***displacement***. Hence, despite Chomsky's commitment to providing a design for F_L grounded in conceptual efficiency, he has been unable to expunge operational (i.e., empirical) explanations from his design principles.

There is one variant of the MP; however, that eliminates operational explanations from the design of Human Language. This variant of the MP is based on Stroik's (1999, 2000, 2009) *Survive Principle*, which delimits structure building operations to Numeration-to-Derivation mappings. Purged of non-local, displacement operations (and all the machinery required to support these operations), this version of the MP, which we call *Survive*-minimalism (following Putnam (2007)), utilizes only conceptually necessary design elements: the external interfaces, External Merge, and the Lexicon/Numeration. Critically, by eschewing mechanisms such as Internal Merge and phases that — as argued by Stroik (2000, 2009) and Putnam & Stroik (to appear) — substitute operational explanations for conceptual ones, *Survive*-minimalism is able to offer a design for Human Language that follows from conceptual necessity. In putting forth such bold claims about *Survive*-minimalism, we realize that the burden of proof rests solely on our shoulders, which provides the primary goal of this introduction. In the sections that follow, we first present our case for the lack of conceptual grounding for some fundamental operations and mechanisms in standard minimalist theory that have simply been "taken for granted"; and we then argue that *Survive*-minimalism circumvents these highly unwanted circular, operational aspects of the MP. The introduction concludes with a discussion of the scope and content of the contributions in this volume that advance our conceptual understanding and empirical coverage of *Survive*-minimalism and its interaction with Human Language.

2. The computational necessity of displacement operations

2.1 Move/internal merge

One of the most enduring assumptions of generative grammar has been that syntactic objects (i.e., constituents) can participate in multiple dependencies within the same grammatical string. Take, for example, sentence (1):

(1) a. What$_i$ did John read t_i?

In example (1), the *wh*-item *what* is simultaneously the direct object of the predicate *read* and the *wh*-item that resides at the leftmost periphery of the clause. Traditional Chomskyan (1965) analyses of the observed phenomenon in (1) postulate that the *wh*-item is physically present in both positions at different stages in the derivation of the sentence. These analyses explain the multiple dependencies of this single item by invoking transformations/movement rules. Over the last four decades, although standard generative theories of syntax have undergone many versions and many revisions, all these approaches have continued to adhere to some notion of transformational rule as a driving force in the theory of F_L, culminating, perhaps, in the GB syntactic model (see Chomsky (1981)) that allowed constituents to move freely as long as the movement satisfied specified well-formedness conditions (Subjacency Condition, Empty Category Principle, etc.). Since the advent of the Minimalist Program (see especially Chomsky (1995)), syntax has been re-conceived: syntactic rules, once the ontological ground of grammar, are now subject to the conceptual requirements of the interfaces. That is, syntactic rules/operations are permitted within the design of Human Language if and only if they are conceptually necessary for the interfaces to execute their interpretative processes. Some rules formerly included within a theory of syntax, such as phrase structure rules, have been abandoned; however, not all of them have disappeared. One, in particular, has persisted in Minimalist syntax — the transformation. It is now called *Internal Merge*, but it is a syntactic operation that moves constituents from one structural position in a derivation to another structural position. Although having "transformations" in Minimalist syntax allows us to explain the multiple appearances of the *wh*-item in (1), the conceptual necessity of having the *Internal Merge* operation is still in question. As pointed out by Stroik (1999, 2009), the argument for the displacement operations in their current form is a circular argument: we point to displacement phenomena to support our position that we need displacement operations in natural language, and we use displacement operations to account for displacement phenomena in natural languages.

The fact that Chomsky's (2004: 110) observation that "displacement is not an imperfection in language" has largely been taken as an argument for the conceptual necessity of both displacement phenomena and displacement operations — which we can see in Chomsky's (2005: 7) remark that "[t]he absence of the [displacement] operation would be an imperfection that has to be explained." Rather than conflate displacement data with displacement operations and use these data to conceptually justify displacement operations, we will look at whether or not displacement operations follow conceptually from the demands of the interfaces. With this in mind, consider the following line of argumentation:

(2) a. Assume that displacement exists (without this assumption, displacement operations will be completely devoid of conceptual necessity). This assumption permits a syntactic object **α** to appear in a given syntactic derivation SD more than once:
SD → [...**α**...**α**...**α**...]

b. Assume that what the Narrow Syntax delivers to the interfaces for interpretation are structured SDs.

c. Question: does it matter to the interfaces how α gets to its various positions?

d. If not, then displacement operations such as *Internal Merge* lack conceptual necessity because displacement itself is not visible to the interfaces. In this situation, the interfaces will interpret each structural merging of **α,** but not the merging operation.

e. If so, then the interfaces must interpret both α (in the various positions it occupies) and the way α gets positioned. However, although syntactic derivations (and the representations they produce) give α in various structural positions, they do not provide any information about the process of getting α from position to position. The process of α-placement, then, is not visible to the interfaces and not interpretable by the interfaces; hence displacement operations have no conceptual necessity.

The essential point we are making here is that if, as is assumed within Minimalism, the only information visible to the interfaces is derivational output, then the derivation processes/operations that are responsible for the various structural appearances of α will not contribute to interface interpretations — which means that these operations will lack conceptual necessity. In other words, if all the interfaces need to see is where α appears and not how it appears where it does, then displacement operations such as *Internal Merge* will not be conceptually necessary and must be expunged from the design elements of Human Language.

So how is it possible to account for displacement data without displacement operations? And, even more importantly, how do we build structure at all if there are no placement operations? These questions are at the heart of the *Survive* framework developed in Stroik (2009) and Putnam and Stroik (to appear). As we discuss below, *Survive*-minimalism builds structure and explains displacement using the one structure-building operation that has conceptual necessity — *External Merge*. For *Survive*-minimalism, the computational system of Human Language does not have any look-back, non-local syntactic operations such as *Internal Merge* (which lack conceptual necessity) and it doesn't have any of the other domain-restricting mechanisms required to limit the computational power of *Internal Merge*, such as the *Phase Impenetrability Condition* (PIC) (Chomsky 2001) and multiple Transfer (which, we shall see, also lacks conceptual necessity).

As we have just argued, a computational system for Human Language that includes displacement operations will necessarily have design flaws. These flaws pertain not only to the fact that *operational* displacement operations are not conceptually necessary, but also to the fact that second-order ***operational*** (as opposed to ***conceptual***) constraints must be placed on displacement operations. Since Ross's (1967) dissertation, it has been recognized that derivation-to-derivation mappings (such as the transformations rules in the 1960's and 1970's, *Move-α* in the GB framework, and *Internal Merge* in the MP framework) are computationally too powerful for Human Language unless their attributes and properties are restricted. Consequently, generative theories of syntax over the last forty years have been dedicated in no small measure to formulating constraints on derivation-to-derivation (or phrase marker-to-phrase marker) operations. In Standard Theory models, transformations had their input and output restricted by the transformational cycle and by various filters on representations (A-over-A Condition, Specified Subject

Condition, etc). In the GB framework, *Move-α* was unconstrained in its application, but its output was constrained by a new set of representational filters (Government Theory, Bounding Theory, Case Theory, etc). The early minimalist framework (see Chomsky 1993, 1995) continued to posit a derivation-to-derivation mapping with the power of transformation rules (*Move*) but used economy-of-derivation conditions to filter the output of this operation (e.g., economy conditions such as Chomsky's (1995) *Shortest Move*, Richards's (2002) *Principle of Minimal Compliance*, and Aoun and Li's (2003) *Minimal Match Condition*). More recently, standard versions of Minimalism have replaced one derivation-to-derivation mapping (*Move*) with another (*Internal Merge)*, but have continued to reduce the input-output domain of this mapping, circumscribing how far constituents in a syntactic derivation can be displaced: constituents cannot be displaced through the Spec of any *phase* (a CP or a *v*P) per application of *Internal Merge*. Two mechanisms have been introduced into the design of HL to ensure that the *Internal Merge* operation is duly constrained — the *Phase Impenetrability Condition* (PIC) and *Phase-Complement Transfer* (PCT). The PIC limits the input-output domain for *Internal Merge* (as described above) and the PCT removes the complements of phase heads from a syntactic derivation and sends them to the interfaces (thereby preventing the *Internal Merge* operation from using any constituents that have undergone Transfer as input for displacement).

What we can conclude from the foregoing discussion is (i) employing displacement operations to account for displacement data, as we noted previously, provides merely an ***operational*** explanation for these data; and (ii) employing displacement constraints such as the PIC or *Shortest Move* to account for the appropriately limited computational power of displacement operations places a second level of ***operational*** explanation on the use of displacement mechanisms. Needless to say, the fact that displacement operations are not ***conceptually*** necessary is already sufficiently problematic to question their viability as a design element in the computational system of HL; however, the fact that these operations necessitate a second level of ***operational*** mechanisms to support them threatens their ***conceptual*** necessity all the more.

2.2 The PIC and TRANSFER

Chomsky (2001, *et seq.*) advances the claim that *phases* (typically, CP or *v*P) are involved in a number of phenomena; for example, successive cyclic movement and multiple Transfer/Spell-Out. The gist of the idea here — as pointed out by Bošković (2003/2007) as well as others — is that a syntactic operation can involve constituent Y in a phase ZP and constituent X not contained in ZP if and only if Y is in the Spec of the phase; this constraint on syntactic operations, which is formally stated as the *Phase-Impenetrability Condition* (PIC), disallows syntactic operations from involving constituent Y contained in the complement of the phase-head Z and constituent X outside of ZP, as is represented in (3) (taken from Bošković 2003/2007: 55; (1)).

(3) [X ... [$_{ZP}$ [Z Y

According to the PIC, the constituent Y is simply "invisible" to X for either *Move/Internal Merge* or *Agree*.

The viability of PIC, however, is not firmly established. Bošković (2003/2007) presents data that suggests that whereas *Move* is subject to the PIC, most instances of *Agree* are not. As a matter of fact, a number of languages allow *Agree* dependencies that clearly violate the PIC. In his article *Agree, Phases and Intervention Effects*, Bošković (2003/2007) provides multiple instances of languages that violate the PIC. One such language that displays these sorts of PIC-violations is Chuckchee, which allows agreement to reach into a finite CP (this data is noted in Inènlikèj and Nedjalkov (1973), also discussed in Mel'̌čuk (1988), Ura (1994), and Stjepanović and Takahashi (2001), cited from Bošković (2003/2007: 57; (2)):

(4) ənan qəlɣ̦ilu ləŋərkə-nin-et [iŋqun Ø-rətəm' ŋəv-nen-at qora-t]
He -inst regrests$_{3PL}$ that $_{3SG}$-lost$_{3PL}$ reindeer-PL(noun)
'He regrets that he lost the reindeer.'

The matrix *v* agrees with the embedded clause object in (4), an *Agree* relation that violates the PIC. Bošković also provides data from several Algonquian languages that also allow agreement into a finite clause, such as Blackfoot (cited again from Bošković (2003/2007: 57; (5)), but also see Frantz (1978) and Legate (2002)).

(5) nít-ssksinoa-**wa** **m**-aníst-sskonata'psspi.
I-know-3 3-manner-strong
'I know how strong he is.'

Of note, both (4) and (5) have long-distance agreement relations that violate the PIC and, as a consequence, these sentences should be ruled out as ill-formed constructions. The grammaticality of these examples, then, is problematic for the PIC.

Recent studies (see Stjepanović & Takahashi (2001), Lee (2003), Fox & Pesetsky (2005), and Bošković (2005)) try to address the foregoing problem by re-analyzing the PIC as a phonological constraint on PF-linearization. Bošković (2003/2007: 74) maintains, in particular, that *phases* induce phonological "freezing effects": "if something will ever move, then it cannot be contained in a unit that is shipped to Spell-Out". That is, Y has to move to Spec,XP (XP a phase) in order not to get caught in a spell-out unit, which would freeze it for pronunciation purposes. Under this analysis, as Fox and Pesetsky (2005), Stjepanović and Takahashi (2001) and Bošković (2005) contend, the PIC does not apply in the Narrow Syntax; rather, it applies at the interfaces to monitor the operation *Move*, ensuring that cyclicity is maintained in displacement operations. The consequence of making the PIC an interface constraint is that the Narrow Syntax can be freed of opaque domains (phases), thereby permitting long-distance *Agree* relations, as well as allowing *Internal Merge* to apply freely.

Although some current versions of the MP that adopt the "Internal Merge is free" approach employ some sort of phase-based and PIC-constrained approach (in the spirit of Chomsky 2000, 2001), there are many variants of the MP that criticize phase-based approaches (see Epstein & Seely (2002), Legate (2002), Boeckx (2003, 2007), and Boeckx & Grohmann (2007), to name a few). Despite Chomsky's claim that phases are somehow "structurally independent" units at the interfaces, Epstein & Seely (2002:78) raise questions about the conceptual necessity of phases: "How can we know that [phases] are *relatively independent* at the interface if Spell-Out applies *before* the interface is reached and without access to interface properties?"

We agree with both Bošković, who argues that phases and the PIC are not applicable to the Narrow Syntax, and Epstein and Seely, who argue that phases and the PIC are not interface material. What would seem to follow from these arguments is that phases and phase-based constraints have no place in the design of Narrow Syntax. This conclusion is correct, we suggest, because phases and phase-based constraints are simply not conceptually necessary in a theory of syntax, especially one that (as alluded to in points laid out in (2) above) relies on a strict notion of interface interpretation of multiple instances of a syntactic unit, in this case α. We can see this if we look carefully at phase-based constraints. For these constraints to be conceptually necessary, they must participate in interface interpretations; that is, they must be visible at the interfaces. Hence, they must constrain what the interfaces "see" — i.e., representations (as Bošković asserts). There are, however, two substantial problems with having interface mechanisms that filter representations. First, the conceptual necessity of filtering representations presupposes that the Narrow Syntax necessarily produces representations that must be filtered. In other words, to require representational filters, it would have to be conceptually necessary (for the interfaces to operate) that the Narrow Syntax derives representations that cannot be used by the interfaces. This notion of conceptual necessity (that the system of Human Language conceptually needs to build structures it can't conceptually use) is patently ridiculous. Second, if the conceptual justification for phased-based constraints is to filter representations (and thereby establish well-formedness conditions on representations) and if well-formed representations are those that satisfy phase-based constraints, then we have a circular relationship between representations and constraints, a relationship incapable of meeting the threshold of conceptual necessity.

Our conceptual arguments against phase-based constraints such as the PIC extend to those other mechanisms required to sustain *Move* or *Internal Merge*, in particular *Phase-Complement Transfer* (PCT) and EPP/Edge-features. PCT is an operation, according to Chomsky (2005) that transfers complements of phase-heads (C or *v*) from a derivation to the interfaces. The essential purpose of the PCT operation is to remove parts of a derivation from the derivation itself so that these parts and the constituents they contain cannot be subsequent targets of the *Internal Merge* operation. In this way, the PCT limits the operational reach of *Internal Merge*. Of note, however, the PCT operation has no interface effects. The phase-complements submitted to the interfaces by the PCT do not undergo interpretation when they are first submitted (if the *v*-complement *hire Pat* is shipped off to the interfaces, how will it be possible to derive the sentence *Hire Pat, I shall;* and if the *v*-complement *hire which woman* is shipped off to the interfaces, how will the interfaces know it should interpret the *wh*-phrase as a *wh*-operator or as a *wh*-in-situ element; and if phase-complements are interpretation-ready, why aren't there phase-complement abortings, such as "given Mary a book"?). If, as we argue here, phase-complements are not ready for interface interpretation, then the PCT lacks a conceptual warrant and should be eliminated from the design of Human Language. A similar analysis applies to EPP/Edge-features. As Chomsky (2005) notes, although these features function to license movement operations, they themselves are not interpreted at the interfaces. That is, they are *operationally*, but not *conceptually*, necessary mechanisms; as such, they are descriptive mechanisms that point to the importance of edges in building structure, but they don't offer an explanation for why edges are important.

If we remove elements such as phases, the PIC, the PCT, and Edge-features from the design of HL, we still have to determine what motivates displacement in an interface-driven, derivational theory of syntax. In what follows, we show how Survive-minimalism divests itself of all unnecessary operations and constraints, and can still provide a derivational model of syntactic theory able to explain displacement.

3. Returning to conceptual necessity: Survive-minimalism

The Minimalist Program is committed to the Strong Conceptual Hypothesis (SCH), which circumscribes the design of HL within the conceptual needs of the performance systems. Under the SCH, all design elements must be ones that are conceptually necessary for the interfaces to operate. As we argue above, the conceptually necessary elements of HL are the Lexicon/Numeration, *External Merge*, and the interfaces themselves. No other elements — including *Internal Merge*, phases, the PIC, phrase structure rules, GB-style sub-theories (Case Theory, etc), or economy conditions such as Shortest Move — are conceptually necessary.

This doesn't mean that these latter elements, or any other versions thereof, are not pertinent to HL. They may, in fact, provide valid operational descriptions of HL — for example, Shortest Move may adequately describe constraints on displacement and displacement operations may operationally describe displacement phenomena. However, if some element Y is not conceptually necessary, then no matter how adequately Y may describe properties of HL, Y itself must be the consequence of the conceptually necessary design elements of HL. That is, all the properties of HL, and all meta-properties we devise to describe the properties of HL, must follow from the conceptually necessary design elements of HL.

Stroik (1999, 2009) investigates what the design of HL should look like if it conforms to the SCH. He argues that this design will have a Lexicon/Numeration, one syntactic operation (*External Merge EM*) that combines syntactic objects, and the interfaces — but nothing else. His design of HL (called Survive-minimalism by Putnam (2007)) locates the Lexicon/Numeration within the interfaces (since learning words and their lexical features must involve the performance systems/interfaces) and it uses the EM operation to combine lexical items LIs selected from the Numeration N, always deriving concatenated items (CIs) that are usable by the interfaces. In Survive-minimalism, LIs are copied from the Lexicon Lex into the Numeration (thereby retaining the original LIs in the Lexicon and not, as a result, ever diminishing one's Lexicon while building a syntactic structure); the EM operation re-copies (an already-copied–from-Lex) LI X in N into a syntactic derivation SD, but only if merging X with syntactic object Y will produce an interface-licit concatenation in which at least one lexical/interface feature of X matches a paired feature of Y. Once X merges into the SD, there are two copies of X within the computational system: one in the SD and one in N. If the X in N has lexical/interface features that have survived the merger of X with Y (because Y could not match these features), then X remains active in N for subsequent (re) merger, via the EM operation, into the SD. X will continue to remerge into an SD as long as it has active (unchecked) features. It is this remerger of X that produces the displacement effects we can see in (6).

(6) What should Peter do (what)

In (6), the *wh*-element *what* will merge into the SD at least twice (see Stroik (1999, 2009) for a detailed discussion of how remergers work).

In Survive-minimalism, the computational system is subsumed within the interfaces and it is designed to meet the conceptual needs of the interfaces: the Lexicon and any Numeration selected from the Lexicon are situated within these performance systems, and the syntactic operations map interface-usable simplex concepts (LIs) onto interface-usable complex concepts (structured CIs). Since the computational system operates on interface-interpretable input to derive interface-interpretable output, this system will be "crash-proof" in the sense of Frampton and Gutmann (2002) and will not, therefore, require any derivational or representational mechanisms to filter unusable output.

But let's take a little closer look at Survive-minimalism. Being a framework that adheres to the core assumptions of the MP outlined previously, Survive-minimalism follows Chomsky's (1995, 2005) assumption that a language L derives (λ, π) representations for sentences that are interpreted at the conceptual-intentional (C-I) and articulatory-perceptual (A-P) interfaces. It initiates this derivation by constructing a Numeration NUM piecemeal — selecting a lexical item X (which bears a CATegory feature α and, let's say, another syntactic feature β) for NUM and placing X in the SD. To build X into an interface-usable CI requires X to merge with another lexical item Y that has an α-matching SUBCATegory feature, as well as other features, including its own CAT feature γ. Once Y is placed in NUM, it will merge with X to form syntactic object K, which will project the γ feature and, consequently will necessitate other LIs to be placed in NUM to build an interface-usable CI for K, and so on. Since X, after it merges with Y, will have a feature β that has not been checked by Y, the copy of X in the Numeration will remain active in NUM, available to remerge into the Narrow Syntax. This remerger will take place when another head (Z) enters the derivation bearing a matching feature with X, namely, the feature β. Under such an analysis, X survives its initial concatenation with Y and must participate in a subsequent concatenation with another head (Z) bearing a matching feature in order to ensure the all of the features of X build CIs that are usable at the interfaces. Stroik formalizes this relationship between the Numeration and a syntactic derivation as the *Survive Principle*:

(7) *The Revised Survive Principle* (based on Stroik 1999: 286)
If X is a syntactic object (SO) in an HP headed by H, and X has an unchecked feature [+ F] that is incompatible with the features of H, X remains active in the Numeration.

Elements in N will continue to merge and remerge into the SD until there are no Xs in N that have active features. At this point, the derivation terminates and is ready for interface interpretation. (For an extended discussion of how derivations begin, continue, and terminate, see Putnam and Stroik (to appear).)

To demonstrate how the computational system works in the Survive-minimalist model, we provide an abbreviated derivation for (8) in (9) — for an expanded discussion of syntactic derivation, see Stroik (2009), Stroik (this volume) and Putnam and Stroik (to appear).

(8) What should Peter do?

(9) a. Merge {do, what} → do what
b. Survive {what}
c. Merge {Peter, {do, what}} → Peter do **what**
d. Survive {Peter}
e. Merge {should, {Peter, {do, what}}} → should **Peter** do **what**
f. Survive {should}
g. Remerge {Peter, {should, {Peter, {do, what}}}} → Peter **should Peter** do **what**
h. Remerge {should, {Peter, {should, {Peter, {do, what}}}}} → should Peter **should Peter** do **what**
i. Remerge {what, {should, {Peter, {should, {Peter, {do, what}}}}}} → what should Peter **should Peter** do **what**
j. Derivation terminates
k. Representation presented to C-I and A-P interfaces for evaluation

The initial application of Merge {do, what} is driven by the necessity to match the categorial feature +D of the *wh*-object with the subcategorization feature of the verb. According to the *Survive Principle,* if a syntactic object bears any additional features not present on the immediately adjacent head (in this case V), it will survive and remain active in the Numeration until all subsequent features have been checked through Remerge. This occurs in (9b): the *wh*-object has a [WH] feature that cannot be matched/checked by the verb; hence, the *wh*-object remains active in NUM for later remerger. Similarly, the Case-feature of the DP **Peter** and the Q-feature of the modal **should** cannot be checked when these elements are originally merged into the derivation, so these syntactic objects also survive for Remerge, as we can see in (9d) and (9f). By (9i), all the features of the syntactic objects in NUM have been appropriately checked, so the derivation will terminate. Needless to say, the example (8) is structurally simple; to see how Survive-minimalism analyzes structurally complex examples, such as those with island effects, see Stroik (2009).

As exhaustively stated in the works of Stroik (2009), Putnam (2007), Stroik (this volume) and Putnam & Stroik (to appear), Survive-minimalism radically departs from mainstream minimalist desiderata in several areas (adapted from Stroik (this volume)):

- Lexical items in the Numeration (NUM) are derivationally selected, not blindly pre-selected by an operation Select to generate sub-arrays of lexical items to be exhausted in the course of a derivation.
- The lexical items in NUM are copied from the Lexicon (LEX) and the lexical items in the syntactic derivation are copied from NUM; hence NUM is never emptied of its lexical material and (λ, π) representations consist of second-order copies of lexical material.

- Lexical items in NUM with active SYNtactic features are available for (Re)merge into the syntactic derivation.
- The available/necessary syntactic operations are External Merge operations — Merge and Remerge — which map lexical material from NUM to DER. Internal Merge/Move which maps Derivation-to-Derivation is not needed.
- There are no movement-based legacies or ontologies. No special boundaries (such as phases or Grohmann's (2003) prolific domains) are necessary.
- No economy conditions that function as filters that police displacement properties are needed; economy conditions fall out as a natural by-product of the Survive Principle.
- Survive-minimalism drastically reduces its machinery to an active Numeration, two variants of External Merge, and the interfaces. This is conceptually more appealing than standard minimalist assumptions in that it is ontologically and computationally simpler in design and function. Survive-minimalism also addresses Brody's (1995, 2002) problem of how to construct a grammar that can be both derivational and non-representational. This is achieved with the removal of Internal Merge and the mapping of iterative derivations. The notion of Merge-alpha is also a non-sequitur in this framework; all syntactic objects created in the syntax are legible at the interfaces and achieve basic semantic compositionality requirements as well. Spell-Out, at least in the form of a mechanism that ships derivationally-generated syntactic units to the interfaces prior to the completion of the entire grammatical string, does not exist in this system. With the removal of (multiple) Spell-Out/Transfer, we also no longer have to worry about issues such as the re-configuration of previously transferred syntactic objects to the interfaces at the completion of the derivation. Survive-minimalism brings us one step closer to the vision of a crash-proof syntax following the ideas of Frampton & Gutmann (2002) and Hinzen (2006) eliminating all remnants of look-ahead and look-back mechanisms in the MP. As can be interpreted from this brief, yet concise, introduction to the potential conceptual advantages of Survive-minimalism, it appears that this framework addresses many problematic issues that plague standard minimalism as outlined at the end of the previous section.

4. Scope and content of this volume

The contributions in this volume explore the advantages and challenges that accompany instantiations of the MP that employ the Survive Principle. Importantly, these chapters are not solely focused on the notion of syntactic 'displacement', but rather address other pressing elements and mechanisms in Minimalist theory such as the relationship among the Lexicon, Numeration, and the Narrow Syntax, the operation *Select*, language change/grammaticalization, bilingualism and code-switching, syntax-semantic interface issues such as temporal chains and type-shifting, ellipsis and derivational identity, (a)symmetries in coordinate structures, long-distance agreement and the structure of the left periphery in West Germanic.

This volume begins with Thomas Stroik's analysis of the relationship between the lexicon/numeration and syntactic derivations in Survive-minimalism in his piece entitled *The Numeration in Survive-minimalism*. In contrast to mainstream approaches to minimalist theory, Stroik engages in a discussion of the often under-discussed pitfalls of assuming the Numeration is used to construct a Lexical Array prior to a syntactic derivation. On the contrary, Stroik champions a Numeration that must be compiled throughout a syntactic derivation and not selected prior to the derivation. The consequence of having a derivationally built Numeration is that the Numeration becomes the domain for both the Merge and Remerge operations, thereby obviating the need for the Internal Merge operation.

Omer Preminger explores the possibility of accounting for long-distance agreement phenomenon without the assistance of Probe-Goal relations in his paper, *Long-distance agreement without Probe-Goal relations*. Here Preminger explores two possibilities to derive agreement-at-a-distance in a Minimalist system void of Probe-Goal relations. The first possibility is that agreement is established in purely local (i.e., sisterhood) relations, followed by the agreeing head moving to a higher position, providing the impression of long-distance movement. Given the computational difficulties associated with the first option, Preminger opts for a second option for long-distance movement without Probe-Goal relations; namely, a system where the phonological component of grammar (PF) is capable of pronouncing the lower copy of a moved element. Such an approach is similar in scope and function to Bobaljik's (2002) *Lower Right Corner* effects.

Gema Chocano's *Musings on the left periphery in West Germanic: German Left Dislocation and Survive* presents an analysis which combines the findings in recent work on the construction with Survive-minimalism, a model that dispenses with the existence of the EPP-feature. The core component of Chocano's argument is that the Merge of both the D-pronoun and the left-dislocated XP with C is triggered by the presence of a logico-semantic feature for referentiality [+ Ref], similar to Putnam's (2007) account of middle field scrambling in West Germanic and Stroik's (2000, 2009) approach of English *wh*-constructions with a pair-list reading.

The introduction of the Survive Principle to account for constituent 'displacement' also forces us to focus on the interpretation of movement chains/legacies and traditional empirical puzzles and their respective theoretical treatments in the

Survive-minimalist framework. In *Tense, finiteness and the survive principle: Temporal chains in a crash-proof grammar*, Kristin Eide elucidates how temporal chains are construed in a syntactic structure in a crash-proof model of syntactic derivation.

Michael Putnam & Maria del Carmen Parafita Couto's contribution, entitled *When grammars collide: Code-switching in Survive-minimalism,* analyses code-switching data from a Survive-minimalist perspective centering on the internal DP structure (e.g., explicitly looking at the ordering of Det(erminers)-Adj(ectives)-N(ouns)) in Spanish-German bilinguals. Putnam & Parafita Couto advance the argument that, due to the dominant trend in Survive-minimalism to eliminate the interpretable/uninterpretable feature distinction on LIs, the question of article selection and gender agreement in mixed DPs (i.e., Spanish article-German noun and German article-Spanish noun) can best be understood as a conflict that is resolved by the operation Select with the help of soft (violable) constraints. Critically, in their analysis the Merge operation does not need to be complicated nor are Probe-Goal relations (*Agree*) necessary for the valuing of features.

John te Velde expands his research on coordinate structures in the MP in his contribution entitled, *Using the survive principle for deriving coordinate (a)symmetries.* Building on his previous work (2005 et seq.), te Velde points out the fault in a phase-based model of computation where the selection of lexical items (LIs) takes place prior to Merge, an ordering that results in the failure of a matching operation prior to Merge. Te Velde presents an alternative approach to the construction of coordinate structures that uses the Survive Principle where LIs are selected as needed for the Merge operations of coordinate structures. The selection process of LIs to enter to Narrow Syntax is guided by algorithms that map certain features from a leading conjunct onto the next conjunct participating in the next instance of Merge. Te Velde delivers a conclusive argument that this revised approach to coordinate (a)symmetries in coordinate structures no longer requires matching to take place across conjunctions.

The final three papers in this volume explore the effects outside of traditional issues of constituent 'displacement' that would be significantly altered with the addition of the Survive Principle into minimalist desiderata. Two papers, Greg Kobele's *Syntactic identity in Survive-minimalism: Ellipsis and the Derivational Identity Hypothesis* and Winfried Lechner's *Evidence for survive from covert movement,* address issues pertaining to the relationship between the Narrow Syntax and the external performance interfaces. Kobele's contribution explicates upon the power contributions that ellipsis structures can provide into the underlying mechanisms of grammar. Kobele makes use of the Derivational Identity Hypothesis (DIH) in tandem with the Survive Principle to demonstrate that linguistic formalisms ideally should develop rich derivational structures to account for ellipsis data. Similar in some respects to Preminger's discussion of push-chain movement chains, Lechner interprets the Survive Principle (and its ilk) as being responsible for motivating displacement effects to reconcile semantic type incompatibility at LF. Secondly, with respect to configurations warranting multiple covert movement operations for scope interpretation in double object and inverse linking constructions, Lechner points out the advantages that Survive-minimalism delivers by expressing ordering restrictions between different movement types (i.e., Case-driven movement vs. Quantifier Raising (QR)) in a natural way.

The final contribution in this volume, Elly van Gelderen's *Language change and survive: Feature economy in the Numeration,* focuses on language change, paying particular attention to the process of grammaticalization whereby certain grammatical features are often lost or reanalyzed, where they then undergo semantic bleaching and grammaticalize. In contrast to much current work in the Survive-minimalist framework to date, van Gelderen argues in favor of maintaining uninterpretable features in the Numeration and Narrow Syntax (contra Stroik & Putnam 2005).

References

Aoun, J. & Li, Y.-H.A. 2003. *Essays on the representational and derivational nature of grammar: The diversity of wh-constructions.* Cambridge MA: The MIT Press.

Bobaljik, J. 2002. A-chains at the PF interface: Copies and 'covert' movement. *Natural Language and Linguistic Theory* 20: 197–267.

Boeckx, C. 2003. *Islands and chains.* Amsterdam: John Benjamins.

Boeckx, C. 2007. *Understanding minimalist syntax: Lessons from locality in long-distance dependencies.* Oxford: Blackwell.

Boeckx, C. To appear. A tale of two minimalisms: Reflections on the plausibility of crash-proof syntax, and its free-merge alternative. In *Exploring crash-proof grammars*, M. Putnam (Ed.). Amsterdam: John Benjamins.

Boeckx, C. & Grohmann, K. 2007. Putting phases into perspective. *Syntax* 10: 204–222.

Bošković, Ž. 2003/2007. Agree, phases, and intervention effects. *Linguistic Analysis* 33(1–2): 54–96.

Bošković, Ž. 2005. On the locality of Move and Agree: The big spring cleaning. Ms, University of Connecticut, Storrs.

Brody, M. 1995. *Lexico-logical form: A radically minimalist theory.* Cambridge MA: The MIT Press.

Brody, M. 2002. On the status of representations and derivations. In *Derivation and explanation in the minimalist program*, S. Epstein & T.D. Seely (Eds), 19–41. Oxford: Blackwell.

Chomsky, N. 1965. *Aspects of the theory of syntax.* Cambridge MA: The MIT Press.

Chomsky, N. 1981. *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, N. 1993. A minimalist program for linguistic theory. In *The view from building 20: Essays in Honor of Sylvain Bromberger*, K. Hale & S.J. Keyser (Eds), 1–52. Cambridge MA: The MIT Press.

Chomsky, N. 1995. *The minimalist program.* Cambridge MA: The MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in Honor of Howard Lasnik*, R. Martin, D. Michaels & J. Urigereka (Eds), 89–115. Cambridge MA: The MIT Press.

Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A life in language*, M. Kenstowicz (Ed.), 1–52. Cambridge MA: The MIT Press.

Chomsky, N. 2004. Beyond explanatory adequacy. In *Structures and beyond*: *The Cartography of Syntactic Structures* 3, A. Beletti (Ed.), 104–31. Oxford: OUP.

Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry* 36(1): 1–22.

Epstein, S. & Seely, T.D. 2002. Rule applications as cycles in a level-free syntax. In *Derivation and explanation in the minimalist program*, S. Epstein & T.D. Seely (Eds), 65–89. Oxford: Blackwell.

Fox, D. & Pesetsky, D. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31: 1–45.

Frampton, J. & Gutmann, S. 2002. Crash-proof syntax. In *Derivation and explanation in the minimalist program*, S. Epstein & T.D. Seely (Eds), 90–105. Oxford: Blackwell.

Frantz, D. 1978. Copying from complements in Blackfoot. In *Linguistic studies of native Canada,* E.D. Cook & J. Kaye (Eds), 89–109. Vancouver: University of British Columbia Press.

Grohmann, K. 2003. *Prolific domains: On the anti-locality movement dependencies.* Amsterdam: John Benjamins.

Hinzen, W. 2006. *Mind design and minimal syntax.* Oxford: OUP.

Inènlikèj, P. & Nedjalkov, V. 1973. Glagoly čuvstva v čukotskom jazyke. *Lingvističeskie issledonanija* 1972: 175–203.

Lee, J. 2003. Phase sensitivity in wh-dependencies. *Korean Journal of Linguistics* 28: 67–89.

Legate, J. 2002. Some interface properties of the phase. *Linguistic Inquiry* 34: 506–516.

Mel'čuk, I. 1988. *Dependency syntax: Theory and practice.* Albany NY: SUNY Press.

Putnam, M. 2007. *Scrambling and the survive principle.* Amsterdam: John Benjamins.

Putnam, M. & Stroik, T. To appear. Structural relations in Survive-minimalism. In *Exploring crash-proof grammars*, M. Putnam (Ed.). Amsterdam: John Benjamins.

Richards, N. 2002. Very local A-bar movement in a root-first derivation. In *derivation and explanation in the Minimalist Program*, S. Epstein & T.D. Seely (Eds), 227–248. Oxford: Blackwell.

Ross, J. 1967. Constraints on variables in syntax. Ph.D. dissertation, MIT.

Stjepanović, A. & Takahashi, S. 2001. Eliminating the phase impenetrability condition. Paper presented at the *Motivating Movement* conference, University of Ulster, Jordanstown (January 26–28).

Stroik, T. 1999. The survive principle. *Linguistic Analysis* 29: 278–303.

Stroik, T. 2000. *Syntactic controversies.* Munich: Lincom.

Stroik, T. 2009. *Locality in minimalist syntax.* Cambridge MA: The MIT Press.

Stroik, T. & Putnam, M. 2005. The lexicon at the interfaces. Paper presented at LASSO (October 7–9).

Te Velde, J. 2005. *Deriving coordinate symmetries: A phase-based approach integrating Select, Merge, Copy and Match.* Amsterdam: John Benjamins.

Ura, H. 1994. Varieties of raising and the feature-based bare phrase structure theory. MIT Occasional Papers in Linguistics 7.

The numeration in Survive-minimalism*

Thomas Stroik
University of Missouri-Kansas City

This paper argues that the computational system of Human Language requires a Numeration; however, the Numeration must be compiled throughout a syntactic derivation and not selected prior to the derivation. The consequence of having a derivationally built Numeration is that the Numeration becomes the domain for both the Merge and Remerge operations, thereby obviating the need for the Internal Merge operation. Furthermore, having a derivationally built Numeration makes it possible to design a crash-proof syntax along the lines developed in Stroik's (2009) Survive-minimalism.

1. Introduction

Over the last fifteen years, Chomsky (1995, 2001, 2004, and 2005) has developed a syntactic framework called the Minimalist Program (MP). This framework assumes that a language L derives structural representations for sentences that are interpreted at the conceptual-intentional C-I and the sensorimotor SM interfaces. Representations in the MP are mapped by the computational system CS of Human Language HL from a NUMeration (of lexical items LIs) to the interfaces, as in (1); and the representations are derived, cumulatively, from the application of two syntactic operations—External Merge (EM), which merges LIs in NUM into a syntactic derivation SD, and Internal Merge (IM), which copies and remerges elements in SD to other positions in SD, as in (2).

(1) CS_{HL}: NUM → ⟨C-I, SM⟩

(2) a. EM: NUM → SD
    b. IM: SD → SD

*Many thanks to Michael Putnam for inviting me to write this paper and for consulting with me on it. Although I've been working on the Survive Principle and on crash-proof grammar for nearly nine years, which culminated in an early draft of Stroik (2009) written in (2004), it's only been since I started working with Michael four years ago that Survive-minimalism has moved forward as a program of inquiry.

The view of syntax briefly described above is widely held—see Collins (1997), Epstein et al. (1998), Lasnik (1999), Hornstein, Nunes, and Grohmann (2005), Epstein and Seely (2006), Hinzen (2006), Boeckx (2008a), and many others.

In Chomsky's (1995, 2004) version of Minimalism, every syntactic derivation begins with the formation of a NUMeration—an array of lexical items LIs selected in pre-derivational blindness from the lexicon.[1] According to Chomsky (1995, 2004), Epstein, Thráinsson, and Zwart (1996), and Hornstein, Nunes, and Grohmann (2005), each LI in NUM consists of a set of lexical features $\langle f_1,\ldots,f_n\rangle$. They further assume that these features fall into three subsets: a semantic-feature set SEM, a phonological-feature set PHON, and a syntactic-feature set SYN. The features in SEM and PHON are intrinsically compatible with the C-I and SM interfaces, respectively. The features in SYN—which include the Case of nouns—are said to be "formal" features that needn't be compatible with either interface. To ensure appropriate interpretation at the interfaces, these "uninterpretable" formal features must be checked/valued and deleted before a derivation reaches the interfaces.

Once the LIs in NUM are randomly (and in pre-derivational blindness) selected from the lexicon, the derivation proceeds by Merging LIs into SD. Importantly, Merge is a syntactic operation that involves two syntactic elements X and Y that have matching features, which become "checked" for interface legibility/compatibility in the course of Merger.[2] Through iterative applications of External Merge and Internal Merge, all the LIs in NUM must be incorporated into SD and all the features of each LI must be checked, or the derivation will "crash" and not be interpreted;[3] if all the features are appropriately checked, the derivation will be submitted to the interfaces for interpretation.

There are three points about this approach to derivations that I would like to emphasize at this juncture. First, although the lexicon participates, through the NUM, in syntactic derivations, it appears to be extraneous to the interfaces.

1. As Chomsky (2001) and Hornstein, Nunes, and Grohmann (2005) observe, the NUM could consist of phase-selected subarrays of LIs, rather than a single array of LIs. Whether NUM consists of a single array or a set of subarrays is not particularly relevant to my analysis of NUM because the problems with the blind selection of LIs, which I discuss at length in this paper, are the same in either case.

2. See Adger (2003) and Putnam and Stroik (2008) for extended discussions of the matching condition on the Merge operation. However, see Boeckx (2008b), Epstein (2007), and Chomsky (2008) for arguments that Merge should be free-Merge, an operation not constrained by matching conditions.

3. See Epstein and Seely (2006) for a discussion of different types of "crashes"—fatal versus non-fatal crashes—and of the relationship between crash-type and interpretability.

This seems to be the case because the lexical features of LIs must be checked for interface compatibility; hence these LIs must potentially possess features that are exotic to (outside of and incompatible with) the interfaces. Second, in standard Minimalism, it is assumed that successful derivations can be built from the blindly selected LIs in NUM, as can be seen in both Chomsky (1995, 2005) and Hornstein, Nunes, and Grohmann (2005). And third, by designing derivations around the checking requirements on the features of individual lexical items, standard Minimalism presents derivations, and their representational outputs, as the structured licensing of LIs. What this suggests is that the interfaces give visibility to individual LIs and that the interfaces perhaps even sensitize their interpretations of syntactic representations to the individuals LIs.

Stroik (2009) offers a view of Minimalism—first called Survive-minimalism by Putnam (2007)—that departs in radical ways from standard Minimalism. In Survive-minimalism, the computational system of Human Language does not function to structurally license LIs, and their features, for interface interpretation because these features are inherently interface features and, as such, don't need to be licensed for interpretation; rather, the computational system uses the features of LIs, which are already interface compatible, to derive interface-ready syntactic objects, which I will call concatenated items (CIs). From this perspective, the computational system builds concatenated structures (by matching/merging features) and the interfaces interpret these concatenations. At issue for Survive-minimalism, then, is not the interface compatibility of LIs, but the interface compatibility of CIs. In accordance with Hinzen's (2006: 250) suggestion that the computational system constructs "complex concepts" from "simple ones," Survive-minimalism assumes that the computational system functions to build and interpret complex concepts and complex forms (i.e., CIs), rather than to merely license strings of simple concepts (i.e., LIs).

By conceiving of the narrow syntax in terms of the construction of CIs, Survive-minimalism requires a re-analysis of the standard approach to lexical features, especially SYN features. In Survive-minimalism, SYN features are not formal features that could be deleted or obviated in some way; rather, SYN features are concatenative (matching) features that are instrumental in the formation and interpretation of CIs—consequently, these features must be visible, and appropriately concatenated, at the interfaces; this is another way of saying that all concatenations must be visible at the interfaces. In fact, all SYN features must be intrinsically interface-interpretable, as are all SEM features and PHON features. (It is important to note here that I am not limiting feature "interpretability" to the C-I interface; rather, for me what it means for a feature to be "interpretable" is that the feature can be read at one of the interfaces.) Since the LIs in the lexicon and the NUMeration contain only interface features, the lexicon and the NUMeration are necessarily contained within the interfaces, and so are all the CIs built out of these LIs.[4] The fact that LIs and CIs are composed of exactly the same interface materials (features) suggests that LIs and CIs are varieties of the same thing: they are Interface Items (IIs). The computational system of Human Language, subsequently, doesn't map interface-extraneous LIs in NUM onto the interfaces; instead, according to Survive-minimalism, the computational system maps interface material onto interface structures at every step in a derivation. In particular, it maps IIs onto II*s, where the structure of any II* properly contains the structure of an II. We can represent this as in (3).

(3) CS_{HL}: II → II*, where II* ⊃ II

Importantly, the computational design given in (3) places significant constraints on the types of syntactic operations permissible in CS. Acting within the domain of the interfaces, syntactic operations take well-formed interface input (starting with LIs) and derive well-formed interface II* output. In other words, these operations, within the Survive-minimalist framework, cannot produce outputs that are not interface-compatible, or interface-interpretable. The design of Survive-minimalism, then, meets the design-threshold hypothesized by Frampton and Gutmann (2002: 93), who posit that the optimal computation system must be crash-proof in that "every derivation converges." Or, as Hinzen (2006: 249) states, "In an optimal system, an impossible meaning would never be produced, and syntax would be 'crash-proof' . . . that

4. An anonymous reviewer has expressed some puzzlement about the architecture I am assuming. The issue of architecture/design is at the heart of my paper. I am arguing throughout this paper that the Numeration/Lexicon and the computational system are situated within the interfaces; I am arguing, in particular, that lexical information is interface material and that the operations of the computational system use this material to construct more elaborate interface material. Meeting interface conditions, given this design, is guaranteed because it is encoded in the very design of Human Language. On the other hand, what is truly puzzling is the design of standard Minimalism. Where the Numeration and Lexicon reside in standard Minimalism and where the computational processes take place aren't ever addressed. It seems to be the case that these design elements in standard Minimalism are all external to the interfaces, and furthermore, it seems that the output of these elements (representations) is also interface external because this output must undergo *Transfer* to reach the interfaces. If this is the case, how can these elements, which are interface external, ever anticipate what interface conditions must be met for a successful derivation? How could any of these interface external elements, such as Internal Merge, ever be conceptually necessary for the performance systems (the interfaces) to operate? And why would extraneously and externally generated representational output ever be usable by the interfaces?

is, the products of locally correct steps of a derivation are always well-formed, and meet interface conditions."[5]

To ensure that the computational system is rigorously crash-proof, producing only well formed II* outputs, Survive-minimalism, as developed in Stroik (2009), Putnam (2007), and Putnam and Stroik (2008), rules out any look-back operations, such as Internal Merge (IM), that depend upon special mechanisms (e.g., Chomsky's *Phase Impenetrability Condition* PIC or *Attract Closest*) to prevent these operations from looking too far back into a syntactic derivation and from producing ill-formed IIs. Since IM-type look-back operations have the intrinsic capability of generating structures that cannot be interpreted at the interfaces (and, consequently, these operations must be constrained), they do not in and of themselves possess the restrictive, domain-encapsulated properties necessary to "meet interface conditions": that is, they are not interface-limited operations (most likely because they assume properties of "movement" that are not interface properties in that movements themselves are not visible, or interpretable, at the interfaces).[6] What is an intrinsic interface property, however, is concatenation; so, the types of syntactic operations allowable in Survive-minimalism are the pure concatenation-type,

5. Not every variety of Minimalism embraces the Crash-proof Principle. Some approaches to Minimalism, such as Chomsky (2008), Epstein and Seely (2006), Epstein (2007), and Boeckx (2008b) are crash-tolerant approaches. Of these, Epstein and Seely (2006) and Epstein (2007) are the most tolerant. In fact, they argue that crashes show up throughout a syntactic derivation—which suggests that syntactic derivations are crash-driven. As I argue in this paper, crash-tolerant versions of Minimalism are problematic because they can't guarantee that an interpretable syntactic derivation will ever reach the interfaces. This means that the computational system could endlessly be producing unusable output.

6. Chomsky (2004) claims that External Merge EM and Internal Merge IM are versions of the same operation, which suggests that these operations should have the same conceptual justifications. However, the fact EM and IM have similar effects on building syntactic structure is not sufficient reason to construe these operations as versions of the same thing. These operations have two inherent properties that make them incommensurable: (i) they function over very different domains—EM functions over derivational-external lexical domains and IM functions over derivational-internal structural domains—and (ii) they build structure in very different ways: EM adds material to syntactic structures, while IM doesn't. From the second property it follows that it is not in principle possible to build syntactic structure without EM; the same is not true of IM. Hence, EM is unquestionably conceptually necessary; on the other hand, the conceptual necessity of IM is quite in question (see Putnam and Stroik (this volume) for a discussion of the conceptual necessity of the IM operation). Furthermore, should it be the case that EM and IM are actually versions of the same operation, then one must wonder why we must have both of them—which would seem to argue for finding the generalization that allows the operations to collapse in a single Merge M operation (given the conceptual necessity of EM, this M operation would have to be much more EM-like than IM-like).

the External Merge type. Hence, Survive-minimalism has only syntactic operations (SynOp) that concatenate elements from NUM to SD, as in (4).

(4) SynOp: NUM → SD

As we can see from (4), in Survive-minimalism, the NUMeration plays a key role in syntactic derivations, as it does for Chomsky (1995), but it has very different properties than does NUM in standard Minimalism. The NUMeration in Survive-minimalism is not built blindly from the lexicon; its LIs are not selected all at once prior to the derivation; its LIs do not have uninterpretable features; and elements in NUM do not move into SD, thereby vacating NUM. Rather, NUM is built piecemeal throughout a syntactic derivation, adding copies of LIs in the lexicon as needed; these copies can be copied multiple times in a syntactic derivation to satisfy the concatenative demands of interface features.

Although there are substantial design differences between Survive-minimalism and standard Minimalism, I will focus, in this paper, on only one key difference: the role that the lexicon and the NUMeration play in a syntactic derivation. I will argue here that, despite Chomsky's (1995, 2001, 2004) assumptions about NUM formation, the lexical items in a NUMeration cannot be selected (blindly) prior to the onset of a syntactic derivation and the lexicon is not extraneous to the interfaces.

2. Selecting a NUMeration

According to Chomsky (1995, 2004), Hornstein, Nunes, and Grohmann (2005), and even Hinzen (2006), the NUM must be exhaustively determined before a syntactic derivation (SD) can begin. That is, LIs are brought into NUM, via a Select operation, from the lexicon prior to the computations of the SD. Furthermore, since these LIs are selected for NUM without look-ahead properties able to anticipate how a derivation might proceed, they must be selected in derivational blindness. Consider, however, the effect that having a NUM with ten blindly (and randomly) selected LIs would have on syntactic computations, where NUM = $\{LI_1,\ldots,LI_{10}\}$. If we assume, as does Chomsky, that every LI_i in NUM must be eventually inserted into the SD, then from our NUM we could, in principle, generate 10! (3.6 million) structured strings of words from the LIs we select.[7] Of these strings/concatenations,

7. As far as I have been able to find, no one, except perhaps Collins (1997), has looked closely at the problem of selecting LIs for NUM in derivational blindness. Those theorists who use NUM in their derivations generally posit NUMs that will work successfully. The statistical improbability of forming such NUMs has gone largely unnoticed.

we must identify, within Chomsky's version of Minimalism, the optimally derived structured string (the one out of 3.6 million). This daunting search could be mitigated, somewhat, if there were some sort of feature-match requirement placed on Merge operations which could dramatically reduce the number of concatenations that can be licitly derived.[8] However, whether we have a feature-matching requirement on Merge or not, to be sure we find the optimal SD for the LIs in NUM, we must compute, at the very minimum, at least 90 SDs. We have to consider the possibility of starting the SD with each of the ten LI_is and we must consider the possibility of Merging each LI with the remaining nine LIs. All of the ninety SDs that are well-formed after the first Merger will have to proceed further, creating additional SDs (if Merge lacks matching requirements, then all first Merges will have to proceed derivationally). Though it may not be the case that all 3.6 million possible SDs will have to be computed to completion, there will, in all likelihood, be hundreds of SDs generated for the LIs in NUM—even with feature-match requirements on Merge. If all SDs, except for the optimal one, will crash and not be interpreted, then every NUM with ten LIs will minimally produce hundreds of crashes, perhaps more. Needless to say, computing all these SDs to find one of them that could be interface compatible would be an amazingly time-consuming ordeal. An even more daunting consequence of following Chomsky's (1995) approach to NUM selection is that there is little possibility of deriving any successful, interpretable SD for NUM. Given that LI selection is random and that it must take place under conditions of derivational blindness, the possibility of finding a single well-formed SD among the millions we could, in principle, generate from the ten LIs in NUM is infinitesimally small (to see the problem, blindly select ten words from the dictionary; then try to find even one way of putting these words together as a sentence). What this suggests is that we will spend an enormous amount of processing time and effort to produce only ***crashes***. The problems described above become all the more poignant when we consider the derivational effects of having a blindly selected NUM with twenty LIs. From such a NUM, we could generate 20! (over 1.0×10^{18} = 1,000,000,000,000,000,000) structured strings of words, with little hope that any of them could be well-formed SDs (again, randomly select twenty words from the dictionary and look for ways to string them together as sentences). As before, we might not have to compute all 20! strings, but we will have to compute at least the first Merger (20x19 = 380) and we will have to continue computing any initially successful derivations. Hence, it is likely that we must compute thousands

8. Without a feature-match condition on Merge, all the possible 3.6 million structured strings will have to be generated and submitted to the interfaces for interpretation. However, as I mention above, it is unlikely that any of these strings will be interpretable.

of derivations that will crash. And all these crashes will occur whether we eventually stumble upon an interface interpretable SD, or not.

Of note, Chomsky (2004) reformulates NUM, making it phase-based. This phase-based approach to NUM, however, fares no better than does the unrestricted NUM formation approach discussed above. In Chomsky (2004), the NUM is not a single set of the LIs that will appear within a syntactic derivation. Rather, it is a set of subsets, where each subset contains a phase-head (C or *v**). Although each subset will have one element (the phase-head) that is possibly derivationally necessary and, consequently, each subset will contain an element that is arguably not selected in derivational blindness, the remainder of LIs in these subsets are selected pre-derivationally. Given that these subsets (especially those with a *v** phase-head) could have ten or more LIs, the problems we discuss above don't disappear. As a consequence, Chomsky's (2004) proposal to construct NUM as a set of phase-selected subarrays will not substantially improve the likelihood of deriving interface interpretable SDs.

Hornstein, Nunes, and Grohmann (2005) recognize that NUMerations can be derivationally uncooperative, that there can be what they call "crazy" NUMerations, such as those in (5), which will not have any convergent, interface-interpretable derivations built from them.

(5) a. NUM = {tree, of, buy}
  b. NUM = {with, about, Mary, John}
  c. NUM = {see, man, Infl}

Such "crazy" NUMerations, however, don't appear to bother Hornstein, Nunes, and Grohmann very much. They say, "Since PF and LF already are responsible for filtering crashing derivations, there's no need to filter out the numerations in (5), since derivations resulting from them will crash at LF and PF" (2005: 71). This rather cavalier attitude towards "crazy" NUMerations seems to miss two very extraordinary facts about blindly selected NUMerations. First, under the assumption that lexical items are selected for NUM blindly and without any look-ahead properties, selecting "crazy" NUMerations will be the rule, not the exception. That is, it will generally be the case, even for NUMs with fewer than ten LIs in them, that for any given NUM, no convergent derivation will be able to be built from them. In other words, the derivations sent to the interfaces will almost always crash because they will come from "crazy" NUMerations. Second, if, in some extremely rare circumstance, a NUMeration is not "crazy," it will still produce a substantial number of crashes. In fact, as we noted previously, should we have a NUM with ten LIs, we will have to compute hundreds of SDs for the LIs in NUM—perhaps as many as 3.6 million SDs—in our quest to find the (single) optimally computed SD. Approaching syntactic derivations from this perspective leaves the interfaces with the problem of always having to rule out a sizable number of crashed derivations. And should it be the case that we attempt to circumvent this problem by proffering a new analysis that "aborts" derivations rather than have them "crash" (i.e., derivations fail/fizzle somehow before they get to the interfaces), we will still have a troubling computational state of affairs in which we would allow all these abortings to be syntactically derived and then rejected at some point in the derivation. From this analysis we can see that constructing syntactic derivations out of blindly selected NUMerations cannot help but be excessively costly as a computational system. That is, standard Minimalism, as it currently conceives of the NUMeration, necessarily produces a grossly crash-rife syntax, not a crash-proof one.

Given the problems with blindly selecting LIs for NUM, perhaps we should surrender our commitment to NUM and assume, instead, that the computational system maps LIs directly from the LEXicon to the C-I and SM interfaces, thereby obviating the need for NUM.[9] There are, however, several reasons for having NUM: (i) Chomsky (1995: 236–238) argues that LIs in LEX lack formal features and that these features are accreted to LIs when they are mapped from LEX to NUM; (ii) as Hornstein, Nunes, and Grohmann (2005: 69–70) maintain, without NUM, the computational system won't be able to determine, in principle, when a derivation terminates—that is, NUM sets computational boundaries for SDs (they note that "…the computational system that builds syntactic structures doesn't work with the whole lexicon at a time, but with collections of lexical items" (2005: 71); (iii) syntactic derivations involve computations in short term memory and NUM, rather than LEX, is the short term memory storage for LIs used in a derivation; (iv) NUM restricts the LI domain of the External Merge operation, without NUM, this operation would have to sort through the lexicon, which is a much more vast domain and sorting through it would be costly in processing terms; and (v) if we surrender all Move-type syntactic operations, as Stroik (1999, 2009) suggests, then, without NUM, it would be impossible to have a derivational account of displacement effects (which arise when a single syntactic constituent can appear in more

9. Chomsky (2008) proposes a computational system that does not include a Numeration. Since he doesn't offer any arguments for why EM should select LIs from the lexicon and not from NUM beyond the broad claim that EM is unbounded, it is difficult to respond to his proposal. The viability of Chomsky's proposal will depend, in part, on whether or nor it can respond to the pro-NUM arguments (i)–(v) I give above. Argument (v) is a particularly important one because it raises questions about the conceptual necessity of Internal Merge and because it points out that NUM must be a component of the computational system if this system lacks IM operations (see Putnam & Stroik (this volume) for arguments against the conceptual necessity of IM). Given that the arguments in (i)–(v) haven't been addressed, I will continue to assume that NUM exists.

than one structural position)—this occurs under Stroik's analysis because, absent Move-type operations such as IM, NUM provides the only domain from which any single LI could be multiply Merged into a syntactic derivation (see (6)). Trying to multiply Merge an LI X from the lexicon will generate structures in which the LI X appears multiple times, but it won't generate displacement effects since each X will be a new X from the lexicon; hence, no single X will appear in more than one structural position.

(6) a. Merge {see, who} → see **who**
b. Merge {Pat, {see, who} → Pat see **who**
c. Merge {will, {Pat, {see, who}}} → will Pat see **who**
d. Remerge {who, {will, {Pat. {see, who}}}} → who will Pat see **who**

The emboldened wh-word in (6) indicates that this word has unchecked features and that, as a result, the LI continues to "survive" in NUM and can be Remerged into the SD to have its remaining features appropriately checked. In (6), the wh-displacement occurs not because of a Move-type operation but because the wh-element in NUM can be merged into the SD more than once. Without NUM, this sort of multiple-merging of a single LI into an SD is not possible (see Stroik (2009) and Putnam (2007) for extended discussions of the syntactic necessity of remerging LIs from the NUM to SD).

Let us assume, in the face of the foregoing arguments, that minimalist syntax must have a NUMeration. However, as we have seen, if we are to have a syntax that is not prone to producing an egregious amount of unusable and uninterpretable derivational output ("crashes"), the lexical items in this NUMeration cannot be blindly and exhaustively selected prior to the onset of the syntactic derivation. But this raises the question of whether or not any LIs in NUM can ever be selected independently of (i.e., blindly of and/or prior to) the derivation? Needless to say, once we allow even one derivationally independent LI to appear in NUM, there will be no principled way of ensuring that other such LIs are not selected for NUM, too. The consequence, then, of permitting any derivationally independent LIs in NUM is that we could have NUMerations with all blindly selected LIs—which is exactly the sort of situation we must exclude. Therefore, it must be the case that no LIs in NUM can be selected derivationally blind. This suggests that the NUMeration is derivationally sighted, or derivationally built. In other words, all LIs enter NUM one at a time, as they are derivationally required.

My analysis here supports Putnam and Stroik's (2008) crash-proof reformulation of phrase structure relations. According to Putnam and Stroik, the ways in which derivations begin, proceed, and terminate are dictated by the hierarchical Feature Matrices (FMs) of lexical items in NUM. Any given LI can have an FM with a CATegory feature, with SUBCATegory features, and with Interface (IF) features, such as wh-features, Case features, etc. All these features are SYN features that participate in the concatenating operations, Merge and Remerge, and that must be checked in the course of a syntactic derivation. Of note, though, is the fact that these features in an FM are structured hierarchically, FM = ⟨SUBCAT ⟨IF-selecting ⟨CAT ⟨IF-selected⟩⟩⟩⟩ and they must be checked in order. The selecting features of an LI—its SUBCAT features and any IF-selecting features, such as the Case feature of the light verb v or the Case feature on the Tense-head or the wh-feature of a C-head—must be checked before the selected features are (see Putnam and Stroik (2008) for an extended discussion of FMs). Checking features in this order guarantees that an LI will not be selected for NUM or be merged into an SD unless it is required to match features already projected in the SD. That is, an LI with a SUBCAT feature f will enter a NUM and will be merged into the SD that is being built from this NUM if and only if there is an X being projected in SD that has a matching f-CAT feature. Since an LI Y with a SUBCAT feature is precluded from being selected for NUM unless the SD already has built an X constituent with a matching CAT feature, this mean that Y cannot be the first element in the NUM of the SD. The first element in any NUM, as a result, must not have a SUBCAT feature.[10]

We can see how this works if we return to the syntactic derivation in (6). Both the NUM and the SD for (6) begin with the LI *who*. The fact that the FM for *who* has no SUBCAT feature—the FM for this LI is ⟨CAT-D ⟨WH⟩⟩ (this FM states that *who* is a DP with a wh-feature)—which makes *who* a suitable lexical selection to initiate an SD (see (7a)).[11] After *who* is placed from NUM into the SD, it is now necessary to select an LI from the lexicon that has a matching SUBCAT-D feature. Since the verb *see* takes a DP argument and, therefore, has the appropriate

10. An anonymous reviewer notes that it not possible to have a first element in a NUM because NUM is a set and the elements in sets are unordered. This point applies to sets that are defined all-at-once (the way pre-derivationally selected NUMs are formed in Chomsky (1995) and Hornstein, Nunes, and Grohmann (2005)). I'm arguing that the NUM is built piecemeal; hence, there can be a first element in NUM.

11. An anonymous reviewer raises the possibility that the quantifier feature of *who,* as well as its selectional requirements (restriction and scope), could be SUBCAT features. If so, the LI *who* would not be without SUBCAT features. I'm using the long-standing notion that SUBCAT features of an LI identify the CAT features of its arguments, as the verb *see* subcategorizes for a DP object argument (i.e., an object with a CAT-D feature). For me, the wh-feature and the selectional features of *who* are semantic/interface features (IFs) and not SUBCAT features (note that separating SUBCAT information from selectional information has also been a long-standing tradition within generative analyses).

SUBCAT-D feature, this verb can be added to the NUMeration (= {see, who}) and merged into the SD, as in (7b).

(7) a. NUM = {who} → who ⟨CAT-D ⟨WH⟩⟩
b. NUM = {see, who} → Merge {see, who} → see **who** ⟨SUBCAT-D ⟨CAT-V⟩⟩

At this juncture in the syntactic derivation, two things happen. First, the LI **who** has an unchecked ⟨WH⟩ feature at this point in the derivation and, as a result, this LI is required, by Stroik's (1999, 2009) *Survive Principle,* to "survive" (remain activated) in the NUM.[12] Second, although the first SUBCAT feature in the verb's FM (⟨SUBCAT-D ⟨SUBCAT-D ⟨CAT-V⟩⟩⟩) has been checked in (7b), the remainder of the FM projects and must be checked. In particular, the next feature in the verb's FM (i.e., the ⟨SUBCAT-D⟩ feature) must be checked. To do so, we must add an LI to NUM with a ⟨CAT-D⟩ feature. The LI *Pat* has the requisite feature: its FM is ⟨CAT-D ⟨CASE⟩⟩. Hence, *Pat* will merge into the SD as in (7c).

(7) c. NUM = {Pat, see, **who**} → Merge {Pat, {see, who} → **Pat** see **who** ⟨CAT-V⟩

In (7c), the ⟨CAT-D⟩ feature of *Pat* will be checked, though its ⟨CASE⟩ feature won't; consequently, this LI will also survive in NUM. Furthermore, the ⟨CAT-V⟩ feature of the verb will project. Checking the ⟨CAT-V⟩ feature being projected requires that an LI with a ⟨SUBCAT-V⟩ be added to NUM and merged into the SD. The modal *will* has this feature. However, it also has a ⟨CASE-NOM-selecting⟩ feature and, as one of its options, it has a selected Q(uestion)-feature, as well as its selected CAT-M feature. The modal will merge into the SD in the following way and with the following consequences.

(7) d. NUM = {will, **Pat**, see, **who**} → Merge {will, {Pat, {see, who}}} → **will Pat** see **who** ⟨CASE-NOM ⟨CAT-M ⟨Q⟩⟩⟩

As we can see, merging the modal into the derivation leaves the SD with a projecting ⟨CASE-NOM-selecting⟩ feature. This feature must be matched by some Y with a ⟨CASE-selected⟩ feature. Given that Y would have such a feature available if and only if its hierarchically higher features, including its SUBCAT and CAT features, have been checked previously in the SD, the appropriate Y must already be in NUM and must be available for Remerge into the SD. The LI **Pat** has the requisite available feature, so it will Remerge into the SD, as in (7e).

12. The *Survive Principle* states that if Y is a syntactic object SO merged into a syntactic projection headed by X and Y has features incompatible with (i.e., cannot be potentially checked by) the features of X, then Y remains active in NUM.

(7) e. NUM = {**will**, Pat, see, **who**} → Remerge {Pat, {will, {Pat, {see, who}}}} → Pat **will Pat** see **who** ⟨CAT-M ⟨Q⟩⟩

The projecting ⟨CAT-M⟩ feature of the modal now requires an LI with a ⟨SUBCAT-M⟩ feature to be added to NUM and merged into the SD. One LI with this SUBCAT feature is a C(omplementizer). Though there is a variety of C-heads, the sort we need here, as is indicated in (7e), is one with a Q-type feature, as well as a WH-selecting feature. Should we introduce any other C into NUM, the derivation will stall because the Q feature on the modal and the WH-feature on the wh-word won't be checked. So, let's assume that the C selected for NUM has a ⟨SUBCAT-M ⟨Q ⟨WH⟩⟩⟩ FM. Notice that this FM has no CAT features, which indicates that the NUM will not be able to add LIs and that the SD will terminate after all the projecting features are checked.

(7) f. NUM = {C, **will,** Pat, see, **who**} → Merge {C, {Pat, {will, {Pat, {see, who}}}}} → C Pat **will Pat** see **who** ⟨Q-selecting ⟨WH-selecting⟩⟩

Although the merger of the C element into the SD checks the CAT-M feature of the modal, it leaves the ⟨Q⟩ feature of the modal unchecked; therefore, the modal will survive in the NUMeration and will be available for Remerger. Since the projecting features ⟨Q-selecting ⟨WH-selecting⟩⟩ that remain in the SD do not involve CAT or SUBCAT features, they must be checked by LIs already in NUM, as in (7g) and (7h).

(7) g. NUM = {C, **will**, Pat, see, **who**} → Remerge {will, {C, {Pat, {will, {Pat, {see, who}}}}}} → will C Pat **will Pat** see **who** ⟨WH-selecting⟩

(7) h. NUM = {C, will, Pat, see, **who**} → Remerge {who, {will, {C, {Pat, {will, {Pat, {see, who}}}}}}} → who will C Pat **will Pat** see **who**

The syntactic derivation is now complete. All the SYN features in all the FMs of all the LIs introduced into NUM have been appropriately checked; consequently, the SD in (7h) is now a structural representation that is interpreted by the Conceptual-Intentional and Sensorimotor interfaces (note: since this representation is the output of lexical items and derivational processes that are contained in the interfaces, it too is contained in the interfaces and won't have to undergo any *Transfer* operation to get to the interfaces—as it must in Chomsky (2001, 2004, and 2008)).

For the purposes of this paper, we need to pay particularly close attention to the significant role that the NUMeration plays in the syntactic derivation provided above. Notice that by conceiving of the NUMeration as a short-term work-space for the syntactic derivation, by building the NUMeration under conditions of derivational necessity, and by positing an interactive relationship between NUM and the SD, we can design a computational system for Human Language that is optimally simple and efficient, ontologically minimal (not requiring Movement-type rules or economy/minimality conditions), and maximally crash-proof.

3. Locating the lexicon

One of the main presuppositions of standard Minimalism is that the LEXicon, and therefore the NUMeration, too, is somehow extraneous to the interfaces. We can observe this presupposition in Lasnik's (1999), Hornstein's (2001), and Chomsky's (1995, 2004) claims that a grammar maps lexical material from the NUMeration to the interfaces, as if this mapping is from one domain to another quite different domain, and in their claims that syntactic operations make the features of LIs legible at the interfaces, which suggests that the features aren't inherently interface-legible. Relatedly, Hornstein, Nunes, and Grohmann (2005) note that "…natural languages, for yet unexplained reasons, have formal features that are legible neither at LF nor at PF" (328). Following Chomsky, they argue that the syntax "…allows these illegible features to be appropriately deleted, allowing the interfaces to read the objects built by the computational system" (2005: 328).

Let us follow the standard Minimalist line of analysis and assume that some, perhaps even all, of the SYN features are extraneous to the interface and are deleted in the course of a derivation to ensure that they don't show up at the interfaces. If this is the case, then, as Stroik (2009) and Putnam and Stroik (2008) contend, it will be impossible for any such deleted formal feature f_n to be learned. Once the f_n feature of an LI is deleted in a derivation, it will not appear in the representation sent to the interfaces and it won't be part of the output sent through the sensorimotor mechanisms to other users of the language L. As a result, when children learn the language, they will not have any access to this feature. They also won't be able to enlist the Cognitive-Intentional and Sensorimotor interfaces in identifying this feature because the feature won't have any interface visibility. It would appear, therefore, that locating the feature f_n and attaching it to a given LI would not be a matter of the computational system of Human Language. Instead, this feature will have to be acquired in some ancillary fashion. We might have to posit an interface-independent parsing mechanism that can deduce the presence of the interface-invisible feature f_n and add it to the FM of the appropriate LI. Going this route to explain interface-extraneous features is problematic, however. We would have to have a "crazy" grammar that allowed features to be added by one mechanism (perhaps the parser) and deleted by another mechanism (the computational system) without ever allowing these features to appear at the only levels of representation (the interfaces) that matter for a grammar. Closeting and cloaking features as described above complicates grammar in ways inconsistent with Minimalist commitments to conceptual/interface necessity, so we should strongly disprefer analyses that assume the existence of interface-extraneous features.

We can observe the problems of having interface-extraneous features closely if we take a look at the sorts of portal features that Chomsky (2001, 2004, and 2008) and others posit to manage constituent movements/remergers. Faced with the need to limit the domain and range of constituent movements, Chomsky assumes that various phrase heads H can have edge features which permit a structurally defined set of constituents to move to the left edge of HPs. These movements, driven by edge features, are short/intermediary movements of a constituent X that license other short movements of X and that cumulatively result in the long-distance movement of X; the edge features, in this way, serve as portals to subsequent movements. We can see these short, portalling movements in (8).

(8) Who does Mary [***t*** believe [***t*** that Pat [***t*** will hire ***t***]]]

In (8), the LI *who* begins its derivational journey as the object of the verb *hire* and it moves leftward up the derivational structure in short steps, each of which is marked by the trace ***t***. The movement of the LI *who* occurs because *v* heads and C heads have edge features that attract the wh-element in (8). Importantly, these portal/edge features, which are not interface interpretable features, delete once they merge appropriately with the wh-element. At the interfaces, the wh-element is interpreted in its highest structural position and in its lowest trace position, but nowhere else, because the portal features and their concatenations are not visible at the interfaces. So how can these portal features of H be learned? As we argue above, these features can't be learned from the interfaces, which suggests that, if they are learnable at all, they must be learned/acquired from some source other than the interfaces—say, the parser. However, a parser would discover the features associated with the intermediate trace positions of the wh-element in (8) only if it has to re-trace the derivational history of a representation in reverse. And why would it do so? More so, how would it do so? To re-trace the feature-driven derivational history of any sentence, a parser would have to know the features that motivated this history. Since all the portal features have been deleted in the derivation and don't appear at the interfaces, there is no way for the parser to know what these features might have been.[13] In other words, reversing feature-driven

13. Another problem with deleting portal features such as edge features before they reach the interfaces is that they have no interface-function; hence, they are not conceptually necessary for the interfaces to operate. Under the Minimalist assumption that all design elements of Human Language must be conceptually necessary at the interfaces, these sorts of features should be disallowed.

movements after the features have been eliminated would be as impossible as accurately back-tracking the journey of some animal after its tracks have been deleted. What this discussion seems to indicate, then, is that it is not possible to learn deleteable, interface-extraneous features.

If my analysis is correct, all lexical features—SEM, PHON, and SYN features—must be interface interpretable. From this it follows that (i) no lexical features can be deleted in the course of a derivation because interpretable information cannot be excised; (ii) the LEXicon, together with the NUMeration, is contained within the interfaces because all lexical material is interface material; and (iii) since all lexical features are intrinsically interface interpretable, syntactic operations, contra Chomsky, do not serve to check (or value) lexical features for interface legibility and the well-formedness of syntactic derivations is not dependent on feature legibility—rather, these operations licitly map interface-interpretable (II) units, including LIs, onto larger interface-interpretable (II*) units by ensuring that all mergings/concatenations licitly link grammatical units via feature agreement. Under this view, a theory of syntax is a theory of II* legibility and II* interpretation.

4. Conclusion

Since Chomsky (1995) introduced the NUMeration into the design of the computational system, the NUMeration has played a very special role in building syntactic structure: it has been used to limit the lexical materials out of which a derivation could be built. That is, the NUMeration establishes a strong "materials" constraint on the architecture of a derivation; this constraint is so strong that it disallows any derivational structure that does not use all and only the lexical materials pre-selected for the NUMeration. In this paper, I revisit this conception of the NUMeration, arguing that the NUMeration should not be treated as a "materials" constraint because doing so will leave the computational system virtually incapable of ever producing any interpretable representations. Needless to say, having a computational system that produces very little usable output from an unbounded set of NUMeration inputs does not reach the "optimal design" threshold that Chomsky posits for a computational system. I propose, instead, that the NUMeration be a work-space area for a syntactic derivation—a space where interface-interpretable materials are added as required by the derivation, where these materials are kept activated if they have features that have been unchecked/unused, and from where these materials are Merged or Remerged into a derivation. As a subset of the LEXicon, a NUMeration houses only materials that are intrinsically interface-interpretable and it serves as the domain over which syntactic operations (Merge and Remerge) function to map larger interface-interpretable units. Conceived of in this way, the NUMeration establishes the sort of work-space required to build a crash-proof syntax. Stroik (2009) and Putnam (2007) design Survive-minimalism around this notion of the NUMeration.

References

Adger, D. 2003. *Core syntax*. Oxford: OUP.

Boeckx, C. 2008a. *Understanding minimalist syntax*. Oxford: Blackwell.

Boeckx, C. 2008b. A tale of two minimalisms. Ms, Havard University.

Chomsky, N. 1995. *The minimalist program*. Cambridge MA: The MIT Press.

Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A life in linguistics*, M. Kenstowicz (Ed.), 1–52. Cambridge MA: The MIT Press.

Chomsky, N. 2004. Beyond explanatory adequacy. In *Structures and beyond,* A Belletti (Ed.), 104–131. Oxford: OUP.

Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry* 36: 1–22.

Chomsky, N. 2008. On phases. In *Foundational issues in linguistic theory*, R. Freidin et al. (Eds), 133–166. Cambridge MA: The MIT Press.

Collins, C. 1997. *Local economy.* Cambridge MA: The MIT Press.

Epstein, S. 2007. On i(nternalist)-functional explanations in minimalism. *Linguistic Analysis* 33: 20–53.

Epstein, S., Thráinsson, H. & Zwart, C.J.-W. 1996. Introduction. In *Minimal ideas,* W. Abraham et al. (Eds), 1–66. Amsterdam: John Benjamins.

Epstein, S. et al. 1998. *The derivation of syntactic relations*. Oxford: OUP.

Epstein, S. & Seely, D. 2006. *Derivations in minimalism*. Cambridge: CUP.

Frampton, J. & Gutmann, S. 2002. Crash-proof syntax. In *Derivation and explanation in the minimalist program*, S. Epstein & D. Seely (Eds), 90–1–5. Oxford: Blackwell.

Hinzen, W. 2006. *Minimal mind design*. Oxford: OUP.

Hornstein, N. 2001. *Move! A minimalist theory of construal*. Oxford: Blackwell.

Hornstein, N., Nunes, J. & Grohmann, K. 2005. *Understanding minimalism*. Cambridge: CUP.

Lasnik, H. 1999. *Minimalist analysis*. Oxford: Blackwell.

Putnam, M. 2007. *Scrambling and the survive principle*. Amsterdam: John Benjamins.

Putnam, M. & Stroik, T. 2008. Syntactic relations in survive-minimalism. Ms, Carson-Newman & University of Missouri-Kansas City.

Stroik, T. 1999. The survive principle. *Linguistic Analysis* 29: 278–303.

Stroik, T. 2009. *Locality in minimalist syntax*. Cambridge MA: The MIT Press.

PART II

Studies of movement phenomena and structure building in Survive-minimalism

Long-distance agreement without *Probe-Goal* relations

Omer Preminger
Massachusetts Institute of Technology

In this paper, I explore the possibility accounting for constructions that appear to instantiate *Long-Distance Agreement* without appealing to a formal operation of *agreement-at-a-distance*. The viability of such an account is particularly important in light of recent theoretical developments that suggest a move away from *Probe-Goal* oriented approaches to movement and towards viewing movement as a response to formal needs of the moving element itself.

Broadly speaking, I consider two possible approaches: (i) agreement is established in a purely local configuration, followed by the agreeing head (and whatever material ends up intervening between this head and the target noun-phrase) moving away, giving the impression of Long-Distance Agreement; and (ii) apparent Long-Distance Agreement is actually an instance of syntactic movement in which the phonological component chooses to pronounce the moved element in its lower position. It is shown that the latter approach fares better with respect to the scopal properties of several constructions, including English expletive-associate constructions, and so-called Long-Distance Agreement in Hindi-Urdu and in Basque.

1. Introduction

Recent work in generative syntax has seen a resurgence in the debate regarding the driving force behind "syntactic movement," broadly construed to include any instance of a mismatch between the locus in which an element is pronounced, and the locus in which an element is (wholly or partially) interpreted. To the uninitiated observer, it may appear rather obvious that syntactic movement is driven by the nature of the moving element—for example, that wh-elements in English questions move because of some factor having to do with their wh-morphology; after all, that seems to be the property that sets them apart from their non-moving counterparts. Let us refer to this avenue of explanation as *push-movement*.

The wake of Rizzi's (1990) "Relativized Minimality" (see also Frampton 1991), and in particular, work done within the Minimalist Program (Chomsky 1995), saw a shift from *push-movement* to *drag-movement* as the prevalent approach to movement in generative syntax—where *drag-movement* refers to the notion that movement is driven by the landing-site, or some element (an *attractor*) that is structurally-adjacent to the landing-site (I use the terms "push" and "drag" following van Riemsdijk 1997).

Nonetheless, there exists a considerable (and growing) body of work suggesting that models that appeal exclusively to *drag-movement*, such as the *Probe-Goal* framework (e.g., Chomsky 2001 *et seq.*), are ultimately untenable; a non-exhaustive list includes: Ackema et al. (1993), van Craenenbroeck (2006), Koeneman (2000), Moro (2007), Nash and Rouveret (1997), Nilsen (1997), Platzack (1996), van Riemsdijk (1997), and Suranyi (2004).

Abandoning *drag-movement* altogether faces several challenges, centered around providing alternative accounts for the phenomena that motivated the shift to *drag-movement* in the first place.[1] In this paper, I address one such phenomenon: instances of so-called *Long-Distance Agreement*.

2. Why long-distance agreement is a challenge

Within *push-movement* approaches, syntactic constituents undergo displacement due to formal requirements of those constituents themselves. This can be implemented by merging an already-present constituent in the derivation at an additional attachment site (*derivation-to-derivation re-merge*—as articulated by Chomsky 2004, albeit within a *drag-movement* model), or by merging an element from the numeration into the derivation more than once (*numeration-to-derivation re-merge*—as articulated by Putnam 2007, Putnam and Stroik to appear, Stroik 1999). Regardless of implementation, however, *push-movement* regards syntactic displacement as a response to the formal requirements of the displaced element. Within such a framework, it would be somewhat puzzling if language were to provide a second, unrelated mechanism for addressing the formal needs of a syntactic constituent—namely, some form of *agreement-at-a-distance*.[2]

Consider the following well-studied and well-documented alternation:

(1) a. A man seem*(s) to me to be sitting outside.
    b. There seem*(s) to me to be a man sitting outside.

1. Perhaps the most prominent such phenomenon is that of superiority effects in multiple-wh questions in languages like English. Prima facie, it seems a *push-movement* model must resort to global computation to determine which of the wh-elements in English multiple-wh questions would undergo overt movement. See Preminger (2007) for a possible alternative, which involves neither *drag-movement* nor global computation.

2. Here and throughout, any reference to "distance" is meant as a primarily structural (rather than linear) measure.

If movement of the noun-phrase *a man* in (1a) occurs to address some formal need of that noun-phrase (plausibly, the checking of its Case-feature), the question immediately arises as to what the fate of this same formal need is in (1b); and if this need—whatever it is—can be addressed at a distance, as in (1b), why is movement obligatory in (1a)? The familiar response is appealing to an EPP-like requirement: stating that the needs of the noun-phrase can indeed be addressed at a distance—via some *Agree*-like mechanism (Chomsky 2000, 2001)—but that there is an independent requirement that the subject position be filled, and this requirement can be satisfied either by movement of the noun-phrase or by insertion of an expletive.

It is important to note that this EPP-based explanation is crucially of a *drag-movement* nature: if the EPP is a requirement of subject positions, and movement in (1a) occurs to satisfy this requirement, this is a bonafide instance of movement being driven by the landing site (or some element that is structurally-adjacent to the landing-site, such as T^0). Thus, if one wishes to abandon *drag-movement*, such an account cannot be maintained.

3. Two alternatives

In this section, I will examine two possible alternatives to the standard *Agree*/EPP account discussed in §2, neither of which invokes *drag-movement*.

3.1 Late evacuation

Broadly speaking, the *late evacuation* account goes as follows: what looks like agreement-at-a-distance is established in a perfectly local configuration; it is the agreeing head (in addition to whatever adjacent material might occur in its vicinity) that later moves away from the site where agreement was established, creating the impression of Long-Distance Agreement.[3] Of course, adopting this approach requires one to assume a far more articulated cartography of clause structure: the agreeing head, as well as any material that ends up in between the agreeing head and the agreed-with noun-phrase, must be moved upwards after agreement has

3. The *late evacuation* account is inspired by Sigurdsson's (2000, 2004, 2006) and Sigurdsson and Holmberg's (2008) treatment of Long-Distance Agreement and dative intervention in Icelandic dialects. However, I would be quick to note that the connection is anecdotal, and theoretically immaterial: the forthcoming discussion of *late evacuation* has no bearing on Sigurdsson and Holmberg's analysis, nor on its adequacy in dealing with the Icelandic facts for which it was intended. Moreover, any kind of "evacuation" involved in Sigurdsson and Holmberg's account is done via head-movement; in the construction discussed in the text (the English expletive-associate construction), head-movement cannot be used to achieve "evacuation", due to the intervening phrasal material—namely, the dative experiencer argument.

been established locally; and these movement operations all require appropriate landing-sites to be available. As a result, the clause must be populated with projections that provide these landing-sites.

More concretely, recall (1b), repeated below:

(1) b. There seem*(s) to me to be a man sitting outside.

The standard analysis of (1b) would be as in (2), below:

(2) [$_{TP}$ There T^0 [$_{vP}$ seems [$_{VP}$ t$_{V^0}$ [$_{PP}$ to me]
[$_{TP}$ to [$_{vP}$ be [$_{SC}$ [$_{DP}$ a man] [sitting outside]]]]]]]

Within an analysis like (2), it appears that the matrix verb *seems* agrees with the noun-phrase *a man* across the boundaries of several maximal projections. However, imagine that the structure in (2) were only the starting point of the derivation of (1b); and that subsequently, the noun-phrase *a man* were to undergo phrasal movement to the specifier position of the matrix TP:[4]

(3) [$_{TP}$ There [$_{DP}$ a man] T^0 [$_{vP}$ seem [$_{VP}$ t$_{V^0}$ [$_{PP}$ to me]
[$_{TP}$ to [$_{vP}$ be [$_{SC}$ t$_{DP}$ sitting outside]]]]]]

At this point, agreement between the finite verb *seem* and the noun-phrase *a man* can be established in a purely local manner, on par with whatever mechanism is used to implement agreement in examples like (1a).[5] Whatever the mechanism behind this relation may be, the crucial point is that the relation is a strictly *local* one: it requires no recourse to a formal operation of *agreement-at-a-distance*. The only operation that applies at any structural distance—other than sisterhood or a *spec-head* configuration—is the movement operation.

Next, suppose the embedded predicate were to undergo phrasal movement out of the matrix *v*P, creating a remnant consisting of the phonological material *seems to me to be*:

(4) [$_{TP}$ There [$_{DP}$ a man] T^0
[$_{XP}$ [$_{Pred}$ sitting outside] X^0 [$_{vP}$ seems [$_{VP}$ t$_{V^0}$ [$_{PP}$ to me] tvs
[$_{TP}$ to [$_{vP}$ be [$_{SC}$ t$_{DP}$ t$_{Pred}$]]]]]]]

4. If one maintains that the expletive *there* syntactically occupies the specifier of TP (see §3.2 for an alternative account for *there*-expletives), then the movement of noun-phrase *a man* would target a second specifier of TP. Alternatively, if one rejects the possibility of multiple specifiers for TP, one could actually base-generate the expletive *there* in a higher position—see fn. 6.

5. If one is inclined to prefer establishing agreement in a *government*-like configuration, rather than *spec-head*, then phrasal movement to [Spec,*v*P] can be substituted for phrasal movement to [Spec,TP] in the analysis present in the text, with negligible effect on the rest of the derivation.

Next, suppose the matrix *v*P undergoes fronting, as in (5):

(5) $[_{YP}\ [_{vP}$ seems $[_{VP}\ t_{V^0}\ [_{PP}$ to me$]\ [_{TP}$ to $[_{vP}$ be $[_{SC}\ t_{DP}\ t_{Pred}]]]]]\ Y^0$
$[_{TP}$ There $[_{DP}$ a man$]\ T^0$
$[_{XP}\ [_{Pred}$ sitting outside$]\ X^0\ t_{vP}]]]$

Finally, the expletive *there* must be moved to the left of *v*P, and into a separate maximal projection than the one hosting the matrix verb (to account for the classic English vs. French contrasts first discussed by Pollock 1989):[6]

(6) $[_{ZP}$ There $Z^0\ [_{YP}\ [_{vP}$ seems $[_{VP}\ t_{V^0}\ [_{PP}$ to me$]\ [_{TP}$ to $[_{vP}$ be
$[_{SC}\ t_{DP}\ t_{Pred}]]]]]\ Y^0\ [_{TP}\ t_{There}\ [_{DP}$ a man$]\ T^0$
$[_{XP}\ [_{Pred}$ sitting outside$]\ X^0\ t_{vP}]]]]$

This approach—which I refer to as *late evacuation*—achieves the effect of Long-Distance Agreement without requiring a formal operation capable of agreement-at-a-distance. However, this achievement comes at a significant cost. First, the various (and elaborate) movement operations required to get from (2) to (6)—not to mention the existence of additional maximal projections, labeled *XP*, *YP* and *ZP*, above—must be independently motivated; otherwise they would be no more than notational devices to "hide" agreement-at-a-distance.[7]

Moreover, there exists an empirical problem with the *late evacuation* approach, having to do with the scopal properties of this derivation. It is by now well-established that in the English expletive-associate construction, the agreed-with noun-phrase (*a man*, in (1b)) has obligatorily narrow scope when it is pronounced downstairs and the matrix subject position is occupied by expletive (Den Dikken 1995; Hornstein 1999). Consider the following examples:

(7) a. Some applicants$_i$ seem to each other$_i$ to be eligible for the job.
b. *There seem to each other$_i$ to be some applicants$_i$ eligible for the job.
[Den Dikken 1995:(2a–b)]

(8) a. Someone$_i$ seems to his$_i$ mother to be eligible for the job.
b. *There seems to his$_i$ mother to be someone$_i$ eligible for the job.
[Den Dikken 1995:(4a–b)]

As (7b, 8b) show, the associate must take narrow scope, at least with respect to material in the subordinating clause.

6. Alternatively, the expletive *there* could simply be base-generated in [Spec,ZP] (see also fn. 4).

7. As an anonymous reviewer points out, the "evacuation" operations themselves (e.g., the operations depicted in (4–6), above) are purely stipulative in nature.

This means that in the example discussed here, the noun-phrase *a man* must scope below the verb *seems*. This is a challenge to the *late evacuation* approach—since in the final step (in (6), repeated below), the noun-phrase *a man* is not within the c-command domain of the verb *seems*:

(6) $[_{ZP}$ There Z^0 $[_{YP}$ $[_{vP}$ seems $[_{VP}$ t_{V^0} $[_{PP}$ to me$]$ $[_{TP}$ to
$[_{vP}$ be $[_{SC}$ t_{DP} $t_{Pred}]]]]]$ Y^0 $[_{TP}$ t_{There} $[_{DP}$ a man$]$ T^0
$[_{XP}$ $[_{Pred}$ sitting outside$]$ X^0 $t_{vP}]]]]$

Obviously, there exists a lower copy of the noun-phrase *a man* which is c-commanded by a lower copy of the verb *seems* (within the base-generated position of the verb-phrase). However, assuming that the noun-phrase *a man* is interpreted in its base-generated position (or anywhere within the base-generated position of the verb-phrase) runs into a problem. Consider the following contrast:

(9) a. [Someone from New York]$_1$ is likely [t$_1$ to win the lottery].
✓[$\exists_{someone\ from\ NY}$ > *likely*], ✓[*likely* > $\exists_{someone\ from\ NY}$]

b. (John said that someone from New York was likely to win the lottery, …)
… and [likely t$_1$ to win the lottery]$_2$ [someone from New York]$_1$ is t$_2$.
✓[$\exists_{someone\ from\ NY}$ > *likely*], ✗ [*likely* > $\exists_{someone\ from\ NY}$]

The sentence in (9a) has two readings: when *[someone from New York]* scopes over *[likely]*, the sentence is about the likelihood of the lottery being won by a particular individual in the actual world (who happens to be from New York); when *[likely]* scopes over *[someone from New York]*, the sentence is about the likelihood of anyone from New York winning the lottery. Interestingly, the sentence in (9b) lacks one of the corresponding readings: it can only refer to a particular individual from New York, rather than referring to anyone from New York. In other words, in (9b), *[someone from New York]* must out-scope *[likely]*; the inverse scopal configuration is unavailable.

As Sauerland and Elbourne show, the scopal properties of sentences like (9b) are an example of a much wider generalization, which they attribute to Barss (1986):

(10) <u>Barss' generalization</u>
Total reconstruction of an A-moved QP to a position X is blocked when the QP does not c-command X in the overt form.
[Sauerland & Elbourne 2002:(36)]

Sauerland and Elbourne account for (10) by assuming that "total reconstruction" is actually a misnomer—that such cases are actually an instance of movement that occurs exclusively within (or during the derivation towards) the phonological component; that such movement is necessarily preceded by movement operations that affect both interfaces (semantic and phonological); and finally, that movement obeys a derivational c-command condition, requiring the landing-site to c-command the launching-site at the point at which movement takes place). (Sauerland and Elbourne assume that lower copies are syntactically inert, in the sense that they cannot themselves launch further movement; for an opposing view, see Chomsky to appear).

Accepting Sauerland and Elbourne's analysis, however, is not crucial for our purposes. Moreover, while (10) mentions "reconstruction", one need not interpret this literally: *Barss' generalization* is an empirical observation about the location in which syntactic objects can and cannot be interpreted; it is formulable even in a framework that completely eschews *reconstruction* as a formal device.[8] We can therefore rephrase *Barss' generalization* as follows:

(11) Barss' generalization (reformulated)
An A-moved QP cannot be interpreted at its base-position if it does not c-command that base-position in the overt form.

Let us now examine the consequences of *Barss' generalization* itself, regardless of its explanation or implementation, for the expletive-associate construction discussed here: movement of the noun-phrase *a man* to [Spec,TP] (in step (3)) is certainly an instance of A-movement; given that the *v*P has been fronted (as in (5)), the landing-site of this A-movement chain has "moved away", in the overt syntax; therefore, reconstruction of the noun-phrase *a man* should be impossible—but this is precisely the position in which the noun-phrase *a man* must be interpreted, in order to generate the correct reading (i.e., narrow scope for the associate noun-phrase, as established by Den Dikken 1995 & Hornstein 1999).

The *late evacuation* approach thus leads to an impasse: to get the correct reading, the associate noun-phrase (*a man*) must be interpreted in a position which runs afoul of *Barss' generalization* (and Sauerland & Elbourne's 2002 account thereof). Obviously, it is conceivable that the expletive-associate construction constitutes the single heretofore-unnoticed exception to *Barss' generalization*. Nevertheless, given that alternative accounts for so-called Long-Distance Agreement avoid this problem, it balances the scales against the *late evacuation* approach.

In fact, while there exist instances of so-called Long-Distance Agreement which do not manifest the obligatory narrow-scope effect exhibited by the expletive associate-construction, the *late evacuation* approach will run into the

8. Such frameworks include Bobaljik's (1995, 2002) "single-output" model of syntax (which forms the basis of the proposal in §3.2), as well as Putnam's (2007) treatment of scrambling in Russian and Japanese within the *Survive Principle* framework.

very same problem with any construction in which narrow scope for the agreed-with noun-phrase is even a possibility—since *Barss' generalization*, when applied to *late evacuation* derivation, will prohibit the agreed-with noun-phrase from taking narrow scope at all. Cases of so-called Long-Distance Agreement in which narrow scope of the agreed-with noun-phrase is a possibility, though not a necessity, are indeed attested. Consider the following example from Hindi-Urdu, given by Bhatt (2005:(65a)):

(12) Naim-ne [har kitaab par:h-nii] chaah-ii (Hindi-Urdu)
Naim-ERG every book$_{\text{FEM}}$ read-INF.FEM want-PRFV.FEM
thii.
be.PAST.FEM.SG
'Naim wants to read every book.'
✓[$\forall_{\text{book}}$ > *want*], ✓[*want* > $\forall_{\text{book}}$]

In (12), the upstairs verb and auxiliary exhibit agreement with the embedded noun-phrase (*har kitaab* 'every book$_{\text{FEM}}$'), whose grammatical gender is feminine. Unlike the English expletive-associate construction, the embedded noun-phrase can have either narrow or wide scope: (12) can mean *for every actual book, Naim has the desire to read it*, or it can mean *Naim's desire is to read every book* (crucially, the latter reading does not require Naim to know which actual books exist).

A potential confound with (12) (and examples like it) is that feminine agreement appears both on the matrix auxiliary and the matrix verb, as well as on the embedded infinitival verb (a pattern typical of Long-Distance Agreement in Hindi-Urdu; see Bhatt 2005). Depending on one's precise assumption regarding the clause-structure of Hindi-Urdu, this might indicate that no *agreement-at-a-distance* takes place whatsoever—but rather each head successively agrees with the head that is structurally adjacent to it (note that this is not the analysis argued for by Bhatt 2005).[9]

An example that does not suffer from this problem is provided by Etxepare (2006:(102a)):

(13) [Leiho guzti-ak ix-te-a] ahaztu (Basque)
Window all-ART$_{\text{PL}}$(ABS) close-NMZ-ART(ABS) forgot
ϕ-zai-zki-o.
3.ABS-be-PL.ABS-3SG.DAT
'(S)he forgot to close all the windows.'
(subject is [*pro*-3SG.DAT]) ✓[$\forall_{\text{window}}$ > *forget*], ✓[*forget* > $\forall_{\text{window}}$]

9. Note that if this were the case—namely, each head successively agreeing with the head that is structurally adjacent to it—then these data from Hindi-Urdu would not bear on the issue at hand, since they would not constitute *agreement-at-a-distance*, even in the superficial sense of the term.

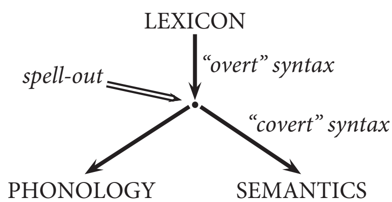
In the Basque sentence in (13), the plurality of the absolutive morpheme on the upstairs auxiliary is determined by the plurality of *leiho guzti-ak* ('window all-ART$_{PL}$(ABS)'), the argument of the downstairs verb *ix* ('close') (Etxepare 2006; Preminger to appear). Again, unlike the English expletive-associate construction, the embedded noun-phrase can have either narrow or wide scope: (13) can mean *for every actual window, (s)he forgot to close it*, or it can mean *(s)he forgot something, and that thing was the closing of the totality of windows* (crucially, the latter reading does not require the subject of *ahaztu* 'forgot' to know which actual windows exist).

Thus, this variety of Basque provides an example of so-called Long-Distance Agreement, in which narrow scope of the agreed-with noun-phrase is a possibility, though not obligatory. As discussed earlier, this too poses a problem for a *late evacuation* account, since deriving the narrow-scope reading would violate *Barss' generalization*.

3.2 *Lower-right corner* chains

At the outset of work within the minimalist program, most research assumed the so-called "inverted-Y" model of syntax (which finds its origin much earlier, in a proposal by Chomsky & Lasnik 1977):

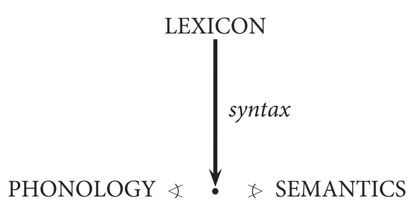
(14) "INVERTED-Y" MODEL:



Within the "inverted-Y" model, *covert* operations are distinguished from *overt* operations by virtue of their ordering, relative to the point of *spell-out*: operations that precede spell-out will be reflected in the phonological output, whereas operations that follow this point will not be reflected in the phonological output.

Bobaljik (1995) later challenged this model, proposing to subsume it under a so-called "single-output" model of syntax:

(15) "SINGLE-OUTPUT" MODEL:



In the "single-output" model, there is no intrinsic ordering between so-called *overt* and *covert* operations; the syntactic derivation proceeds along a single path, ultimately creating a single representation (modulo the cycle). This representation is then read-off by each of the interfaces, phonological and semantic (I will sometimes refer to this as *interpreting* this representation, which is meant to cover both phonological pronunciation and semantic interpretation).

The "single-output" model, along with the *Copy Theory of Movement* (Chomsky 1993, 1995, 2000), allows *overt* and *covert* movement to be recast as mere epiphenomena: *overt* movement describes a situation in which both phonology and semantics interpret the highest copy/link in the movement chain, whereas *covert* movement describes a situation in which the semantics interprets a higher copy/link than does phonology.[10] Crucially, each of the two interfaces (phonological and semantic) is relatively free in choosing which copy it will privilege for interpretation—though not completely free, as will be discussed below.

In his discussion of expletive-associate constructions and A-movement, Bobaljik (2002) points out that given the assumptions of the "single-output" model of syntax—and in particular, that each interface is relatively free in choosing which link in a movement chain to "interpret"—one would expect to find instances of movement in which both interfaces happen to privilege the same lower copy for pronunciation/interpretation. Bobaljik argues that this state of affairs, though very similar to non-movement, can in fact be distinguished from non-movement by virtue of the effect that higher copies have on purely syntactic operations (among which he includes the checking of formal features).

Bobaljik proposes that this state of affairs—which he terms a *Lower-Right Corner* configuration—is in fact the correct analysis of the expletive-associate construction in English.[11] Recall (1a–b), repeated below:

(1) a. A man seem*(s) to be sitting outside.
    b. There seem*(s) to be a man sitting outside.

---

10. This contrasts with the traditional Minimalist Program framework (Chomsky 1995 *et seq.*), in which the difference between overt movement and covert movement derives from "features-strength": *strong* features must be checked prior to spelling out the derivation to PF, and therefore result in movement that is phonologically detectable; checking of *weak* features, on the other hand, can be deferred until after spell-out, resulting in phonologically undetectable movement. Thus, within traditional minimalism, overt movement and covert movement are essentially primitives, encoded in the form of the feature-strength of those features responsible for triggering each particular instance of movement.

11. See Bobaljik (2002), regarding the reason for the term *Lower-Right Corner*.

Let us assume that (2), repeated below, is the correct structural analysis of (1b):

(2) [$_{\text{TP}}$ There [$_{v\text{P}}$ seems [$_{\text{VP}}$ [$_{\text{TP}}$ to [$_{v\text{P}}$ be [$_{\text{SC}}$ [$_{\text{DP}}$ a man] [sitting outside]]]]]]]

Bobaljik assumes that the EPP in English is a phonological requirement on having pronounced material in subject position (an idea developed in greater detail by Landau 2007)—thus relegating the insertion of *there* in (1b, 2) to the phonological component—and that in (1b), there actually is an A-chain relating the base-position of the noun-phrase *a man* to [Spec,TP] of the matrix clause:[12]

(16) [$_{\text{TP}}$ [$_{\text{DP}}$ a man]$_1$ [$_{v\text{P}}$ seems [$_{\text{VP}}$ [$_{\text{TP}}$ to
[$_{v\text{P}}$ be [$_{\text{SC}}$ [$_{\text{DP}}$ a man]$_1$ [sitting outside]]]]]]]

If the phonological component pronounces the lower copy of the noun-phrase *a man*, rather than the higher copy, the aforementioned phonologically-oriented EPP requirement imposes pronunciation of an expletive element (*there*) in the matrix subject position. The scopal properties of this construction (as discussed in §3.1) are handled in terms of an economy-based preference for both interfaces—phonological and semantic—to privilege the same copy in a chain (see also Diesing's 1997 *Scope Principle*):

(17) Minimize Mismatch
(To the extent possible) privilege the same copy at PF and LF.
[Bobaljik 2002:(70)]

According to Bobaljik, the principle in (17) is evaluated with respect to a particular intended interpretation. When the intended interpretation is one in which the agreed-with noun-phrase has narrow scope, the principle in (17) stands in competition with the phonologically-oriented EPP: pronouncing the higher copy violates (17) but satisfies the EPP; pronouncing the lower copy satisfies (17), but requires insertion of an expletive element in the matrix subject position (which Bobaljik assumes is economically dispreferred). On the other hand, when the intended interpretation is one in which the agreed-with noun-phrase has wide scope, pronouncing the higher copy satisfies both constraints; pronouncing the lower copy is therefore ruled out.

Crucially, regardless of whether the noun-phrase *a man* is pronounced in the higher position or the lower position, syntactic agreement is established between the verb *seems* and the higher copy of the noun-phrase *a man* (which is present in the

12. The schematization in (16) is meant to abstract away from whether or not there is an intermediate copy of the noun-phrase *a man* in the embedded [Spec,TP] position.

matrix [Spec,TP] regardless of whether it is or is not phonologically pronounced there). Just like the *late evacuation* approach, this analysis requires no recourse to a formal operation of *agreement-at-a-distance*. The only operation that applies at any structural distance—other than sisterhood or a *spec-head* configuration—is the movement operation.

To summarize, the *Lower-Right Corner* analysis takes so-called Long-Distance Agreement to be an instance of movement in which both the phonological and the semantic components privilege the lower copy for pronunciation/interpretation. Such movement chains are distinguished from instances of non-movement precisely by virtue of the effects of higher copies on purely syntactic operations—and in particular, feature-checking. Thus, agreement (an instance of feature-checking) can obtain in a purely local fashion, targeting the unpronounced (and semantically uninterpreted) higher copy of the movement chain—and as a result, the impression of Long-Distance Agreement is created.

Moreover, since this approach handles the scopal properties of the expletive-associate construction in a way that is intimately tied to the EPP, it might extend to other instances of so-called Long-Distance Agreement that have different scopal properties, depending on the the particular constraints that come into play. For example, neither Hindi-Urdu nor Basque (both of which exhibit Long-Distance Agreement-like constructions with different scopal properties than the English expletive-associate construction; see §3.1) exhibit anything reminiscent of a phonologically-oriented EPP constraint. Therefore, the set of interacting constraints in those languages cannot be the same as in English—a difference which may offer the beginnings of an account for their different behavior with respect to scope.[13]

It is worth pointing out that a very similar approach is taken by Putnam (2009) in his discussion of apparent Long-Distance Agreement and dative intervention phenomena in Icelandic, within the *Survive Principle* framework. Obviously, the full range of data addressed by Putnam cannot be repeated here in their full detail; however, one example of the construction he discusses is given below:

(18) θað hafa komið hingað θrír málvísindamenn (Icelandic)
there have.3PL come here three linguists.NOM.MASC.PL
'There came three linguists.' *[Sigurðsson 2008:61;(1)]*

13. The details of such an account would obviously need to be worked out carefully—a task that is beyond the scope of this paper. For the current purposes, the important point is that the *Lower-Right Corner* approach can account for the scopal properties of the English expletive-associate construction, without completely inhibiting the ability to extend the account to other instances of so-called Long-Distance Agreement, which exhibit different scopal properties (and of course, without appealing to an operation of *agreement-at-a-distance*).

Putnam argues that the downstairs nominative subject *málvísindamenn* ('linguists. NOM.MASC.PL') is in fact remerged in a position structurally adjacent to the matrix auxiliary (or more accurately, adjacent to the position in which the number morpheme that surfaces on the auxiliary is base-generated). Thus, agreement between the matrix auxiliary *hafa* ('have.3PL') and this nominative noun-phrase is established in a purely local manner. It is the phonological component, however, that determines that this element will be pronounced in its base-position. Just as in the *Lower-Right Corner* approach sketched here, Putnam takes this decision to be the result of computation by the phonological component, rather than narrow-syntax.

4. Summary

In this paper, I have examined two alternative accounts for so-called Long-Distance Agreement, both of which do not involve actual recourse to a formal operation of *agreement-at-a-distance*. One account, which I have labeled the *late evacuation* approach, involves establishing agreement in a local fashion, followed by the agreeing elements moving away, thus creating the appearance of Long-Distance Agreement. The second account, which I have labeled the *Lower-Right Corner* approach (following Bobaljik 2002), involves a particular kind of covert movement: a movement chain in which both the phonological and the semantic component privilege the lower copy of the chain for pronunciation/interpretation; therefore, the higher copy has an effect only on syntax itself, not on the derivation. A prime example of such narrowly-syntactic effects, of course, is feature checking—and in particular, checking of ϕ-features, otherwise known as *agreement*.

As shown in §3.1, the *late evacuation* approach suffers from a problem, as far as the semantic scope of the agreed-with noun-phrase is concerned: deriving narrow scope for the agreed-with noun-phrase violates *Barss' generalization*. This is particularly problematic in the English expletive-associate construction, where the agreed-with noun-phrase has obligatory narrow scope (Den Dikken 1995; Hornstein 1999); but the problem extends to any instance of so-called Long-Distance Agreement in which narrow scope of the agreed-with noun-phrase is even a possibility (such as Hindi-Urdu and Basque; Bhatt 2005 & Etxepare 2006, respectively). As discussed in §3.2, the *Lower-Right Corner* approach can account for the scopal properties of the English expletive-associate construction (as proposed by Bobaljik 2002), and in principle, may also extend to instances of so-called Long-Distance Agreement where the agreed-with noun-phrase has narrow scope only optionally—especially if these occur in languages that do not exhibit an English like phonologically-oriented EPP (and both Hindi-Urdu and Basque fall within this category).

It therefore appears that the *Lower-Right Corner* approach offers a promising avenue of accounting for so-called Long-Distance Agreement, without appealing to a formal operation of *agreement-at-a-distance*.[14]

With respect to the theoretical landscape outlined in sections §1–§2, this is a particularly beneficial result. It alleviates the duplication problem discussed in section §2—namely, the need for a second, unrelated mechanism of addressing the formal needs of a structurally-embedded constituent, alongside syntactic movement—allowing an account of Long-Distance Agreement without appealing to a formal operation of *agreement-at-a-distance*.

As a result, abandoning *drag-movement* in favor of other models of movement no longer incurs a "penalty", in terms of theoretical simplicity. This removes Long-Distance Agreement as an empirical obstacle to frameworks that eschew drag-movement, such as the Survive Principle framework (Putnam 2007; Putnam & Stroik to appear, Stroik 1999), "repel"-based movement systems (e.g., van Craenenbroeck 2006; Platzack 1996; van Riemsdijk 1997), and others.

References

Ackema, P., Neeleman, A. & Weerman, F. 1993. Deriving functional projections. In *Proceedings of the 23rd Conference of the North EastLinguistic Society (NELS 23)*, A. Schafer (Ed.), 17–31. Amherst MA: GLSA.

Barss, A. 1986. Chains and anaphoric dependence: On reconstruction and its implications. Ph.D. dissertation, MIT.

Bhatt, R. 2005. Long-distance agreement in Hindi-Urdu. *Natural Language and Linguistic Theory* 23: 757–807.

Bobaljik, J.D. 1995. Morphosyntax: The syntax of verbal inflection. Ph.D. dissertation, MIT.

Bobaljik, J.D. 2002. A-chains at the PF interface: Copies and 'covert' movement. *Natural Language and Linguistic Theory* 20: 197–267.

Chomsky, N. 1993. A minimalist program for linguistic theory. In *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, K. Hale & S.J. Keyser (Eds), 1–52. Cambridge MA: The MIT Press.

Chomsky, N. 1995. *The minimalist program*. Cambridge MA: The MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In *Step by step:Essays on minimalist syntax in honor of Howard Lasnik*, R. Martin, D. Michaels & J. Uriagereka, 89–155. Cambridge MA: The MIT Press.

Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A life in language*, M. Kenstowicz (Ed.), 1–52. Cambridge MA: The MIT Press.

14. As pointed out by an anonymous reviewer, the same general approach may provide the basis of an account for apparent Long-Distance Agreement effects in non-ϕ features, such as [Topic] or [Focus].

Chomsky, N. 2004. Beyond explanatory adequacy. In *The cartography of syntactic structures,* Vol. 3, *Structures and beyond*, 104–131. New York NY: OUP.

Chomsky, N. To appear. On phases. In *Foundational issues in linguistic theory*, R. Freidin, C. Otero & M.-L. Zubizaretta (Eds). Cambridge MA: The MIT Press.

Chomsky, N. & Lasnik, H. 1977. Filters and control. *Linguistic Inquiry* 8: 425–504.

van Craenenbroeck, J. 2006. Transitivity failures in the left periphery and foot-driven movement operations. In *Linguistics in the Netherlands 2006*, J. van de Weijer & B. Los, 52–64. Amsterdam: John Benjamins.

Den Dikken, M. 1995. Binding, expletives, and levels. *Linguistic Inquiry* 26: 347–354.

Diesing, M. 1997. Yiddish VP order and typology of object movement in Germanic. *Natural Language and Linguistic Theory* 15: 369–427.

Etxepare, R. 2006. Number long distance agreement in (substandard) Basque. In *Studies in Basque and historical linguistics in memory of R.L. Trask. Supplements of the Anuario del Seminario de Filologia Vasca "Julio de Urquijo" XL:1–2*, ed. Joseba A. Lakarra & Jose Ignacio Hualde, 303–350.

Frampton, J. 1991. Relativized minimality: A review. *The Linguistic Review* 8: 1–46.

Hornstein, N. 1999. Minimalism and quantifier raising. In *Working minimalism*, S.D. Epstein & N. Hornstein, 44–75. Cambridge MA: The MIT Press.

Koeneman, O. 2000. The flexible nature of verb movement. Ph.D. dissertation, UiL-OTS, Utrecht: LOT Dissertation Series.

Landau, I. 2007. EPP extensions. *Linguistic Inquiry* 38: 485–483.

Moro, A. 2007. Some notes on unstable structures. http://ling.auf.net/lingBuzz/000454, Ms, Cambridge, MA Milan: Harvard University & Universita Vita-Salute San Raffaele.

Nash, L. & Rouveret, R. 1997. Proxy categories in phrase structure theory. In *Proceedings of the 27th Conference of the North East Linguistic Society (NELS 27)*, Kiyomi Kusumoto (Ed.), 287–304. Amherst MA: GLSA.

Nilsen, O. 1997. Adverbs and A-shift. In *Working Papers in Scandinavian Syntax* 59: 1–31.

Platzack, C. 1996. Germanic verb second languages – Attract vs. repel: On optionality, A-bar movement and the symmetrical/asymmetrical verb second hypothesis. In *Deutsch – Typologisch*, E. Lang & G. Zifonun, 92–120. Berlin: Walter de Gruyter.

Pollock, J.-Y. 1989. Verb movement, UG and the structure of IP. *Linguistic Inquiry 2*0: 365–424.

Preminger, O. 2007. Toxic syntax. Ms, MIT.

Preminger, O. To appear. Breaking agreements: Distinguishing agreement and clitic-doubling by their failures. *Linguistic Inquiry*.

Putnam, M.T. 2007. *Scrambling and the survive principle* [Linguistik Aktuell/Linguistics Today 115]. Amsterdam: John Benjamins.

Putnam, M.T. 2009. Late agreement in Icelandic from a Survive perspective. *Ms. Carson. Newman College*.

Putnam, M.T. & Stroik, T.S. To appear. Surviving reconstruction. In *Phases at the interface*, K.K. Grohmann (Ed.). Berlin: Mouton de Gruyter.

van Riemsdijk, H. 1997. Push chains and drag chains: Complex predicate split in Dutch. In *Scrambling*, S. Tonoike (Ed.), 7–33. Tokyo: Kurosio Publishers.

Rizzi, L. 1990. *Relativized minimality*. Cambridge MA: The MIT Press.

Sauerland, U. & Elbourne, P. 2002. Total reconstruction, PF movement, and derivational order. *Linguistic Inquiry* 33: 283–319.

Sigurdsson, H.A. 2000. The locus of case and agreement *Working Papers in Scandinavian Syntax* 65: 65–108.

Sigurdsson, H.A. 2004. Agree and agreement: Evidence from Germanic. In *Focus on Germanic typology*, W. Abraham, 61–103. Berlin: Akademie Verlag.

Sigurdsson, H.A. 2006. Remarks on features. Ms, http://ling.auf.net/lingBuzz/000334.

Sigurdsson, H.A & Holmberg, A. 2008. Icelandic dative intervention: Person and Number are separate probes. In *Agreement restrictions*, R. d'Alessandro, G.H. Hrafnbjargarson & S. Fischer, 251–280. Berlin: Mouton de Gruyter.

Stroik, T.S. 1999. The survive principle. *Linguistic Analysis* 29: 278–303.

Suranyi, B. 2004. The left periphery and cyclic spell-out: The case of Hungarian. In *Peripheries: Syntactic edges and their effects*, D. Adger, C. de Cat & G. Tsoulas, 49–74. Dordrecht: Kluwer.

Musings on the left periphery in West Germanic

German Left Dislocation and 'Survive'*

Gema Chocano
Universidad Autónoma de Madrid

An analysis of German Left Dislocation is proposed which combines the findings in recent work on the construction with a theoretical model of syntactic derivation that dispenses with the existence of the EPP-feature. The findings at stake are basically those in Frey (2000, 2004a), where it is shown that the topical character of left-dislocated structures is the result of movement of the D-pronoun to a middle-field, and not a pre-field, topic projection. The theoretical model adopted is 'Survive Minimalism' (Stroik, 2009; Putnam, 2007; Putnam and Stroik, in progress). The core of the proposal is that 'Merge' of both the D-pronoun and the left-dislocated XP with C is triggered by the presence of a [+REF] feature on the three heads, as in Stroik's (2009) account of English *wh*-constructions with a pair-list reading.

1. Introduction

According to the traditional analysis of verb second (V2) as proposed in Thiersch (1978), den Besten (1989) and many others, root clauses in German are the result of movement of both the finite verb and the XP that precedes it to the CP projection. With the advent of Chomsky's 'Minimalist Program', where feature checking is a requirement for strictly syntactic movement, the focus has been on the exact nature of the features involved in V2 (Fanselow, 2003; Gärtner, 2002; Roberts & Roussou, 2002; Bayer, 2004; Brandner, 2004; Truckenbrodt, 2006; among others). Following that trend, two recent studies on the German left

*I am indebted to Michael Putnam for inspiring and constructive suggestions, as well as for helpful discussion of some of the basic tenets of 'Survive'. I also thank Carlos Piera and an anonymous internal reviewer for useful comments. Any shortcomings are my sole responsibility.

periphery, Fanselow (2004) and Frey (2004, 2005, 2006), conclude that V2 structures fall into two main types:

i. V2 structures in which displacement to the CP projection correlates with the checking of a feature of semantic/pragmatic import on both C and the preposed XP, as in (1) below, a standard instance of focalization:

(1) a. *Ein BUCH hat er gelesen*
a-ACC book has he-NOM read
'A book, he has read'

b. $_{CP}[XP_{j[+focus]}\ _{C'}[V_{fin\ i}\ _{IP}[t_j \ldots t_i \ldots]]]$

ii. V2 structures in which the only feature satisfied is an EPP feature on C (2a), along the lines of the proposal in (2b), originally made by Bhatt (1999) for Kashmiri. Fanselow calls the syntactic operation at stake 'Stylistic Fronting',[1] whereas Frey labels it 'Formal Movement'.

(2) a. *Einem Studenten ist ein Fehler aufgefallen*
a-DAT student is a-NOM mistake struck
'A student noticed a mistake'

b. $_{CP}[XP\ _{C'}[V_{fin}\ _{IP}[\ldots.]]]$ can be derived from $_{C'}[_{IP}[XP\ldots V_{fin}]]$ (by moving the finite verb to C and) by moving XP as the highest element of IP to SpecCP (Minimal Link) just in order to fulfill the EPP requirement associated with C, i.e., without additional intonational, semantic, or pragmatic effects.

'Stylistic Fronting'/'Formal Movement' is, according to both Fanselow and Frey, responsible for what has been traditionally considered topicalization in the German pre-field, on the basis of the combination of (2b) and the findings in Frey's previous studies (Frey, 2000; 2004b), where the existence of a topic projection in the German middle-field is defended:

(3) In the middle field of the German clause, directly above the base position of sentential adverbials (SADVs), there is a designated position for topics (in the *aboutness*-sense):[2] all topical phrases occurring in the middle field, and only these, occur in this position (2004a: 208).[3]

1. For reasons of parallelism with the analysis of 'Stylistic Fronting' in Icelandic developed in Holmberg (2000).

2. According to Reinhart's (1981) definition in (i), grounded, in turn, on Hockett's (1958) famous characterization in (ii):

(i) A topic is an expression whose referent the sentence is about. The concept "topic" is a category of pragmatic aboutness.

Thus, a sentence such as (4a) reduces to the structure in (4b), where, as in Frey (2006), the cartographic approach to the left periphery in Rizzi (1997) is adopted. The EPP feature on C is now an EPP feature on Fin, the head of Finite Phrase (FinP), and verb movement is possible only to the highest realized head that carries an EPP feature:

(4) a. *Den Hans wird erfreulicherweise nächtes Jahr eine polnische*
the-ACC Hans will happily next year a-NOM Polish-NOM
Gräfin heiraten
countess marry
'Hans, a Polish countess will marry (him) next year'

b. $_{CP}$[$_{C'}$[$_{FinP}$[den Hans$_{j}$ $_{Fin'}$[wird$_{i}$ $_{TopP}$[t$_{j}$ $_{Top'}$[t$_{i}$ $_{TP}$[...t$_{i}$ t$_{j}$ ...]]]]]]

(ii) "the most general characteristic of predicative construction is suggested by the terms 'topic' and comment for their ICs [immediate constituents]: the speaker announces a topic and then says something about it" (1958: 201).

3. Frey's main arguments are those examples where the context forces 'aboutness' (i), the ill-formedness of the structures in which a non-referential expression precedes a sentence adverbial (ii), and the fact that co-reference with a cataphoric pronoun is restricted to elements occupying that high position:

(i) *Da wir gerade von Hans sprechen*
since we right now about Hans talk
'Since we are talking about Hans right now'

Nächstes Jahr wird √***den Hans*** *erfreulicherweise* ****den Hans*** *eine*
next year will the-ACC Hans fortunately the-ACC Hans a-NOM
vornehme Dame heiraten
fine lady marry
'Next year a fine lady will marry Hans'

(ii) *Heute hat* ****fast jeder*** *erstaunlicherweise* √***fast***
today has almost everyone-NOM surprisingly almost
jeder *gearbeitet*
everyone-NOM worked
'Today almost everyone has worked surprisingly'

(iii) *Sein$_{i}$ Vater wird* √***dem Otto***$_{i}$ *wahrscheinlich* ****dem Otto***$_{i}$ *das*
his-NOM father will the-DAT Otto probably the-DAT Otto the-ACC
Auto ausleihen
car lend
'Probably, Otto's father will lend him the car'

The proposal above has the empirical advantage of accounting for asymmetries such as the one in (5), where it is shown that, in contrast with the optionality found in the case of simple sentences, long-distance topicalization obligatorily correlates with a pitch accent on the topicalized constituent. Notice that the interpretation of the preposed constituent as a topic is guaranteed by the presence of the introductory sentence. (Pitch accent signaled by capital letters):

(5) *Ich erzähle dir was über Max*
'I will tell you something about Max'

a. *Den Max / Den MAX sollte der Chef mitnehmen*
the-ACC Max / the-ACC Max should the-NOM boss take-along
'Max / MAX, the boss should take along'

b. **Den Max* / √*Den MAX meint Eva, dass der Chef mitnehmen sollte*
the-ACC Max / the-ACC Max thinks Eva that the-NOM boss take-along should
'Max / MAX, Eva thinks that the boss should take along'

According to Frey, the variant of (5a) in which *den Max* receives a pitch accent is possible only if *den Max* is a member of a set of alternatives. This leads Frey to conclude that movement to the German pre-field is semantically/pragmatically restricted to the checking of the feature [+contrast],[4] and to treat standard cases of focalization such as (1a) in these terms.[5]

4. The [+contrast] feature on the preposed XP will be checked by a matching feature of the head of Kontr(ast) Phrase, a functional projection below CP and above FinP.

5. Fanselow's (2004) account of focalization structures maintains the existence of a [+focus] feature on C, but departs from the traditional view in the light of data like those in (i)–(iii), which Fanselow considers equally well-formed as answers to the question *What did you do last week*?, which requires focus on VP:

(i) *Ich hab ein BUCH gelesen*
I-NOM have a-ACC book read
'I read a book'

(ii) *Ein BUCH gelesen hab ich*
a-ACC book read have I-NOM
'I read a book'

(iii) *Ein BUCH hab ich gelesen*
a-ACC book have I-NOM read
'I read a book'

Although empirically well-motivated, the analysis of V2 along the lines shown in the preceding paragraphs is inextricably linked to a highly controversial theoretical construct, namely the EPP feature. This feature has been criticized mainly on two grounds. On the one hand, as is argued in Lasnik (2001) and Epstein and Seely (2002, 2006), the EPP is incompatible with the 'Interpretability Condition', which requires that "lexical items have no features other than those interpreted at the interface, properties of sound and meaning" (Chomsky, 2000: 113). On the other, it exhibits what Stroik (2009) calls "an unusual strength", illustrated by the contrast in (6) (from Stroik, 2009):

(6) a. *Where does Mary wonder (where) John put the money (where)?
b. Sam wants PRO to appear (PRO) to have (PRO) left

The *wh*-feature of the matrix clause is not strong enough to attract the *wh*-feature on *where*, in the specifier of the embedded clause. As Stroik puts it, no feature interpreted at the interfaces undergoes iterated attraction. However, in orthodox minimalist accounts, the EPP feature of the higher T suffices to make PRO undergo iterated attraction to its final landing site.

The present paper is an attempt to keep the empirical advantages of Frey's (2004, 2006) proposal for the German left periphery while dispensing completely with the EPP feature. For reasons of space, it will be limited to the analysis of German Left Dislocation (GLD), which, I hope, will ultimately prove useful, in so far as an account for this construction presents some difficulties that are not found in 'plain' topicalization or focalization/contrast structures. The theoretical framework adopted will be that of 'Survive Minimalism' as developed in Stroik (2009), Putnam (2007), Putnam and Stroik (in progress), and references therein.

2. German left dislocation and the EPP

German Left Dislocation combines the presence of an XP in the left periphery with the presence of an agreeing D-pronoun, sometimes viewed as a resumptive (RP). Whereas the former invariably occupies the highest position of the clause, the latter may immediately follow it (7a) or even appear in a lower, middle-field

Fanselow analyzes (iii) as an instance of the operation he calls 'Part pro toto movement', in which only part of the phrase endowed with the attracted feature moves. In the case at stake, movement of the object suffices to establish the required checking relation between the [+focus] feature on the VP and that on C. A critical evaluation of Fanselow's innovative proposal lies outside the scope of this paper (but see Frey, 2005). Therefore, in the following, I will simply refer to Frey's more conventional analysis.

position (7b). Notice that, at first sight, both cases constitute an exception to the V2 constraint, in that the finite verb is preceded by two phrasal constituents.[6]

(7) a.

[Den	*Studenten],*	√*[den] /*				
the-ACC.MASC.SING	student	RP-ACC.MASC.SING				
**[das]*	*der*	*Peter*	*zum Glück*	*gesehen*	*hat*	
RP-ACC.NEUT.SING	the-NOM	Peter	fortunately	seen	has	

'The student, fortunately Peter saw (√him/*it)'

b.

[Den	*Studenten],*	*der*	*Peter*	*hat*
the-ACC.MASC.SING	student	the-NOM	Peter	has
[den]	*zum Glück*	*gesehen*		
RP-ACC.MASC.SING	fortunately	seen		

'The student, fortunately Peter saw (him)'

Most researchers argue that the RP reaches its final landing site by movement (Grohmann, 1997, 2000, 2003; Grewendorf, 2002; Frey, 2004a, among others). This conclusion is based on the observation that GLD structures show island sensitivity[7] (8a) as well as connectivity effects (8b), (examples from Grohmann, 2000):

(8) a.

**Den*	*schönen*	*Mann,*	*den*	*haßt*	*Martin*	*die*	*Tatsache*	*daß*
the-ACC	handsome	man	RP-ACC	hates	Martin	the	fact	that
die	*kluge*	*Frau*	*geküßt*	*hat*				
the-NOM	smart	woman	kissed	hat				

*'The handsome man, Martin hates the fact that the smart woman kissed'

b.

Den	*Wagen*	*von*	*sich*$_i$,	*den*	*hat*	*er*$_i$	*verkauft*
the-ACC	car	of	himself	RP-ACC	has	he-NOM	sold

'The car of himself, he sold'

Researchers, however, disagree as to the nature of the trigger for that movement as well as the type of syntactic operation responsible for placing the preposed constituent in the position it appears in. In the following, we will examine two proposals that capitalize on the presence of an EPP feature with the aim to show that, besides

6. A potential treatment of the sequence XP-RP as a complete, unbroken constituent—a possibility in (7a)—is absolutely excluded in (7b).

7. Although Grewendorf (2002) does not consider GLD constructions to be islands. His conclusion is, nevertheless, based on examples that may be explained otherwise, as Frey (2004) correctly notices.

the theoretical problems posed in the previous section, accounting for GLD in those terms also leaves some empirical issues unexplained.

2.1 Grewendorf (2002)

Grewendorf (2002) concurs with other analyses of GLD in considering it a clearly topic marking construction. Grewendorf thus aligns himself with the proposal in Vat (1997), where it is assumed that both the left dislocated XP and the D-pronoun are base-generated, forming a single constituent (a "Big DP", with the D-pronoun as its head and the subsequently left dislocated XP as its specifier). This "Big DP" undergoes movement to the specifier of FinP, the lowest projection in Rizzi's (1997) enriched CP-layer, where it checks an EPP feature. After that, the XP will undergo further movement to the specifier of TopP, which would explain the topical character of the construction:

(9) a. *Den Studenten, den hat der Peter gestern gesehen*
the-ACC student RP-ACC has the-NOM Peter yesterday seen
'The student, Peter saw (him) yesterday'

b. $_{CP}[\ _{C'}[_{TopP}[\text{den Studenten}_{k\ Top'}[_{FinP}[_{DP}[t_k\ \text{den}]_j\ [_{Fin'}[\text{hat}_{i\ TP}[\ldots t_j\ t_i \ldots]]]]]]]$

Grewendorf's proposal is criticized by Frey (2004a) on several grounds, among which are: (i) the lack of an account for those structures in which the RP does not precede the finite verb (7b); and (ii) the unexpected grammaticality of the extraction of an XP out of a "Big DP" sitting in a pre-field position, given that, as illustrated in (10), pre-field positions seem to be strict islands for movement in German:

(10) **Über wen glaubt Hans ein Buch hat Maria*
about whom-ACC thinks Hans-NOM a-ACC book has Maria-NOM
heute gelesen?
today read
*'About whom does Hans think that a book Maria has read today?'

A further objection comes from the contrast found in V2 embedded clauses, illustrated in (11), and the way Grewendorf accounts for it:

(11) a. **Den Studenten hat der Peter den gestern gesehen*
the-ACC student has the-NOM Peter RP-ACC yesterday praised
'The student, Peter saw (him) yesterday'

b. *Den Studenten glaubt Hans den hat der Peter*
the-ACC student thinks Hans-NOM RP-ACC has the-NOM Peter
gestern gesehen
yesterday seen
'The student, Hans thinks Peter saw (him) yesterday'

The example in (11a) is an ill-formed variant of the grammatical structure in (7b), where the RP follows the finite verb instead of immediately preceding it. Notice that the only difference between them is that related to the V2 constraint, which, apparently, cannot be satisfied by the mere presence of the left dislocated XP. But, crucially, such presence suffices to comply with it in (11b).[8] Grewendorf solves the paradox by resorting to an embedded TopP which is responsible for checking the [+topic] feature of the XP, which is subsequently moved to satisfy the EPP feature on the matrix Fin, as in standard cases of topicalization. But this explanation fails for two reasons: (i) the undesirable complications that derive from the unexpected asymmetry between the left dislocated topic licensed in the embedded TopP and those topics that, in root clauses, are licensed in the matrix FinP (Haegeman, 1997; Roberts, 1999; Roberts & Roussou, 2002); and (ii) as Frey notices (p.c. to Grewendorf), the absence of an explanation why the lexical XP cannot check the EPP feature on the embedded Fin, given that this feature is not associated to any other semantic/pragmatic feature. In this respect, it must be noticed that, in the absence of further stipulations, (11a)[9] should be fully grammatical.

2.2 Frey (2004a)

Contrary to Grewendorf, who leaves the structures with the D-pronoun in a lower position unexplained, Frey (2004) capitalizes on them to defend the idea that GLD is uniformly derived by movement of the D-pronoun to a topic projection in the middle-field (see (3) above), followed by 'Formal Movement' of the highest constituent to satisfy the EPP feature on Fin. In this light, the differences between (7a), i.e., a structure with the D-pronoun immediately preceding the finite verb, and (7b), i.e., a structure with the D-pronoun following the finite verb, would reduce

8. An anonymous reviewer suggests that the fact that the left-dislocated XP satisfies the V2 constraint should be taken to indicate that *glaubt Hans* in (11b) is a parenthetical, along the lines of Reis (1996). Since Grohmann (2000) shows that the parenthetical analysis is problematic in some respects, the present paper adopts the traditional view and treats examples such as (11b) as true instances of embedding, thus following Grohmann himself, von Stechow and Sternefeld (1988), Müller and Sternefeld (1993), Grewendorf (2002), and Frey (2004, 2004a, 2005), among others.

9. Grewendorf tentatively suggests that the feature set of the Fin head is complete, which would provoke the checking of all features on the displaced constituent. In this view, the ungrammaticality of (11a) would reduce to unchecking of the uninterpretable [+topic] feature on Top.

to the absence or the presence of an additional, higher element, probably a topic too, as shown in (12):[10]

(12) a. $_{CP}[\ _{C'}[_{FinP}[\text{den}_{Studenten;\ Fin'}[\text{hat}_{i\ TopP}[t_i\ [_{TopP}[\text{der Peter}_{k\ Top'}[\ t_{i\ TP}[\dots\ t_k\ t_j\ t_i\ ]]]]]]]]]$

b. $_{CP}[\ _{C'}[_{FinP}[\text{der Peter}_{k\ Fin'}[\text{hat}_{i\ TopP}[t_k\ [_{TopP}[\text{den}_{Studenten;\ Top'}[\ t_{i\ TP}[\dots\ t_k\ t_j\ t_i\ ]]]]]]]]]$

With respect to the left dislocated XP, Frey concurs with previous analyses (Demirdache, 1991; Anagnostopoulou, 1997; Wiltschko, 1997) and assumes that it is base-generated in its surface position, where it forms an A-bar-CHAIN with the D-pronoun.[11] Thus, connectivity effects and island sensitivity are derived.

As in Grewendorf's approach, one of the main problems for Frey is the set of data in (11), repeated here for convenience as (13):

(13) a. **Den Studenten hat der Peter den gestern gesehen*
the-ACC student has the-NOM Peter RP-ACC yesterday praised
'The student, Peter saw (him) yesterday'

b. *Den Studenten glaubt Hans den hat der Peter gestern gesehen*
the-ACC student thinks Hans-NOM RP-ACC has the-NOM Peter yesterday praised
'The student, Hans thinks Peter saw (him) yesterday'

The underlying assumption, made explicit in Frey (2004, 2005, 2006), is that movement of the finite verb to the C-domain is possible only to a head which carries an EPP feature and, at the same time, is the highest head in the structure. Since in (13a) *den Studenten* is base-generated as an adjunct to the CP, the verb must appear in the activated Fin, whose EPP would have attracted either the subject *(Den Studenten, keiner hat den gelobt)* or the D-pronoun *(Den Studenten, den hat keiner gelobt)*. On the other hand, according to Frey's own arguments, (13b) must necessarily be a case in which *den Studenten* occupies a position inside the CP-projection. Notice, however, that there is no apparent trigger for movement to that position: (i) the preposed XP lacks a topic feature that may trigger its movement to the matrix TopP, since it is a topic only to the extent that it forms a CHAIN with the D-pronoun; and (ii) if it were base-generated as an adjunct to

10. An anonymous reviewer suggests that there would be a prosodic difference between (12a) and (12b), namely the presence of focus accent on the object *den Studenten* in (12a), absent in the subject *der Peter* in (12b). However, as Frey (2004a, 2005) correctly notices, focus accent on the object is obligatory only in the case that the object is contrastive.

11. A-CHAIN $\langle \alpha_i \dots \alpha_n \rangle$ is a sequence of nodes sharing the same θ-role such that for any i, $1 \leq i < n$, α_i c-commands and is coindexed with α_{i+1} (Frey, 2004: 223).

the embedded CP, it would be too distant from the matrix Fin, which would make it an unsuitable candidate for the checking of the relevant EPP feature.

2.3 Interim conclusions

As shown in the preceding paragraphs, Grewendorf's and Frey's analyses of GLD structures are similar in some important respects: (i) GLD is treated as a topic-marking structure; (ii) high occurrences of the D-pronoun are linked to Spec, Fin, the lowest projection of Rizzi's (1997) enriched CP-layer; and (iii) movement of the D-pronoun to FinP is triggered by an EPP feature. A further similarity is that neither of them succeeds in accounting for the problems posed by embedded structures.

On the other hand, they also differ in two related aspects: (i) whereas for Grewendorf the element that bears the topic feature is the left dislocated XP, for Frey that element is the D-pronoun; and (ii) whereas Grewendorf assigns the left-dislocated XP a CP-internal position (Spec, Top), Frey treats it as CP-adjoined. In my opinion, Frey's account is superior in both respects, for reasons that will be given in the following.

Regarding (i), Frey himself (2004a) provides solid evidence for the topical character of the D-pronoun. His argument is based on the contrast between GLD and Hanging Topic Left Dislocation (HTLD) constructions as far as the position occupied by the D-pronoun is concerned. Although one of the variants of HTLD[12] is superficially identical to GLD, it may be distinguished by the absence of connectivity effects, which is, in turn, crucially linked to the presence of the D-pronoun lower than a sentence adverbial. This is illustrated by (14a), where the left dislocated

12. The other variants are those in which the D-pronoun is replaced by a personal one (i), and those in which the hanging topic XP appears in the nominative case, irrespectively from the case exhibited by the D/personal pronoun (ii):

(i)	*Den*	*/Der*	*Studenten,*	*der*	*Peter*	*hat*	*ihn*
	the-ACC	/ the-NOM	student	the-NOM	Peter	has	RP-ACC
	gestern	*gesehen*					
	yesterday	seen					

'The student, Peter saw him yesterday'

(ii)	*Der*	*Studenten,*	*den*	*hat*	*der*	*Peter*	*gestern*	*gesehen*
	the-NOM	student	RP-ACC	has	the-NOM	Peter	yesterday	seen

'The student, Peter saw him yesterday'

XP is bound by the subject, and (14b), where the hanging topic cannot be interpreted as bounded (examples taken from Frey, 2004a):

(14) a. *Seinem*$_i$ *Doktorvater,* *würde* ***dem*** ***wenigstens*** *ein* *Linguist*$_i$
his-DAT supervisor would RP-DAT at least one-NOM linguist
Geld *ausleihen?*
money-ACC lend
'His supervisor, would at least one linguist lend (him) money?'

b. ??*Seinem*$_i$ *Doktorvater,* *würde* ***wenigstens*** *ein* *Linguist*$_i$ ***dem***
his-DAT supervisor would at least one-NOM linguist RP-DAT
Geld *ausleihen?*
money-ACC lend
'His supervisor, would at least one linguist lend (him) money?'

Recall that in Frey's previous work (Frey, 2000, 2004b) a connection is established between topicality in the 'aboutness' sense and the position immediately preceding sentence adverbials in German (for evidence, see footnote 2). Thus, Frey's conclusion that the D-pronoun is the element that bears the [+topic] feature seems to be well-founded.

As for Frey's second conclusion, namely that the left dislocated XP is base-generated in the position it appears in, it seems to find some support in two different sets of data. The first is constituted by structures in which a left-dislocated XP co-occurs with a *wh*-element. As noticed by Grohmann (1997, 2000), the result is grammatical only if the D-pronoun stays in a lower position ((15a) vs (15b)). In this respect, notice that standard topicalization is incompatible with *wh*-movement (15c):

(15) a. *Den* *Studenten,* *wer* *hat* *den* *gestern* *gesehen?*
the-ACC student who-NOM has RP-ACC yesterday seen
'The student, who saw him yesterday?'

b. **Den* *Studenten,* *den* *hat* *wer* *gestern* *gesehen?*
the-ACC student RP-ACC has who-NOM yesterday seen
'The student, who saw him yesterday?'

c. **Den* *Studenten,* *wer* *hat* *gestern* *gesehen?*
the-ACC student who-NOM has yesterday seen
*'The student, who saw yesterday?'

The second set of data corresponds to instances of extraction out of a transparent embedded clause. In this case, the left dislocated constituent obligatorily precedes the complementizer, which is traditionally considered to occupy the highest head position within the CP-projection (that is, Rizzi's (1997) Force)[13] (Grohmann, 1997, 2000; Grewendorf, 2002):

(16) a. *Ich glaube, den Studenten, daß der Peter den*
I-NOM think the-ACC student that the-NOM Peter RP-ACC
gestern gesehen hat
yesterday seen has
'I think that Peter saw the student yesterday'

b. **Ich glaube, daß den Studenten der Peter den*
I-NOM think that the-ACC student the-NOM Peter RP-ACC
gestern gesehen hat
yesterday seen has
'I think that Peter saw the student yesterday'

In view of all this, the analysis to be developed in the following will keep Frey's proposal with respect to the two conflicting issues discussed in this section. However, it will crucially differ from it in the theoretical framework adopted, an issue to which we turn now.

3. German left dislocation and 'Survive Minimalism'

The main tenet of 'Survive Minimalism' (Stroik, 2009; Putnam 2007; Putnam & Stroik, in progress; and references therein) is that syntactic relations are strictly local, that is, they are restricted to the combination ('External Merge') of two syntactic objects that will form a new, larger one. From this perspective, 'second-order locality relations', i.e., minimality relations that state the structural conditions making that combination possible, must be completely dispensed with. The consequence is the dismissal of syntactic operations such as 'Internal Merge' ('Move'), 'Attract' and 'Agree' as well as the economy conditions that constrain them (the 'Minimal Link Condition', the 'Phase Impenetrability Condition', etc.). 'Narrow Syntax' is thus reduced to 'External Merge'.

13. In order to account for the grammatical (16a), Grewendorf argues that the complementizer is, in fact, the head of Fin, which allows the higher TopP to host the left dislocated constituent. Embedded TopP is, according to Grewendorf, linked to the non-defective CPs selected for by bridge-verbs. However, it is a well-known fact that standard topics cannot appear in the pre-complementizer position in West Germanic independently from the nature of the selecting verb.

The trigger for 'External Merge' is, as in orthodox Minimalism, feature checking. But, since in 'Survive Minimalism' the existence of uninterpretable features is rejected, feature checking is not driven by interface incompatibility, but rather by what Stroik (2009) calls concatenative integrity, that is, the well-formedness of the form and meaning derived from the inter-relationships of all the features of all the lexical items in the derivation.

As for 'Internal Merge', on which Chomsky (1995 and subsequent work) capitalizes for accounting for the pervasive 'Dislocation Property' of language, it is simply replaced by 'Remerge', i.e., iterated 'External Merge'. 'Remerge' is, in turn, a consequence of the 'Survive Principle' (Putnam, 2007:6):

(17) The Revised 'Survive Principle' (based on Stroik, 1999: 286)
If Y is a syntactic object (SO) in an XP headed by X, and Y has an unchecked feature [+F] that is incompatible with the feature X, Y remains active in the Numeration.

The disappearance of uninterpretable features along with the dependence of 'Remerge' on the 'Survive Principle' results in a theoretical model in which syntactic operations are necessarily correlated with semantic/pragmatic effects. Notice what this entails for an analysis of GLD structures such as Frey (2004a), which we have partially assumed: (i) general movement of the D-pronoun to the middle-field TopP is 'Remerge' for concatenative interpretation of the interpretable [+Top] feature on both the D-pronoun and Top; (ii) further movement of the D-pronoun to the CP-domain cannot be triggered by an EPP feature on the relevant head, but rather by the semantic/pragmatic feature responsible for keeping the D-pronoun active in the Numeration; and (iii) since, as far the left dislocated XP and the D-pronoun are concerned, there are no semantic/pragmatic interpretive differences between those structures in which the D-pronoun immediately precedes the finite verb and those in which it follows it, the same concatenation processes ('Merge' and 'Remerge' operations) must be involved in them. I will not add anything else as far as (i) is concerned, since it is, in my opinion, unproblematic. The issues in (ii)–(iii) will be addressed in the following sections.

3.1 A semantic/pragmatic trigger for the 'Remerge' of D-pronouns

This present section is devoted to showing that, contrary to what is assumed in current analyses of GLD, German D-pronouns must undergo 'Remerge' in the CP-domain for interpretive reasons. The main argument will be that the notion of syntactic referentiality (Stroik, 2000, 2009; Putnam, 2007), which will be introduced in Section 3.1.1, provides a straightforward explanation for some additional properties of German D-pronouns (Section 3.1.2).

3.1.1 *[+Ref] as a syntactic feature*

Stroik (2000, 2009, on previous work) and Putnam (2007) have defended in different studies the existence of a [+Ref] feature as a concatenation feature, that is, a feature that must be structurally licensed. Although, as shown by Putnam (2007), [+Ref] seems to find solid empirical support in several structures attested in an array of languages,[14] the discussion that follows will focus on Stroik's analysis of multiple *wh*-movement in English (Stroik, 2000, 2009).

Stroik argues that English *wh*-elements pattern with pronouns in that they show up in three different varieties, according to the chart in (18):

(18)

Type	**Pronominal elements**	***Wh*-element**
R-expressions	nominal DP	*wh*-operator
	He saw the bastard	*What did he see?*
Structure-dependent elements	anaphors	paired *wh*-in situ elements
	He saw himself	*Who read what?*
Discourse-linked elements	pronouns	Echoic *wh*-in situ elements
	He saw her	*He saw what?*

According to Stroik, the distinguishing property of R-expressions is that they may take a discourse-free antecedent (or referent). This is not so with structure-dependent elements, whose referentiality is crucially dependent on the referentiality of another element in the sentence. Finally, as is evident from their label, discourse-linked elements are necessarily associated with antecedents (or referents) from the backgrounded discourse.

Stroik analyzes in detail *wh*-in situ elements in English, with special attention to their well-studied semantic/pragmatic properties as well as their syntactic distribution. On these grounds, he concludes that the tripartite classification of *wh*-elements in (18) finds its correlate in the three different types of features associated to them in English: [+OP(erator)], [+REF(erential)], and [DISC(ourse)] features. In strictly syntactic terms, a [+OP] feature is a concatenation feature, which means that it must be structurally licensed through 'Remerge' of the constituent that bears the feature [+OP] with C, as illustrated in (19h):

14. According to Putnam, axiomatic negative predicates in Xhosa, DPs in Old Norse, and scrambled phrases in West Germanic.

(19) *Who(m) did John see?*
   a. Merge ⟨see, whom⟩ -⟩ see whom
   b. Merge ⟨v, ⟨see, whom⟩⟩ -⟩ v see whom
   c. Merge ⟨John, ⟨v, ⟨see, whom⟩⟩⟩ -⟩ John see whom
   d. Merge ⟨T, ⟨John, ⟨v, ⟨see, whom⟩⟩⟩⟩ -⟩ T John see whom
   e. Merge ⟨did, ⟨T, ⟨John, ⟨v, ⟨see, whom⟩⟩⟩⟩⟩ -⟩ did John see whom
   f. Remerge ⟨John, ⟨did, ⟨T, ⟨John, ⟨v, ⟨see, whom⟩⟩⟩⟩⟩⟩ -⟩ John did John see whom
   g. Merge ⟨C, ⟨did, ⟨T, ⟨John, ⟨v, ⟨see, whom⟩⟩⟩⟩⟩⟩⟩ -⟩ C John did see whom
   h. Remerge **⟨whom, ⟨C**, ⟨did, ⟨T, ⟨John, ⟨v, ⟨see,whom⟩⟩⟩⟩⟩⟩⟩⟩ -⟩ whom John did John see whom

[+REF] features may appear alone, or clustered with a [+OP] feature. These are the features that characterize multiple *wh*-constructions receiving a pair-list reading, where the moved *wh*-phrase is [+OP, +REF] and the one that appears in situ is [+REF]. Stroik provides empirical support for his claim that the [+REF] features of the in-situ element are licensed only by their concatenation with those of the *wh*-operator, in a structural configuration like the one in (20h):

(20) *Who saw whom?*
   a. Merge ⟨saw, whom⟩ -⟩ saw whom
   b. Merge ⟨v, ⟨saw, whom⟩⟩ -⟩ v saw whom
   c. Merge ⟨who, ⟨v, ⟨saw, whom⟩⟩⟩ -⟩ who saw whom
   d. Merge ⟨T, ⟨who, ⟨v, ⟨saw, whom⟩⟩⟩⟩ -⟩ T who saw whom
   e. Remerge ⟨who, ⟨T, ⟨who, ⟨v, ⟨saw, whom⟩⟩⟩⟩⟩ -⟩ who T who saw whom
   f. Merge ⟨C, ⟨who, ⟨T, ⟨who, ⟨v, ⟨saw, whom⟩⟩⟩⟩⟩⟩ -⟩ C who T who saw whom
   g. Remerge ⟨whom, ⟨C, ⟨T, ⟨who, ⟨v, ⟨see, whom⟩⟩⟩⟩⟩⟩⟩ -⟩ whom C who T who saw whom
   h. Remerge **⟨who, ⟨whom, ⟨C,** ⟨T, ⟨who, ⟨v, ⟨see, whom⟩⟩⟩⟩⟩⟩⟩⟩ -⟩ who whom C who T who saw whom

Notice that the [+OP, +REF] *who* c-commands the referentially dependent *whom*, which results in the appropriate concatenative configuration. Stroik claims that this fact receives a straightforward explanation if the part of the Numeration where elements for 'Remerge' are stored is structured vertically, in a top-down structure that places new elements at the bottom: in other words, order of 'Remerge' must strictly follow the order in which (first) 'Merge' took place. Notice also that, contrary to what is standard in orthodox Minimalism, the referentially dependent object *whom* is not phonologically realized in its 'Remerge' position, but rather in the position in which its theta-role and Case are licensed. Stroik just assumes that, in contrast to the case of [+OP] features, [+REF] is never realized morphophonetically in English

Finally, Stroik assigns a [+DISC] feature to echoic *wh*-elements. Since the syntactic distribution of such *wh*-in situ elements is much less constrained than that of their counterparts with pair-list interpretations, Stroik concludes that [+DISC] is not a concatenation feature, which entails its inability to trigger syntactic 'Remerge':

(21) *John saw whom?*
   a. Merge ⟨saw, whom⟩ -⟩ saw whom
   b. Merge ⟨v, ⟨saw, whom⟩⟩ -⟩ v saw whom
   c. Merge ⟨John, ⟨v, ⟨saw, whom⟩⟩⟩ -⟩ John saw whom
   d. Merge ⟨T, ⟨John, ⟨v, ⟨saw, whom⟩⟩⟩⟩ -⟩ T John saw whom
   e. Remerge ⟨John, ⟨T, ⟨John, ⟨v, ⟨saw, **whom**⟩⟩⟩⟩⟩ -⟩ John T John saw whom

As it stands, Stroik's analysis of *wh*-structures in English makes two interesting contributions: (i) referentiality correlates with an interpretable feature on the relevant lexical item; and (ii) referentiality may be licensed either syntax-internally, thus triggering a structural operation that concatenates two [+REF] features, or syntax-externally, namely at the Discourse level. The account of GLD constructions that will be offered in the following also capitalizes on these two assumptions.

3.1.2 *The referentiality of D-pronouns*

The D-pronouns that appear in GLD structures are traditionally treated as belonging to one of the demonstratives series attested in the grammar of German. This series, labeled 'simple' in Cardinaletti & Starke (1999), differs from those demonstratives they call 'intensified' in that it does not make spatial distinctions of the near/far type. As noticed by Wiltschko (1998), Cardinaletti and Starke (1999), Déchaine and Wiltschko (2002) and Grewendorf (2002), among others, both demonstrative series differ, in turn, from personal pronouns in being subject to Principle C of the Binding Theory, and not to Principle B. This observation is illustrated by the examples in (22), taken from Grewendorf (2002):

(22) a. *Er*$_i$ *glaubt, keiner habe ihn*$_i$ *gelobt*
   he-NOM thinks nobody hat him-ACC praised
   'He thinks that nobody praised him'

   b. **Er*$_i$ *glaubt, keiner habe den*$_i$ *gelobt*
   he-NOM thinks nobody has D-ACC praised
   'He thinks that nobody praised him'

   c. *Maria hat ihm*$_i$ *gesagt, ihn*$_i$ *habe keiner gelobt*
   Maria-NOM has him-DAT said him-ACC has nobody-NOM praised
   'Maria said to him that nobody praised him'

   d. **Maria hat ihm*$_i$ *gesagt, den*$_i$ *habe keiner gelobt*
   Maria-NOM has him-DAT said D-ACC has nobody-NOM praised
   'Maria said to him that nobody praised him'

In his study of West Germanic D-pronouns, mostly based on Dutch data, Hoekstra (1999) explains the facts in (22) as due to the requirement that D-pronouns move to the beginning of the sentence that contains them. Hoekstra establishes a parallelism between this kind of movement and the movement Romance clitics undergo to the functional projection that hosts them. In Hoekstra's view, the only difference between the two processes is that, whereas the former takes place covertly, the latter does overtly.

A different position is taken in Wiltschko (1998) and Déchaine and Wiltschko (2002), who defend the idea that the notion of pronoun is not a primitive of linguistic theory, and explain the data above by claiming that German D-pronouns are pro-DPs, as the structure for German nominative *der* in (23) shows. Their basic argument is that DPs are the only syntactic constituents with the binding theoretic status of R-expressions. In this light, D-pronouns would differ from other DPs only in the presence of an empty NP:

(23) $_{DP}[\ _{D'}[\text{d-}\ _{\varphi P}[\ _{\varphi P'}[\text{er}\ _{NP}[\ \emptyset\ ]]]]]$[15]

I would like to argue here that, despite their apparent incompatibility, both approaches may be reconciled if we assume that, although the DP-layer is responsible for the observed Principle C effects, by itself it can only license referentiality understood in a broad sense.[16] But the kind of referentiality a D-pronoun encodes is more restricted, as shown by the fact that it necessarily requires an antecedent within what Hoekstra (1999) calls 'the structural level of the text':[17]

(24) a.

Kennst	*du*	*Peter?*	*Den*	*kenne*	*ich.*
know-2SG.PRES.IND	you-NOM	Peter-ACC	RP-ACC	know	I-NOM

'Do you know Peter? I know him'

b.

#*Kennst*	*du*	*den?*	*Den*	*kenne*	*ich*
know-2SG.PRES.IND	you-NOM	RP-ACC	RP-ACC	know	I-NOM

'Do you know him? I know him'

15. Wiltschko and Déchaine and Wiltschko propose φP as an intermediate Agreement projection to explain the asymmetry between those languages in which the empty NP is possible (for example, German) and those in which it is not (for example, English).

16. Along the lines of, for example, the definition in Putnam (2007), based on Fodor and Sag (1982):

(i) A nominal expression is understood to be referential if it has a 'fixed referent' in the (model of the) world, meaning that it can be identified by the speaker and/or by one of the people whose propositional attitudes are being reported.

17. According to Hoekstra, 'the structural level of the text' is the coherent and structured whole where matrix sentences connect up with each other. Given that most well-known syntactic processes are clause-bound, operations different from 'Merge' and 'Move' must be responsible for the creation of textual structure.

Furthermore, as expected, that antecedent must itself be referential (Grohmann, 1997; Hoekstra, 1999; Grewendorf, 2002):

(25)	*Kennst*	*du*	*jemanden?*	**Den*	*kenne*	*ich*
	know-2SG.PRES.IND	you-NOM	someone-ACC	RP-ACC	know	I-NOM

On the other hand, recall that, in Wiltschko's and Déchaine and Wiltschko's analysis, the only structural difference between a standard R-expression and D-pronouns lies on the empty nominal that appears in the latter. According to Uriagereka (1995), a similar distinction is found between nominal DPs and clitics in Western Romance. Uriagereka assigns the latter the structure in (26), where *pro* is associated to the relevant notion of referentiality, which, in turn, triggers movement to an appropriate functional projection:

(26) $_{DP}[\ _{D'}[\text{clitic}\ _{NP}[\ \textit{pro}]]$

Due to the complex issues involved in a proper treatment of *pro* within the framework of 'Survive Minimalism', I will not discuss whether the structure in (26) should replace (23) in the case of German D-pronouns, leaving the issue for further research. However, as in Uriagereka's approach, I will assume that the notion of referentiality at stake is encoded in this "defective" nominal, and forces 'Remerge' of the entire DP with a high functional head of the German pre-field, namely C.

3.2 A unified treatment of GLD structures

As we saw in Section 2, there are two types of German GLD structures: (i) those in which the D-pronoun immediately precedes the finite verb (27a), and (ii) those in which the D-pronoun follows the finite verb but, crucially, must precede a sentence adverbial (27b). Recall also that, although one of the advantages of Frey's (2004a) analysis was that it provided a uniform account for both, it did so at the cost of resorting to the theoretically problematic EPP feature.

(27)	a.	*Den*	*Studenten,*	***den***	*hat*	*der*	*Peter*	*zum Glück*	*gesehen*
		the-ACC	student	RP-ACC	has	the-NOM	Peter	fortunately	seen
		'The student, fortunately Peter saw (him)'							
	b.	*Den*	*Studenten,*	*der*	*Peter*	*hat*	***den***	*zum Glück*	*gesehen*
		the-ACC	student	the-NOM	Peter	has	RP-ACC	fortunately	seen
		'The student, fortunately Peter saw (him)'							

On the other hand, we have assumed that D-pronouns are endowed not only with the [+Topic] feature Frey correctly assigns to them, but also with a referential feature that forces them to find an antecedent in what Hoekstra's (1999) calls the 'structural level of the text'. Notice what this entails in 'Survive' terms: if the referential feature is a concatenative feature that requires structural licensing, i.e., [+REF], the

D-pronoun will remain active in the Numeration until it undergoes 'Remerge' with an appropriate, [+REF], head. But if the referential feature is non-concatenative, i.e., [+DISC], it will be interpreted at the conceptual-intentional interface but exempted from syntactic checking, which goes along with its complete disappearance from the Numeration right after the licensing of its [+Topic] feature. The proposal here is that German grammar, in fact, allows for both possibilities. The latter appears in so-called Hanging Left Dislocation structures, where, in contrast to GLD structures, the left dislocated constituent may appear in the nominative case. As may be observed in (28), the presence of the nominative case correlates with the absence of both island constraints (28a) and connectivity effects (28b) (examples from Grohmann, 2000):

(28) a. *Der schöne Mann, den haßt Martin die Tatsache daß*
the-ACC handsome man RP-ACC hates Martin the fact that
die kluge Frau geküßt hat
the-NOM smart woman kissed hat
'The handsome man, Martin hates the fact that the smart woman kissed him'

b. **Der Wagen von sich$_i$, den hat er$_i$ verkauft*
the-ACC car of himself RP-ACC has he-NOM sold
'The car of himself, he sold it'

Note that notions such as islandhood or reconstruction are inextricably linked to the structural dependencies Narrow Syntax creates. If, as argued by Stroik (2000, 2009), [+DISC] features are licensed in a different component, we expect this kind of licensing to be blind to such structural requirements.

We find exactly the opposite situation in GLD structures, which exhibit island sensitivity as well as connectivity effects (see the examples in (8)). The conclusion is, as in Stroik's (2000, 2009) analysis of *wh*-in situ with a pair-list interpretation, that their more limited distribution is due to the requirement that the relevant feature be concatenatively, i.e., structurally, licensed. I will assume that the concatenation at stake is the one between the [+REF] features that appear on C, the left dislocated constituent, and the D-pronoun, which, we recall, is referentially dependent on the left dislocated XP. The relevant part of the derivations of the two German GLD structures in (27a) and (27b) are given in (29a) and (29b):

(29) a. *Den Studenten, den hat der Peter gesehen*
the-ACC student RP-ACC has the-ACC Peter seen
'The student, fortunately Peter saw (him)'
Merge ⟨den Studenten, ⟨ den ⟨**C** ⟨<u>den</u>, ⟨hat, ⟨Top

b. *Den Studenten, der Peter hat den gesehen*
the-ACC student the-NOM Peter has RP-ACC seen
'The student, fortunately Peter saw (him)'
Merge ⟨den Studenten, ⟨ den ⟨**C** ⟨der Peter ⟨<u>den</u>, ⟨hat, ⟨Top

If this account is on the right track, the number and the type of the concatenative operations which the left dislocated XP and the D-pronoun undergo are exactly the same in both structures. Thus, it is not surprising that they obey the same structural requirements and that, as far as the two relevant elements are concerned, they receive a uniform interpretation at LF.

There are, however, two prominent problems in our analysis in (29). The first one is that it cannot accommodate the verb-second constraint. Recall that current analyses of GLD such as Frey (2004a) derive V2-effects from the requirement that the finite verb move to the higher head of the C-domain endowed with an EPP feature. This is clearly incompatible with a 'Survive' approach, where 'Remerge' is triggered exclusively on interpretive grounds. Therefore, we are left with no explanation for the concatenative operation that remerges the finite verb with the [+TOP] head in (29a), and neither can we account for the presence of the finite verb between the two topical elements in (29b).

The second problem is theoretical, and relates to the PF component. We have assumed, following Stroik's (2000, 2009) analysis of English *wh*-structures, that concatenated [+REF] features do not receive morphophonetic interpretation in German either, which would fit the data in both (29a) and (29b): as shown by underlining, the D-pronoun is pronounced in a hierarchically lower position, where its [+TOP] feature is semantically and pragmatically interpreted. In this respect, I would like to draw the reader's attention to Stroik's account of the English example below:

(30) a. What did Sam fix?
b. Who did Sam tell how to fix what?

According to Stroik's model, the different positions the *wh*-element *what* ocuppies in (30a) and (30b) derive from its characterization as [+OP, +WH] in (37a) and [+REF, +WH] in (37b). In the former case, as Stroik puts it, "since the only overt morphophonetic feature of the DP is its [WH] feature, the DP will be spelled out where this feature has been checked in the Spec, CP position" (2007: 89). In the latter, however, he states that "the *wh*-in-situ element is interpreted morphophonetically within the embedded *v*P or VP (depending on where its Case feature is checked)" (2007: 79). Now the question is how to prevent checking of the morphophonetic feature associated to [+Case] (30b) if the *wh*-element undergoes further 'Remerge' to Spec, C (30a). Perhaps, one may assume that checking of morphophonetic features is generally incompatible with 'surviving', unchecked features: thus, the presence of the unchecked [+OP] feature on the *wh*-element in (30a) would block checking of the morphophonetic features associated to its [+Case] feature.[18] Note, however, that

18. The null hypothesis is that a lexical item undergoes 'Merge' and 'Remerge' bearing all its features, and all its features must be visible to the head with which the lexical item merges or remerges.

such an assumption makes an incorrect prediction for (30b): the *wh*-element is also endowed with an additional, unchecked feature ([+REF]), and, nevertheless, the morphophonetic features associated to [+Case] are checked in that position. Thus, if this reasoning is correct, the grammar must be endowed with the kind of 'look ahead' properties 'Survive Minimalism' tries to dispense with.

The following sections are devoted to explaining how the problems just discussed may be circumvented.

3.2.1 *GLD and V2*

As shown in the introduction, some current approaches to V2 treat the phenomenon as a side-effect of the fronting operation that places an XP in the sentence initial position. Such a treatment is problematic in so far as the connection between the movement of the finite verb and the movement of the XP is defined in strictly syntactic terms: verb movement is not triggered by any feature on the verb itself, but rather by a feature on the also displaced constituent. However, this connection must be defined otherwise, as Zwart (2005) demonstrates.

Zwart hypothesizes that 'Merge', the only structure generating procedure he assumes, invariably results in the creation of a dependency relation between an element *x* (the antecedent or non-dependent) and an element *y* (the dependent). Moreover, he claims that such a dependency may be morphologically or positionally marked on one of the members of the dependent: in the former case, morphological marking, the marked element must spell out morphological agreement with the non-dependent; in the latter case, positional marking, the marked element—a linker—must be spelled out as the left-most member of the dependent. Note that, from this perspective, V2 reduces to the application of the positional spell-out rule of dependent marking to the finite verb.

According to Zwart, positional marking is diversely constrained. First, it applies only in some languages (for example, V2 languages vs V1 languages), and only in some constructions (for example, English *wh*-questions vs English declarative sentences). Second, it seems to be restricted, at least in West Germanic, to dependencies that mark the end of a cycle, where cycle is defined as in (31) (Zwart, 2005: 13):

(31) Cycle
A cycle is constituted: (a) when no further operation Merge takes place, or
(b) when the non-dependent is a lexical item (i.e., a noun, a verb, or adjective), or
(c) in elsewhere cases

While case (a) is, according to Zwart, the case of West Germanic root clauses, case (b) applies to embedded clauses, where the complementizer, and not the finite verb, is the linker. Finally, case (c) is needed for language and construction specific variation.

Zwart notices that West Germanic languages allow for structures that deviate from the V2 pattern, and tries to accommodate them within his system. One of them is GLD, which, as its Dutch counterpart, is V3 in its two variants, given that the finite verb is preceded by the left dislocated constituent and the D-pronoun (*Den Studenten, den hat der Peter gesehen*) or the left dislocated constituent and the subject (*Den Studenten, der Peter hat den gesehen*). Zwart accounts for it as a case in which the dependent, not a cycle in the general case, functions as such (32). Note that, since positional marking applies within the cycle, it is completely blind to the extra-cyclic fronted XP:

(32) a. *Den Studenten, den hat der Peter gesehen*
the-ACC student RP-ACC has the-NOM Peter seen
'The student, Peter saw (him)'
$_{\text{CYCLE-1}}$[Den Studenten] $_{\text{CYCLE-2}}$[den hat der Peter gesehen]

b. *Den Studenten, der Peter hat den gesehen*
the-ACC student the-NOM Peter has RP-ACC seen
'The student, Peter saw (him)'
$_{\text{CYCLE-1}}$[Den Studenten] $_{\text{CYCLE-2}}$[der Peter hat den gesehen]

In Zwart's view, the presence of extra-dependent elements is not restricted to initial positions. In this regard, he discusses the case of Dutch sentence connecting adverbs like *echter* 'however', *nu* '(non-temporal) now', *dan* '(non-temporal) then', *immers* 'as is known', and *daaregen* 'in contrast', which may intervene between the sentence initial constituent and the finite verb in Dutch V2 structures (examples from Zwart, 2005):

(33) *Dit voorstel* ***echter*** *is onacceptabel*
this proposal however is unacceptable
'This proposal is, however, unacceptable'

Zwart argues that this kind of extra-dependency finds its justification in the prosodic properties of the extra-dependent element, which being unstressed, must necessarily group with the first constituent, as evidenced in (34), a LD construction:

(34) a. *Dit voorstel* ***echter*** *dat is onacceptabel*
this proposal-NOM however RP-NOM is unacceptable
'This proposal, however, is unacceptable'

b. *Dit voorstel dat* ***echter*** *is onacceptabel*
this proposal-NOM RP-NOM however is unacceptable
'This proposal, however, is unacceptable'

However, he also suggests that the facts in (34) may be alternatively explained by resorting to the character of the positional marking rule as a spell-out rule, which would mark V2 as a process sensitive to prosodic grouping. At this point, I would like to make a stronger claim and assume that, as a spell-out rule, V2 must be also sensitive to other phonological processes, as well as to the mere presence of phonological features on the fronted constituent.[19]

In the remainder of the section, I will show how an analysis of V2 along the lines proposed above is mostly compatible with our 'Survive' account of GLD. I say 'mostly' because, as the reader has probably noticed, the rule of positional marking Zwart defends is crucially dependent on the notion of cycle, a theoretical construct 'Survive Minimalism' dispenses with. Notice, however, that Zwart's cycle differs from Chomsky's (2001, and subsequent work) in important respects. Here I will just note that, while Chomsky's cycle is based on the notion of PF and LF integrity, Zwart's is based instead on the only strictly syntactic operation he assumes, i.e., Merge,[20] which may make it compatible with 'Survive Minimalism'. In any case, despite its potential importance, I must leave the issue for further research.

On these grounds, full derivations for the sentences in (32) can be given as in (35):[21,22]

(35) a. *Den Studenten, den hat der Peter gesehen*
the-ACC student RP-ACC has the-NOM Peter seen
'The student, Peter saw (him)'
$_{\text{CYCLE-1}}$[den Studenten] $_{\text{CYCLE-2 CP}}$[~~den~~ [C $_{\text{TopP}}$[den [Top $_{\text{TP}}$[der Peter $_{\text{AuxP}}$[$_{v\text{P}}$[der Peter $_{\text{VP}}$[~~den gesehen~~] gesehen] ~~hat~~] hat]]]]

19. *Contra* Zwart's account of Dutch V1, which he takes as the spell-out positional marking of the dependency between an empty operator (the non-dependent) and the clause itself (the dependent).

20. At least in the case of root clauses, the type of structure on which most of our analysis of GLD structures focuses.

21. Following the traditional analysis of German clause structure, I assume that TP, AuxP and *v*P are head-final projections, and that finite lexical Vs undergo 'Remerge' with *v* and T. Auxiliaries are base-generated in their own projection, and remerged with T if they are finite. With respect to TP, although Zwart (2005) rejects the idea that it heads its own functional projection and treats it as an operator, I will follow the conventional view.

22. Non-realization at PF is marked by crossing.

b. *Den Studenten, der Peter hat den gesehen*
the-ACC student the-NOM Peter has RP-ACC seen
'The student, Peter saw (him)'
$_{\text{CYCLE-1}}$[den Studenten] $_{\text{CYCLE-2 CP}}$[~~den~~ [C $_{\text{TopP}}$[der Peter [den[Top $_{\text{TP}}$[~~der Peter~~ $_{\text{AuxP}}$[$_{vP}$[der Peter $_{\text{VP}}$[~~den gesehen~~] gesehen] ~~hat~~] hat]]]]

If this analysis is on the right track, it may be concluded that the position of the finite verb in (35) is the product of the combination of three different factors: (i) the extra-dependent nature of the left dislocated constituent; (ii) the lack of phonological features on the D-pronoun (shown by crossing) that goes along with the checking of its [+REF] feature; and (iii) the application of Zwart's spell-out rule for left dependent marking. In other words, the finite verb must be spelled out immediately after the first visible non-dependent, that is, the topic *den* in (35a), and the topic *der Peter* in (35b). The issue of how this visibility must be derived is the focus of the following section.

3.2.2 *GLD in embedded V2 clauses*

As seen in Section 3.2, the assumption that, following Stroik, we have adopted so far with respect to the non-realization of phonological features in [+REF] concatenations faces some difficulties in the case of English multiple *wh*-constructions. On the other hand, if the analysis of GLD proposed here is correct, it does not make the right predictions for this structure either. The main problem comes from embedded left dislocation in German, which, in the general case, is compatible only with V2 structures.[23] Although a discussion of the phenomenon of embedded V2 in German lies outside the scope of this paper, I should point out here that it must

23. As shown in Section 2.3, left dislocation is also acceptable with an embedded *daß*-clause if the matrix verb is a bridge verb (i), and the left dislocated XP is outside the subordinate clause (ii) (examples from Grohmann, 2000):

(i) (?)*Der Bauer glaubt, diesen Frosch, daß sie den gestern*
the-NOM farmer thinks this-ACC frog that she RP-ACC yesterday
geküßt hat
kissed has
'The farmer believes that this frog she kissed yesterday'

(ii) **Der Bauer glaubt, daß diesen Frosch, den sie gestern geküßt hat*
the-NOM farmer thinks that this-ACC frog that RP-ACC yesterday kissed has
'The farmer believes that this frog, she kissed yesterday'

In this respect, notice that the ungrammaticality of (ii) squares nicely with Zwart's treatment of the left dislocated XP as extra-dependent.

be strictly distinguished from the embedded V2 found in Scandinavian languages: the presence of a complementizer is completely forbidden in the former, while required in the latter. Therefore, I will follow the conventional view and syntactically treat German embedded V2 clauses as root clauses.

As pointed out by Grewendorf (2002), a striking fact about left dislocation with embedded V2 is that the left-dislocated constituent seems to be able to satisfy the verb-second constraint in the highest and intermediate clauses (36a), but not in the lowest one (36b):

(36) a. *Den Studenten glaubt Hans meinte Maria den hat*
the-ACC student believes Hans thinks Maria-NOM RP-ACC has
der Peter gesehen
the-NOM Peter seen
'The student, Hans believes that Maria thinks that Peter saw him'

b. **Den Studenten glaubt Hans meinte Maria hat*
the-ACC student believes Hans thinks Maria-NOM has
der Peter den gesehen
the-NOM Peter RP-ACC seen
'The student, Hans believes that Maria thinks that Peter saw him'

The reader will also recall that failure of the left dislocated constituent in complying with V2 is likewise observed in simple root clauses, which seriously defied both Grewendorf's (2002) and Frey's (2004a) accounts of GLD:

(37) **Den Studenten hat der Peter den gesehen*
the-ACC student has the-NOM Peter RP-ACC seen
'The student, Peter saw him'

(36a) is problematic for the proposal the present paper defends: on the one hand, we have followed Zwart in assuming that the left dislocated XP is extra-cyclic, hence invisible to the spell-out rule for left-dependent marking (V2); on the other, if we contend, as we do, that it is the 'Remerge' of the D-pronoun with the successive functional heads that satisfies V2, we are faced with the contradiction that an element checking a [+REF] feature and thus lacking phonological realization is visible for the spell-out positional rule. Note that, in principle, this could be a possibility in the case of the intermediate Cs, on the assumption that they do not themselves bear a compatible [+REF] feature. But this would be empirically wrong: the silent D-pronoun cannot satisfy V2 in the supposedly incompatible lowest C. Moreover, a solution of this sort must be absolutely excluded in the case of the highest C, which must necessarily bear the [+REF] feature that serves to establish the required concatenative relation between the [+REF] on the D-pronoun and the [+REF] feature on the left dislocated XP.

3.2.2.1 *Some previous considerations* I submit here a tentative proposal that modifies Stroik's design of the PF component in important ways. Recall that such a design is based on the assumption that morphophonetic features are checked in the position in which the interpretive feature they are associated with is checked. Recall also that, as we discussed in Section 3.2, Stroik's model does not explain why, whereas a *wh*-operator is pronounced in Spec, CP, a referential *wh*-element receives phonological realization in the position linked to the checking of its [+CASE] feature. In a sense, it is as if Narrow Syntax knew that only the operator, and not the referential *wh*-element, will have another chance to have its morphophonetic features checked. Therefore, the syntactic component is endowed with a 'look ahead' property 'Survive Minimalism' rejects.

My proposal assumes, as in Stroik (2009), that the 'Copy Theory of Movement' must be dispensed with, for the obvious reason that, once 'Internal Merge', i.e., movement, is not a syntactic operation anymore, no copies of moved lexical items are needed. However, it will depart from 'Survive' tenets in two relevant aspects: (i) as shown in the brief discussion above (see also Section 3.2), Stroik's claim that checking of morphosyntactic features takes place in Narrow Syntax faces some empirical difficulties; therefore, I will stick to the conventional view and claim that all merged and remerged constituents are visible to PF,[24] which, in turn, entails the need for an additional mechanism that may determine which instance of 'Remerge' must be pronounced. As in orthodox Minimalism, my claim is that morphophonological realization is the product of 'Chain Reduction',[25] and 'Formal Feature Elimination', as they are defined in Nunes (2004).[26] A potential

24. I abstract away from the issue of 'automatic' or 'blind' Remerge', i.e., 'Remerge' independent from feature compatibility.

25. 'Chain Reduction' is completely unnecessary for Stroik, given his claim that checking of morphophonetic features takes place in Narrow Syntax. But chain formation as such is still needed in both the interpretive and sensorimotor components. While LF chains are typically nontrivial, PF chains consist, in the general case, of only one link, given that morphophonetic features are checked only once.

26. (i) Chain Reduction
Delete the minimal number of constituents of a nontrivial chain CH that suffices for CH to be mapped into a linear order in accordance with Kayne's (1994) Linear Correspondence Axiom.
(ii) Formal Feature Elimination (FF-Elimination)
Given the sequence of pairs $\sigma = \langle (P, F)_1, (F, P)_2, \ldots, (F, P)_n \rangle$ such that σ is the output of Linearize, F is a set of formal features, and P a set of phonological features, delete the minimal number of features of each set of formal features in order for σ to satisfy Full Interpretation at PF.

complication is that, whereas uninterpretable features are the focus in Nunes's system, the 'Survive' model adopted in the present analysis forces us to deal with interpretable features exclusively. In this respect, it must be noticed that Nunes's system may well be extended to interpretable features. In fact, he addresses the issue in passing (2002: 167, fn. 25), and suggests that interpretable features become invisible at PF after being checked. At first sight, this proposal does not fare better than Stroik's account of the contrast of the English *wh*-structures in (30), repeated here for convenience:[27]

(38) a. What did Sam fix?
$_{CP}$[what [did $_{TP}$[Sam $_{vP}$[what $_{[OP]}$ $_{VP}$[fix what $_{[CASE]}$ $_{[OP]}$]]]]

b. Who did Sam tell how to fix what?
$_{CP}$[who [what [did $_{TP}$[Sam $_{vP}$[$_{VP}$[tell you $_{CP}$[how $_{TP}$[PRO [to $_{vP}$[what $_{[REF]}$ $_{VP}$[fix what $_{[CASE]}$ $_{[REF]}$]]]]

As (38) illustrates, 'Chain Reduction' should result in the pronunciation of the highest link in both instances, thus making 'FF-Elimination' unnecessary. The conclusion is obviously wrong. My suggestion, at this point, is that, unlike the [+OP] feature in (38a), [+REF] in (38b) retains its phonological visibility even after having been remerged with C. In this connection, I will briefly address Stroik's (2009) subtle distinction between feature checking (feature deactivation) and feature licensing.

As extensively shown elsewhere, Stroik's analysis of English multiple *wh*-constructions with pair-list reading capitalizes on the idea that the *wh*-in situ element is referentially dependent on the *wh*-operator (that is, in (38b) above *what* is referentially dependent on *who*). Additionally, its [+REF] feature is compatible with the [+REF] feature on C, with which it undergoes 'Remerge'. On this basis, Stroik makes a distinction between 'Remerge' that results in feature checking (or feature deactivation), and 'Remerge' that results in feature licensing: although both require feature compatibility and prevent the feature at stake from undergoing further 'Remerge', only the latter has interpretive effects. Note that, for the case at hand, this simply means that the feature concatenation involving C, *who*, and *what* is interpreted only after the 'Remerge' of *who*.

On these grounds, the representation of (38b) should be as in (39):

(39) Who did Sam tell how to fix what?
$_{CP}$[who [what $_{[REF]}$ [did $_{TP}$[Sam $_{vP}$[$_{VP}$[tell you $_{CP}$[how $_{TP}$[PRO [to $_{vP}$[what $_{[REF]}$ $_{VP}$[fix what $_{[CASE]}$ $_{[REF]}$]]]]

27. Following Nunes (2002), boldfaced subscripts represent features visible at PF.

The question that immediately arises is why checked/deactivated but otherwise unlicensed features are not deleted at PF. Although I must admit that I do not have an answer at this moment, I would like to refer to Nunes's observation that only interpretable features that are not deleted at PF are allowed to enter into another checking relation (2002: 167, fn. 25). In our example in (39), notice that, after being remerged with the [+REF] feature on C, the [+REF] feature on *what* still must be licensed by the [+REF] feature on *who*, which would constitute an additional licensing relation. I leave for further research the issue whether this intuition, if correct, may be formalized somehow.

3.2.2.2 *GLD in embedded V2 clauses revisited* As the reader will recall, the main problem that GLD in embedded V2 clauses poses is illustrated by the set of examples in (36) and (37) above, repeated here for convenience as (40):

(40) a. *Den Studenten glaubt Hans meinte Maria den hat*
the-ACC student believes Hans thinks Maria-NOM RP-ACC has
der Peter gesehen
the-NOM Peter seen
'The student, Hans believes that Maria thinks that Peter saw him'

b. **Den Studenten glaubt Hans meinte Maria hat*
the-ACC student believes Hans thinks Maria-NOM has
der Peter den gesehen
the-NOM Peter RP-ACC seen
'The student, Hans believes that Maria thinks that Peter saw him'

c. **Den Studenten hat der Peter den gesehen*
the-ACC student has the-NOM Peter RP-ACC seen
'The student, Peter saw him'

In a nutshell, for an analysis of GLD such as the one we have presented so far, the question is why the silent D-pronoun may comply with the V2 constraint in the case of the matrix higher clauses in (40a), but it may not do so in its own clause, whatever its syntactic status, i.e., root (40c) or embedded (40b), may be.

As discussed in previous sections, it is clear that our proposal, which contends that 'Remerge' of the D-pronoun with C lacks phonological realization, fares well with both (40b) and (40c), given that, as we have argued, the left dislocated XP is an extra-dependent element, and V2 is phonologically constrained. But it is also clear that the extension of such an account to cases like the grammatical structure of (40a) gives the wrong results.

I would like to argue here that the data in (40a) may also receive a principled explanation if they are analyzed from the perspective we have adopted so far. Specifically, we take these data to be the product of the interaction of 'Remerge' of the

[+REF] feature on the D-pronoun with the [+REF] feature on C, 'Chain Reduction', and Zwart's (2005) cyclic rule of positional marking.

Our first assumption is, as in Zwart, that V2 applies within a cycle. From this perspective, the structure in (40a) would be constituted by three cycles plus the extra-dependent left dislocated XP:

(41) a. $_{\text{CYCLE-1}}$[*den hat der Peter gesehen*]
b. $_{\text{CYCLE-2}}$ [*meinte Maria*]
c. $_{\text{CYCLE-3}}$ [*glaubt Hans*]
d. $_{\text{CYCLE-4}}$ [*den Studenten*]

In the first cycle, the [+TOP] feature on the D-pronoun undergoes 'Remerge' with the [+TOP] on the head of Frey's (2000, 2004b) TopP, but, as required by the 'Survive Principle', remains in the Numeration for licensing of its [+REF] feature. Notice, however, that the lower-C is not endowed with a compatible feature, as we may derive from the comparison of (40a) with the also grammatical structure in (42):

(42)	*Hans*	*glaubt,*	*den*	*Studenten,*	*den*	*hat*	*Peter*	*gesehen*
	Hans-NOM	thinks	the-ACC	student	RP-ACC	has	Peter-NOM	seen

'The student, Hans thinks that Peter saw'

The D-pronoun must undergo further 'Remerge', first with the incompatible embedded C, then with the compatible matrix one, where 'Remerge' of the left dislocated XP finally licenses its [+REF] feature. The full derivation of (40a) is given in (43):[28]

(43) a. $_{\text{CYCLE-1 CP}}$[~~den~~ [C $_{\text{TopP}}$[den [Top $_{\text{TP}}$[der Peter $_{\text{AuxP}}$[$_{v\text{P}}$[der Peter $_{\text{VP}}$[~~den gesehen~~] gesehen] ~~hat~~] hat]]]]
b. $_{\text{CYCLE-2 CP}}$[den [C $_{\text{TP}}$[Maria $_{v\text{P}}$[Maria $_{\text{VP}}$[~~meinte~~] ~~meinte~~] meinte]]]
c. $_{\text{CYCLE-3 CP}}$[den [C $_{\text{TP}}$[Hans $_{v\text{P}}$[Hans $_{\text{VP}}$[~~glaubt~~] ~~glaubt~~] glaubt]]]
d. $_{\text{CYCLE-4}}$ [den Studenten]

The analysis in (43) tries to reflect our claim that the D-pronoun is visible at PF in the intermediate cycles. We contend that this is possible on the basis that (a) as he argues, Zwart's rule for positional marking applies within the cycle, and (b) that 'Chain Reduction', which, on principled grounds, is relevant only for non-trivial chains, applies as early as possible. We hypothesized that deletion of the phonological features of the intermediate D-pronouns takes place at the moment the relevant parts of the structure are assembled. In contrast, deletion of the morphophonetic features of the D-pronoun in the lowest clause is simply impossible,

28. See footnote 20.

given that they are associated with the only feature licensed in that cycle and, in fact, in the whole sentence, before 'Merge' of the extra-dependent constituent. Nevertheless, that the D-pronoun may sometimes comply with the V2 constraint in the lowest clause is demonstrated by the grammaticality of (44):

(44) *Den Studenten, den glaubt Hans, hat der Peter gesehen*
the-ACC student RP-ACC thinks Hans-NOM has the-NOM Peter seen
'The student, fortunately Peter saw (him)'

Examples of this kind are straightforwardly accounted for on Frey's (2004, 2005, 2006) assumption that long movement (in our terms, 'Remerge' with the matrix C) always correlates with contrastive interpretation. From this perspective, (44) reduces to a standard instance of 'Chain Reduction': as shown in (45), in the absence of formal features attached to the highest, [+CONTRAST] D-pronoun, it becomes the morphophonetically realized link.

(45) [den Studenten] $_{CP}$[den [C $_{TP}$[Hans $_{vP}$[Hans $_{VP}$[~~glaubt~~] ~~glaubt~~] glaubt]]]
[~~den~~ $_{[CONTR]\ [REF]}$ [C $_{TopP}$[~~den~~ $_{[CONTR]\ [REF]}$ [Top $_{TP}$[der Peter
$_{AuxP}$[$_{vP}$[der Peter $_{VP}$[~~den~~ $_{[CONTR]\ [REF]\ [TOPIC]}$ ~~gesehen~~] gesehen] ~~hat~~] hat]]]]

4. Conclusions

The present paper has explored some of the possibilities 'Survive Minimalism' may offer for the research on the core properties of language. Its focus has been GLD constructions, whose complexity mainly derives from the diverse and poorly understood processes involved in them. There is a considerable number of problems and pending issues that the analysis presented here leaves unsolved, among them: the exact syntactic structure of German D-pronouns, the potential relevance of Stroik's (2009) distinction between feature checking/deactivation and feature licensing, the question whether a notion of cyclicity different from that in Chomsky (2001, and subsequent work) may be compatible with the fundamental tenets of 'Survive', and issues related to an optimal design for the PF-component. There are other interesting topics that, mainly due to space limitations, I have not addressed, such as the kind of concatenative feature that is responsible for 'Merge' of the left dislocated XP, or the intriguing reasons for the distinction between substantive and functional lexical items with respect to the 'Survive Principle'.

Despite so many loose ends, I hope to have shown that accounts of the German left periphery that capitalize on the theoretical framework of 'Survive Minimalism' may benefit from this new look at old problems. This new perspective, in particular, contributes to a better understanding of the relation between the (conceptual) interpretive and the strictly syntactic components. I am confident that an equivalent enriching work on the PF side will follow in the future.

References

Anagnostopoulou, E. 1997. Clitic left dislocation and contrastive left dislocation. In *Materials on left dislocation,* E. Anagnostopoulou, H.C. van Riemsdijk & F. Zwarts (Eds), 151–192. Amsterdam: John Benjamins.

Bayer, J. 2004. Decomposing the left periphery: Dialectal and cross-linguistic evidence. In *The syntax and semantics of the left periphery*, H. Lohnstein & S. Trissler (Eds), 59–95. Berlin: Mouton de Gruyter.

Besten, H. den. 1983. On the interaction of root transformations and lexical deletive rules. In *On the formal syntax of the Westgermania,* W. Abraham (Ed.), 47–131. Amsterdam: John Benjamins.

Bhatt, R.M. 1999. *Verb movement and the syntax of Kashimiri.* Dordrecht: Kluwer.

Brandner, E. 2004. Head-movement in minimalism, and V2 and FORCE-marking. In *The syntax and semantics of the left periphery,* H. Lohnstein & S. Trissler (Eds), 313–341. Berlin: Mouton de Gruyter.

Cardinaletti, A. & M. Starke. 1999. Responses and demonstratives. In *Clitics in the languages of Europe,* H. van Riemsdijk (Ed.), 273–290. Berlin: Mouton de Gruyter.

Chomsky, N. 1995. *The minimalist program*. Cambridge MA: The MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In *Step by Step: Essays in minimalist syntax in honor of Howard Lasnik,* R. Martin, D. Michaels & J. Uriagereka (Eds), 85–155. Cambridge MA: The MIT Press.

Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A life in language,* M. Kenstowicz (Ed.), 1–52. Cambridge MA: The MIT Press.

Déchaine, R.M. & Wiltschko, M. 2002. Decomposing pronouns. *Linguistic Inquiry* 33: 409–442.

Demirdache, H. 1991. Resumptive chains in restrictive relatives, appositives and dislocation structures. Ph.D. dissertation, MIT.

Epstein, S.D. & Seely, T.D. 2002. Rule applications as cycles in a level-free syntax. In *Derivation and explanation in the minimalist program,* S.D. Epstein & T.D. Seely (Eds), 65–89. Oxford: Blackwell.

Epstein S.D. & Seely, T.D. 2006. *Derivations in minimalism.* Cambridge: CUP.

Fanselow, G. 2003. Münchhausen-style head movement and the analysis of verb second. *UCLA Working Papers in Linguistics.*

Fanselow, G. 2004. Cyclic phonology-syntax interaction: Movement to first position in German. In *Information Structures* 1 [Working Papers of the SFB 632], S. Ishihara, M. Schmitz & A. Schwarz (Eds), 1–42. Postdam: University of Postdam.

Fodor, J.D. & Sag, I.A. 1982. Referential and quantificational identities. *Linguistics and Philosophy* 5: 335–598.

Frey, W. 2000. Über die syntaktische Position der Satztopiks im Deutschen. *ZAS Papers in Linguistics* 20: 137–172.

Frey, W. 2004. The grammar-pragmatics interface and the German pre-field. *Sprache & Pragmatik* 52: 1–39.

Frey, W. 2004a. Notes on the syntax and the pragmatics of German Left Dislocation. In *The Syntax and semantics of the left sentence periphery,* H. Lohnstein & S. Trissler (Eds), 203–233. Berlin: Mouton de Gruyter.

Frey, W. 2004b. A medial topic position for German. *Linguistische Berichte* 198: 153–190.
Frey, W. 2005. Zur Syntax der linken Peripherie im Deutschen. In *Deutsche Syntax: Empirie und Theorie,* F.J. D'Avis (Ed.). *Göteborger Germanistische Forschungen* 46.
Frey, W. 2006. Contrast and movement to the German pre-field. In *The architecture of focus* [Studies in Generative Grammar 82], V. Molnár & S. Winkler (Eds), 235–265. Berlin: Mouton de Gruyter.
Gärtner, H.M. 2002. On the force of V2 declaratives. *Theoretical Linguistics* 28: 33–42.
Grewendorf, G. 2002. Left dislocation as movement. *Georgetown University Working Papers in Theoretical Linguistics* 2: 31–81.
Grohmann, K.K. 1997. On left dislocation. *Groninger Arbeiten zur germanistischen Linguistik* 40: 1–33.
Grohmann, K.K. 2000. A movement approach to contrastive left dislocation. *Rivista di Grammatica Generativa* 25: 3–65.
Grohmann, K.K. 2003. *Prolific domains: On the anti-locality movement dependencies* [Linguistik Aktuell/Linguistics Today 66]. Amsterdam: John Benjamins.
Haegeman, L. 1997. Negative inversion and the structure of the CP. Paper presented at the Linguistic Colloquium, University of Wuppertal.
Hockett, C.F. 1958. *A course in modern linguistics.* New York NY: MacMillan.
Hoekstra, E. 1999. On D-pronouns and the movement of topic features. *Folia Linguistica* 33: 59–74.
Kayne, R. 1994. *The antisymmetry of syntax.* Cambridge MA: The MIT Press.
Lasnik, H. 2001. A note on the EPP. *Linguistic Inquiry* 32: 356–361.
Müller, G. & Sternefeld, W. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 24: 461–507.
Nunes, J. 2004. *Linearization of chains and sideward movement* [Linguistic Inquiry Monograph 43]. Cambridge MA: The MIT Press.
Putnam, M.T. 2007. *Scrambling and the survive principle* [Linguistik Aktuell/Linguistics Today 115]. Amsterdam: John Benjamins.
Putnam, M.T. & Stroik, T. In progress. The Survive manifesto.
Reinhart, T. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27: 53–94.
Reis, M. 1996. Extraction from verb-second clauses in German? In *On extraction and extraposition in German*, U. Lutz & J. Pafel (Eds), 45–88. Amsterdam: John Benjamins.
Rizzi, L. 1997. The fine structure of the left periphery. In *Elements of Grammar,* L. Haegeman (Ed.), 237–279. Dordrecht: Kluwer.
Roberts, I. 1999. The C-system in Brythonic Celtic languages, V2 and the EPP. Ms, University of Cambridge.
Roberts, I. & Roussou, A. 2002. The Extended Projection Principle as a condition on the tense-dependency. In *Subjects, expletive and the EPP*, P. Svenonius (Ed.), 125–156. Oxford: OUP.
Stroik, T. 1999. The Survive principle. *Linguistic Analysis* 29: 278–303.
Stroik, T. 2000. *Syntactic controversies.* Munich: Lincom.
Stroik, T. 2009. *Locality in minimalist syntax*. Cambridge MA: The MIT Press.
Thiersch, C. 1978. Topics in German syntax. Ph.D. Dissertation, MIT.
Truckenbrodt, H. 2006. On the semantic motivation of syntactic verb movement to C in German. *Theoretical Linguistics* 32: 257–306.
Uriagereka, J. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26: 79–123.

Vat, J. 1997. Left dislocation, connectedness and reconstruction. In *Materials on left dislocation*, E. Anagnostopoulou, H. van Riemsdijk & F. Zwarts (Eds), 67–92. Amsterdam: John Benjamins.
von Stechow, A. & Sternefeld, W. 1988. *Bausteine syntaktischen Wissens*. Opladen: Westdeutscher Verlag.
Wiltschko, M. 1997. Parasitic operators in German left dislocation. In *Materials on left dislocation*, E. Anagnostopoulou, H. van Riemsdijk & F. Zwarts (Eds), 307–339. Amsterdam: John Benjamins.
Wiltschko, M. 1998. On the syntax and semantics of (relative) pronouns and determiners. *Journal of Comparative Germanic Linguistics* 2: 143–181.
Zwart, J.W. 2005. Verb second as a function of Merge. In *The function of function words and functional categories*, M. den Dikken & C.M. Tortora (Eds), 11–40. Amsterdam: John Benjamins.

Tense, finiteness and the survive principle

Temporal chains in a crash-proof grammar*

Kristin M. Eide
Norwegian University of Science and Technology

This paper describes how temporal chains are construed in a syntactic structure. The links in T-chains are local T-heads, where every main verb and auxiliary brings its own tense package. The semantic difference between finite and non-finite T-elements consists in the choice of first argument, the speech event S (finite tense) or any preceding verbal event e (non-finite tense). Overt inflectional markings encoding finiteness are a crucial difference between Modern English and Mainland Scandinavian languages (MSc). Middle English, like MSc, encoded finiteness. MSc retained the finiteness distinction but lost the agreement markings; English main verbs lost the finiteness distinction but retained their tense and agreement markings. This development fuelled many syntactic differences between MSc and English, e.g., *do*-support versus verb-second.

0. Introduction

Chomsky's (1955, 1957) analysis of English verbal morphology is considered one of the most brilliant achievements of the Generative Enterprise. In Lasnik's words (1999: 98):

> Much of the apparent chaos of this central portion of English morphosyntax was rendered systematic by the fundamental insight that the tense-agreement inflectional morpheme ("C") is syntactically independent, even though always a bound morpheme superficially.

---

*I thank the audiences at various conferences and seminars where this material has been presented, in particular the ScanDiaSyn Grand Meeting in Mýatn, Iceland, in August 2007. I also want to thank Howard Lasnik, Terje Lohndal, Fritz Newmeyer, and John McWhorter for generously sharing their insights and commenting on previous versions of this paper. I want to thank the anonymous reviewer, who also gave me fruitful comments and suggestions, and Mike Putnam for his encouraging comments, his patience, and his hard work in putting this volume together. As always, thanks to Viktorija Todorovska for her help with the final touches.

The syntactically independent tense head T continues to play a crucial role in most Chomskyan theories of clausal syntax. T is ascribed important case and agreement properties, it selects (little) V, and is selected by other heads and predicates, such as C and matrix V. Certain properties of T are assumed to condition its selection by other heads: whether T is finite or non-finite, tensed infinitival or non-tensed infinitival. Moreover, the temporal properties and (non-)defectiveness of a given T head are claimed to be responsible for the formation of control and raising structures, the presence and absence of PRO, DP-traces and EPP effects, the finiteness of the clause and its transparency vs. opacity to several syntactic operations. In short, almost every aspect of our theory needs to refer to the properties of T, and almost every head in a proposed clausal structure is, in one way or another, T-related.

Given these facts, it is reasonable to say that the temporal properties of T, understood as the morphosyntactic and semantic properties of this pivotal clausal head, have received less attention than they deserve. With some very notable exceptions,[1] our proposals often contain little information about what a T head contains, how it behaves in a syntactic structure, and what it does semantically. Traditionally, our proposals simply refer to T as an abstract syntactic head, in spite of its central place. This proposal addresses what happens when a T-head enters a derivation, what temporal features this T-head has, and what kind of argument it takes.

Although Chomsky's analysis of English verbal morphology is a brilliant achievement, several issues need to be resolved before it can reach its full explanatory power. For instance, we still need a robust explanation for a) what distinguishes English main verbs from auxiliaries, allowing the latter to raise e.g., past negation; b) what accounts for the different patterns of auxiliary types in non-finite domains; and c) what makes English different from its Germanic siblings, leading to the lack of generalized V2 in declaratives and the need for *do*-support?

The present proposal outlines a theory in which the properties often ascribed to "independent T" are attributed to other sources. Firstly, the temporal properties of T are determined by a T-head locally associated with each and every V (including auxiliaries). I argue that every V comes with its own tense package. These tense elements inevitably create temporal chains imposing a relative ordering of events, at least between the verbal event and the speech event S. Secondly, AGR(eement) is partly unrelated to the FIN(iteness) category, except that they tend to appear in the

1. I will mention specifically the work of Tim Stowell, Jacqueline Guéron and Jacqueline Lecarme, Hamida Demirdache and Myriam Uribe-Extebarria. Cf. e.g., the collection of papers in Guéron and Lecarme (2004).

same contexts. I argue that this can be accounted for by Rizzi's (1990b) generalization that anaphors and agreement are typically mutually exclusive. Thirdly, what separates English from its Germanic siblings is the different distribution of the morphosyntactic feature of non-derivative finiteness. The trigger for V2 (whatever it is) requires checking by the morphosyntactic feature [+FIN]. In most Germanic languages, any verb may satisfy this requirement. In English, however, this feature is encoded only in auxiliaries. Although the assertion-forming head of English declaratives at some point in history stopped asking for this feature, certain other heads in the syntactic structure of English still require it, e.g., inversion-triggering operators such as the question operator Q and sentential negation (Σ negation, cf. Laka 1990). These can only be checked by a finite auxiliary. The distribution of the FIN feature is also responsible for the differing behaviour of main verbs and auxiliaries in Present Day English: main verbs do not have this feature, whereas some auxiliaries have only [+FIN] forms (modals, auxiliary *do*) and others have full paradigms with both finite and non-finite forms (*have* and *be*).

The theory is implemented in a crash-proof grammar approach of the type outlined in Putnam (2007) and Stroik (2009) and shares many of the fundamental views of Frampton and Gutman (2002).[2] First, I address the notion of finiteness, since it is key to many of the theoretical considerations I will subsequently invoke. In Section 2, I present a different approach to finiteness, one based on the concept of temporal chains. This concept is further developed in Section 3. In Section 4, I flesh out some important consequences of the distribution of the finiteness feature, explaining the aforementioned differences between English and other Germanic languages and between English main verbs and auxiliaries. Section 5 summarizes and concludes the paper.

1. Finiteness

Discussions of finiteness features in natural language usually start out with the observation that the term *finiteness* is used to cover many different (albeit partly overlapping) concepts. Some authors have also suggested that the term is not amenable to a theory-proof definition (Koptjevskaja-Tamm 1994). As Cowper (2002: 1) states, "The term 'finite' has been used in the grammatical literature for centuries, but its meaning is difficult to pin down." However the term is used, most authors agree that the two occurrences of *be* in (1) are non-finite, and *is* is finite.

2. Although my outline of T-heads will be very different from Frampton and Gutman's (2002) TE-heads (temporal-event heads) and also serving a very different purpose.

(1) To be or not to be – that is the question.

Nikolaeva (2007: 1 ff) offers an informative overview of the origins and uses of the term: she explains that the term *finite* in European linguistics goes back to the Latin *finitus* in the sense of 'referring to a particular person'. The term was first applied to personal pronouns and later to "verbs expressing person and number." Still later, tense is taken into account, and today the typical diagnostic for finiteness is the presence of morphologically expressed tense and agreement features on the verb.

The main objections to this view are easily recognized. The first one is the lack of universality: not all languages express agreement or tense features overtly (i.e., morphologically) and, even in languages that do, verb forms such as the imperative are considered finite even if there are no tense or agreement markings. Even if T(ense) and AGR(eement) features are considered covertly present in these 'finite' syntactic domains, T and AGR features are not always indicators of finiteness. There are languages that employ agreement features in what are obviously non-finite domains (e.g., inflected infinitives in European Portuguese); also, semantic and morphological tense distinctions are found with non-finite verbs (e.g., Latin infinitives). These facts have led some authors to give up on isolating the specific contributions of finiteness. Adger (2007: 58), for instance, states that since the canonical markers of finiteness, T and AGR, are not confined to the topmost layers of the clause, but may appear "very low down" (i.e., in the VP), we may conclude that "there is no clear mapping from the traditional notion of finiteness to the categories of formal grammar."

The most widespread view in Principles and Parameters Theory (P&P), even in the most recent approaches, maintains that finiteness is a binary category which, in addition to controlling the realization of the subject argument and creating domains opaque to certain syntactic processes, regulates the tense and agreement features on the verb (Nikolaeva 2007: 6). When morphologically expressed, these features are considered reliable indicators of a finite domain; when the features are not overtly realized, we could be dealing with a finite domain with only abstract (i.e., covert) T and AGR features or a truly non-finite domain. Likewise, overtly realised T and AGR markings in non-finite domains are sometimes claimed to be subject to operations that 'cancel them out' under certain conditions, i.e., make them count as syntactically non-finite, as in 'finite control constructions' like Balkan embedded subjunctives, cf. Landau (2004). This means that current P&P theory accepts that there are 'finite' verbs with no T or AGR marking and 'non-finite' verbs with T, AGR, or both. Although there clearly is no one-to-one correspondence between overtly realized T-AGR and finiteness, this approach to finiteness still appeals to inflectional criteria.

There are also distributional criteria for finiteness: finite verbs occur in independent contexts and non-finite ones in dependent contexts. This would result in the alleged generalization that only finite verbs can head independent utterances, and that each independent utterance can have one and only one finite verb.[3] The second part of this alleged generalization will be disputed below. Counterexamples to the first part of the generalization are the so-called Mad Magazine Sentences in (2), where seemingly non-finite verbs are the only verbs in independent utterances (Akmajian 1984; Siegel 1987). I will present additional evidence that such sentences are truly non-finite (Section 4.4; cf. also Grohmann 2000).

(2) a. *John eat caviar?* (*I thought he hated the stuff.*) Siegel (1987)

b. *Jeg ikke drikke øl på en fredag?* (*Særlig.*)
I-NOM not drink-INF beer on a Friday? As-if
'Me not drink beer on a Friday? As if.'

Even definitions that rely on distributional criteria consider finiteness a binary category with the values *dependent* and *independent*. In functionalist frameworks, however, finiteness is often seen as a scalar phenomenon signalled by features that combine to yield a specific degree of finiteness. These features are tense, modality, and aspect markers, agreement, case markings on the subject and object, articles, determiners, and topic markers (cf. Givón 1990: 853; Bisang 2001: 1401–2).

The controversies do not end here. In all approaches considered so far, finiteness could be considered a purely syntactic category or a combination of purely syntactic categories. Some authors, however, side with Barron (2000: 2), who claims that finiteness "rather than being a morphosyntactic category…is a semantic category related to time." Most generativists would probably agree with Platzack (1998: 371) that finiteness is a syntactic category with obvious semantic effects (e.g., instantiated by uninterpretable and interpretable features), and many functionalists would agree with Gretsch and Perdue (2007: 433) that finiteness must be analysed from structural as well as semantic and pragmatic points of view.

It seems uncontroversial that finiteness is accompanied by specific morphological markings (in the languages to be considered here, the markings typically appear on the verb); that it has syntactic characteristics (with regard to subject licensing and the V2 effect in relevant languages); and that it gives rise to specific interpretation effects (Platzack 1998: 371, for instance, advocates that finiteness expresses the anchoring of the clause in time and space). The observations mentioned so far constitute the necessary background for advancing a different

3. This definition also has ancient roots, going back to Stoic logic, cf. Nikolaeva (2007), quoting Luhtala (2000).

approach to finiteness. Crucially, I emphasize that finiteness must be distinguished from tense. Holmberg and Platzack (1995: 23) argue that although finiteness is a prerequisite for tense, it is not identical to tense since there are untensed finite verbs (imperative) as well as non-finite tensed verbs (past participles).

2. A different approach

The present approach to finiteness emerges from Comrie's (1985: 36) observation that in natural languages we find examples of *absolute* and *relative tenses.* Absolute tenses (such as the preterit) take the moment of speech as their deictic centre and relative ones (such as participles) take some other contextually-given time point as their point of reference. While in many languages specific verb forms encode the distinction between absolute and relative tenses, I argue that there are also languages where the same verb form is used for both. The context indicates whether the tense should be interpreted relative to the moment of speech (absolute interpretation) or relative to the previous verb (relative interpretation). The structurally highest verb or TMA marker in a verbal chain takes on the meaning of an absolute tense; the next verb takes the previous verb (or TMA marker) as its point of reference and yields a relative temporal reading. Note that the form of the verb remains the same in the absolute and relative function in the sentences in (3) from Capeverdean Creole (Baptista 1997).

(3) a. *El* ***ta*** ***sta*** *na* *kaza.*
S/he FUT be at home
'S/he'll be at home.'

b. *El* ***sta*** ***kume***.
She be eat
'S/he is eating'

c. *N* ***kume*** *tudu* *katxupa.*
I eat all katxupa
'I ate all of the katxupa.'

What I describe as the lack of formal distinctions between absolute and relative tense forms has been considered a lack of formal finiteness distinction in the literature on Creoles, e.g., Muysken and Law (2001). According to Romaine (1993: 62), Creoles typically lack non-finite verb forms, i.e., they use the same form in finite and non-finite functions. In most Germanic languages, however, the finiteness distinction is morphologically encoded. I argue that even in Germanic languages the distinction between finite and non-finite verb forms corresponds to the distinction between absolute and relative tense forms.

As Comrie (1985: 48) observes, most European languages have only a two-way split in the tense systems, past vs. non-past, "with subdivisions within non-past (especially future as opposed to the present) being at best secondary; thus the so-called present tense in such languages is frequently used for future time reference." Pairing this assumption with my claim that the absolute-relative distinction equals the finiteness distinction in the relevant languages, we arrive at the paradigm in (4) for the garden variety Germanic language. Here, each and every verb form encodes a tense element consisting of two pieces of information: [±Past] and [±Finite]. In this analysis, being tensed is an essential part of being a verb (present participles are adjectival, at least in Norwegian, cf. Faarlund, Lie, and Vannebo 1997: 119).

(4)

	+Finite	–Finite
+Past	**Preterite**	**Participle**
–Past	**Present**	**Infinitive**

I will argue that whereas the Mainland Scandinavian languages adhere to the paradigm in (4), in English this paradigm has collapsed into two two-way paradigms, both of which have lost the finiteness distinction. Hence, Present Day English main verbs are (productively) inflected according to the paradigm in (5a) and modals and the auxiliary *do* according to the paradigm in (5b). In fact, the only verbal elements in English that still adhere to the non-collapsed paradigm in (4) are the auxiliaries *have* and *be* (a more detailed discussion of the paradigms of *have* and *be* is given in Eide 2009).

(5) a.

+Past	**Preterit/Participle**
–Past	**Present/Infinitive**

b.

[+Past, +FIN]	**Preterit**
[–Past, +FIN]	**Present**

Whereas many theories in formal linguistics conceive of tense as a semantically (and in a sense syntactically) autonomous element tied to a specific position in the clause, the present theory claims that all verbs are inherently tensed. There is thus no reason to assume that certain verbs (such as infinitives or epistemic modals) are immune to tense marking since tense is part of the definition of being a verb (at least for the languages presently under consideration). Of course, the expression of *finiteness* is sometimes tied to specific positions in a clause, e.g., to the V2 position of root clauses in V2 languages. As emphasized earlier, however, finiteness is not tense. Instead, I will argue that finiteness has much in common with pronominality: finiteness in the verbal domain behaves like pronominality in the nominal domain. If these claims have merit, one would also expect verbal tense forms that behave like anaphors. This is argued in Eide (2007) as well as here. I argue that finite and non-finite tense elements are subject to the principles of Binding Theory. In languages productively employing the finiteness distinction, such as Mainland

Scandinavian, finite forms (absolute tenses) behave like temporal pronouns and non-finite ones (relative tenses) like temporal anaphors.[4]

An anaphor requires a local, c-commanding antecedent;[5] the typical non-finite verb also requires a local, structurally higher verb to which it relates temporally. A pronoun, in the words of Apollonios Dyskolos of Alexandria (2nd century AD), quoted in von Heusinger (2002), is either deictic or anaphoric.[6] The same is true of finite tense elements. In the deictic case, a finite tense element finds its point of reference in the immediate context, usually the moment of speech or speech event S. In the anaphoric case, the finite tense element relates to a salient, linguistically expressed element, e.g., a verb in a preceding clause. This is what happens in sequence-of-tense constructions and presumably also in the 'historic present'.

3. Temporal referential chains

In the present approach, all verbs are tensed, and being tensed is an essential part of being a verb. However, not all verbs are finite; typically, only the structurally highest verb in a verbal sequence carries finiteness markings. The overtly finite verb typically conveys an absolute tense and takes the speech event as its reference point; the next verb in the sequence takes the previous verb as its reference point, and so on. This gives rise to a tongue-and-groove system, where each verb contains a tense element hooking up to the previous event, and each verb (or auxiliary) provides a temporal anchor for the next verb (or auxiliary).

One way of implementing this tongue-and-groove system incorporating the semantics of tense elements is to consider tense elements dyadic predicates of temporal ordering (Stowell 1995; Julien 2001). I will adopt this analysis and will also

4. Cf. Partee (1973, 1984), Enç (1987), and Stowell (2007) for other, different approaches to the anaphoric properties of tense. Non-finite verb forms are not considered tense marked in these approaches (except for the past participle in Stowell's paper). It follows that non-finite forms are irrelevant for their discussion. The proposal of Guéron and Hoekstra (1995) is similar to the present one in spirit, if not in terminology and form.

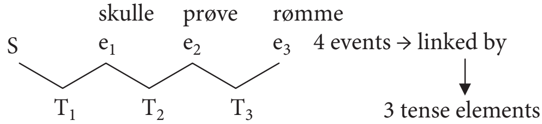
5. I will ignore long-distance anaphora in this discussion although I believe that this type of anaphora also exists in the temporal domain. A good candidate, I believe, is the Swedish embedded perfect participle in constructions like: *Han sager att hon inte kommit* 'He says that she not arrivePERF.PART'.

6. I will ignore the bound-variable reading in this discussion. Cf. e.g., Büring (2007) for a discussion and von Heusinger (2002) for a historical overview of the deictic-anaphoric distinction of pronouns.

follow Giorgi and Pianesi (1997) and Julien (2001) in assuming that the arguments of these dyadic predicates are events (where the term *event* should be understood to encompass dynamic events and states). As we will see, this does not conflict with the central claim that tense elements behave like pronouns and anaphors, or that absolute tense amounts to finite and relative tense to non-finite in languages that maintain the finiteness distinction.

I furthermore assume that each verb contains an event argument (*e* in the sense of Davidson 1967) in addition to a tense element and that each tense element anchors the event argument of its own verb to the preceding event. The speech event S also counts as an event, which means that in a clause with *n* verbs, the total number of relevant events is *n+1*. In typical cases, the speech event functions as the anchor for the first (i.e., finite) verb, the event argument of the first verb is the anchor for the tense element of the second verb, and so on. This gives rise to referential chains of temporality, in principle not very different from nominal referential chains; see the temporal referential chain in (6) versus the nominal one in (7).

(6) *Marit [skulle] [prøve] å [rømme].* **Norwegian**
'Marit would try to escape.'
skulle prøve rømme



(7) *[Marit] sa at [hun] kunne høre [seg selv] le av [seg selv].* **Norwegian**
'Marit said that she could hear herself laugh at herself.'

A non-finite tense element behaving like an anaphor needs to find its antecedent in the same clause to be referentially bound. This antecedent thus serves as the first argument of the tense element. Given Relativized Minimality (Rizzi 1990a),[7] this implies that a non-finite tense element must be bound by the verb immediately preceding it in the same clause. Compare this to the anaphor *seg selv* in (7), which also must be bound by a local antecedent. Just like *seg selv,* the temporal anaphor must be part of a chain where the topmost element is referential in order to yield a specific reference. In the nominal referential chain [*hun, seg selv*$_1$, *seg selv*$_2$] in (7), *seg selv*$_2$ is bound by *seg selv*$_1$, which is bound by *hun.* In the temporal referential chain [*skulle, prøve, rømme*] in (6), the tense element of *rømme* is bound by *prøve,* which

7. Relativized Minimality imposes a locality constraint in a structure [..X..Y..Z..] such that the relation between X and Z is licit only if there is no Y (with the relevant feature) structurally closer to Z than X. Stated in Chomsky's (2001) terms, α in [… β …[α …[… γ]]] prevents β from probing γ for any active feature shared by α and γ.

gets its tense element bound by *skulle.* Likewise, since there is a verbal sequence in (7), it too gives rise to a temporal referential chain. In the temporal chain of the embedded sentence in (7) [*kunne, høre, le*], the tense element of *le* is bound by the event of *høre,* which has its tense element bound by the event denoted by *kunne.*

A finite tense element behaves like a pronoun and must be unbound within its clause; it finds its first argument outside its clause. It may be bound by something in the context (deictic) or by some syntactically realised antecedent in the matrix clause. Consider the nominal and temporal referential chains of (7). Either *hun* has a deictic reference or it is bound by *Marit.* The choice of binder for *hun* affects the reference of *seg selv*$_2$ and *seg selv*$_1$. Likewise, the finite tense element of *kunne* may be bound deictically (by the speech event S[8]) or by an antecedent in the matrix clause; in (7), the verb *sa* is a likely candidate. The latter choice gives rise to a sequence-of-tense construal. In either case, the referential construal of the finite tense element of *kunne* affects the reference of the non-finite, anaphoric forms *høre* and *le,* just like in a nominal referential chain.[9]

---

8. According to the anonymous reviewer, there are many examples where embedded finite verbs are unable to access the temporal coordinates of S, as in this Hebrew example:

(i)	*Dan*	*xashav*	*she-Dina*	*maclixa*	*ba-mivxan.*
	Dan	thought.PAST.3SG.M	that-Dina	succeeding.PRES.3SG.F	in.the-exam
	'Dan thought that Dina succeeded on the exam'.				

Here the time of the succeeding event can only be simultaneous to the thinking event, and cannot be related to the time of utterance, according to the reviewer; adding that the embedded verb in such constructions is "uncontroversially finite". While I am not familiar with the facts in Hebrew, I know from Landau's (2004) work that such constructions in Hebrew also allow for control into embedded "finite" clauses, given that the embedded verb is future and the embedding verb is directive or comissive (i.e., "future-projecting"; cf. Bybee et al. 1994). This leads Landau to classify these constructions in Hebrew as subjunctives syntactically (albeit not morphologically). It is a typical trait of subjunctives crosslinguistically to be temporally dependent on their selecting predicate. I am not in a position to claim that this must be the case in the Hebrew constructions, but this is the direction I would start looking for an explanation of the 'no S-related reading' restriction for these constructions.

9. Although it would take another paper just to begin to do justice to this question, there are alternative ways of approaching the Sequence-of-Tense phenomena (SoT). One approach that would allow us to assume that finite tenses are always absolute is to assume that even the finite embedded tense in SoT contexts takes S as its point of reference (like other finite tenses) and also that the SoT reading of an embedded finite past is simply a sub-case of the independent interpretation of PAST. Stowell (2007: 450) argues against this approach and says that "The main challenge for this independent-tense theory is to explain

In this system, a possible complex tense or temporal chain is not restricted by an upper number of reference times R, as it would be in a (neo)Reichenbachian

why the complement clause E(vent) T(ime) cannot be understood to be subsequent to the main clause ET." In other words, the embedded event in SoT allegedly cannot be construed as future with respect to the main clause event. However, this is not quite accurate. Firstly, SoT is possible in these contexts e.g., in Hebrew; cf. Landau (2004: 820). Secondly, there are many so-called future-projecting predicates (like *foresee, predict*) even in English that may take embedded main clauses with simple past (without the modal *would*), where the embedded clause gives rise to exactly the type of SoT construal Stowell claims to be impossible. For instance, the data in (i) through (iv) were found via a quick Google search.

(i) *In 1978 scientists predicted that Mount St. Helens was soon to erupt.*
(ii) *We predicted that one team had a 60% chance or better to cover the spread.*
(iii) *He foresaw that the fate of Zionist settlement in Palestine depended on the creation of a strong Jewish economy.*
(iv) *[…] Antonia Stone foresaw in the 1980s that people without access to computers risked being left behind.*

Although these examples can easily be construed with a relative present reading of the embedded event, a relative future reading is also possible. Thus, an independent-tense theory of SoT is not ruled out.

On the other hand, there is evidence that the SoT reading of the embedded past needs to be licensed by the embedding predicate. This does not necessarily imply that SoT is embedding of tense; instead, it might be a type of quotative modality reminiscent of the German Konjunktiv II (past subjunctive). One piece of evidence is that SoT is impossible if the verb licensing it is not a claim-type of verb, but a verb selecting a (semi-)factive predicate; compare the SoT in (v) (from Wurmbrand 2007: 5) to my own non-SoT example in (vi):

(v) *John promised me yesterday to tell his mother tomorrow that they were having their last meal together.*
(vi) *I decided yesterday to let John discover tomorrow that he was terminally ill.*

In (v), the event of having their last meal together can easily be construed as simultaneous to the embedding telling-event, but in (vi), John's being terminally ill cannot be construed as simultaneous to the discovering-event (unless this is construed as a narrative text in the past tense). I believe this is due to the different embedding predicates, *tell* vs. *discover*. Likewise, it is not sufficient to have the semantically right kind of embedding predicate; there is also a requirement that the embedding verb form could be interpreted as a derived past. Compare (v), where *tell* is a relative tense form taking *promised* as its point of reference and is not explicitly, i.e., formally, present or future, to (vii) and (viii), where the telling event is specified as a non-past (data from Wurmbrand 2007: 5, 8).

(vii) *John promised me yesterday that he will tell his mother tomorrow that they were having their last meal together.*
(viii) *John will promise me tonight to tell his mother tomorrow that they were having their last meal together.*

tense system (e.g., Vikner 1985; Hornstein 1990, and many others). Instead, the number of possible temporal relations depends entirely on the number of verbs in the chain (cf. also Julien 2001). Just like in a nominal referential chain, a pronominal element breaks the chain, since it has 'independent', or absolute, reference; it is +**R** in the sense of Reinhart and Reuland (1993) and Reuland and Reinhart (1995); an anaphoric element simply extends the chain. This property of referential chains is restricted by the Condition on A-chains:

> *Condition on A-chains* (Reuland & Reinhart 1995: 255):
> A maximal A-chain ($\alpha_1, \ldots, \alpha_n$) contains exactly one link – α_1 – which is +R.

As long as this condition is satisfied, and each temporal referential chain contains exactly one +R, finite, link, nothing prevents language from creating temporal chains of absurd length; (8) shows a temporal chain with seven verbs.

(8) a. *Han måtte ha villet prøve å la dem se henne svømme.*
He mustPAST have wanted try to let them see her swim
'He had to have wanted to try to let them see her swim.'

måtte	[+PAST,+FIN]	$(S > e_{MÅTTE})$
ha	[–PAST,–FIN]	$\neg(e_{MÅTTE} > e_{HA})$
villet	[+PAST,–FIN]	$(e_{HA} > e_{VILLET})$
prøve	[–PAST,–FIN]	$\neg(e_{VILLET} > e_{PRØVE})$
la	[–PAST,–FIN]	$\neg(e_{PRØVE} > e_{LA})$
se	[–PAST,–FIN]	$\neg(e_{LA} > e_{SE})$
svømme	[–PAST,–FIN]	$\neg(e_{SE} > e_{SVØMME})$

This tense-chain consists of eight events, one encoded by each verb plus S, the speech event. These events are ordered by seven tense elements, each encoded by a tense affix on a verb. Only the top-most tense element is finite; this is the only tense element that takes S as an argument. All other tense elements in this chain are non-finite; each takes the event denoted by the c-commanding verb as its first argument. Two are 'past' relations (noted here as ($e_1 > e_2$), i.e., e_1 *after* e_2) and five are 'non-past' (noted here as $\neg(e_1 > e_2)$, i.e., e_1 *not-after* e_2). All non-past relations may in principle be construed as 'future' or 'present', depending on the aspectual properties of the predicates involved (both the selector and the selectee).[10] The default strategy is to construe the

10. It is relevant to note here that Lasser (1997: 35) proposes a Non-Completedness Constraint on root infinitives in child and adult use. Root infinitives encode future or ongoing situations; they never signify past or completed events. An adult root infinitive can be placed in a 'historic narrative' and denote a relative future situation as in the German example in (i) or it can denote an ongoing state as in the Dutch example in (ii). However, root infinitives

non-past predicate as (relative) present if it is stative and as a (relative) future if it is dynamic. This is the case for the finite non-past; i.e., the present in (9a) and (9b) and the non-finite non-past, i.e., the infinitive in (9c) and (9d). Whereas the finite non-past is 'present' or 'future' relative to S, the non-finite non-past is 'present' or 'future' relative to the embedding predicate (i.e., the previous verb in the temporal chain).

(9) a. *Marit kommer.* *kommer* [–PAST,+FIN] ¬(S >e_{KOMME})
Marit comesPRES
'Marit will come.'

b. *Marit liker Jon.* *liker* [–PAST,+FIN] ¬(S > e_{LIKE})
Marit likesPRES Jon.
'Marit likesPRES Jon.'

c. *Marit må komme.* *komme* [–PAST,–FIN] ¬($e_{MÅ}$ > e_{KOMME})
Marit must come
'Marit must come.'

d. *Jon må være på kontoret* (*før ni*) *være* [–PAST,–FIN] ¬ ($e_{MÅ}$ > $e_{VÆRE}$)
Jon must be in his office before nine
'Jon must be in his office (by nine).'

e. *Jon spiser (når han kommer).* *spiser* [–PAST,+FIN] ¬(S > e_{SPISE})
Jon eatPRES when he arrivePRES
'Jon is eating/Jon will eat when he arrives.'

The default construal can easily be overridden by an adverbial forcing an iterative meaning (hence stative or 'present' reading) of the otherwise dynamic predicate or a future-denoting adverbial forcing a future construal of the otherwise stative predicate. The latter is shown in (9d) for the non-finite non-past and in (9e) for the finite non-past. Without the adverbial, the 'present' reading is more natural, both for *være* 'be' and *spiser* 'eats'. Note that the tense element itself does not refer to the aspectual properties of the predicate; it contains only two bits of information: [±Past], determining the temporal ordering, and [±Finite], determining whether the first argument is clause external (S) or clause internal.

Thus, one possible construal of the temporal chain in (8) *Han måtte ha villet prøve å la dem se henne svømme* 'He had to have wanted to try to let them see her swim' gives rise to a time line like the one in Figure (1), with a number of arbitrary choices.

can never denote relative past or completed events. These findings are corroborated in Blom (2003: 76) for root infinitives in Dutch child language.

(i) *Und dann noch die U-bahn nehmen.*
And then still the subway take
'And then we still had to take the subway.'

(ii) *Zeker weten?*
sure know
'Do you know that for sure?'

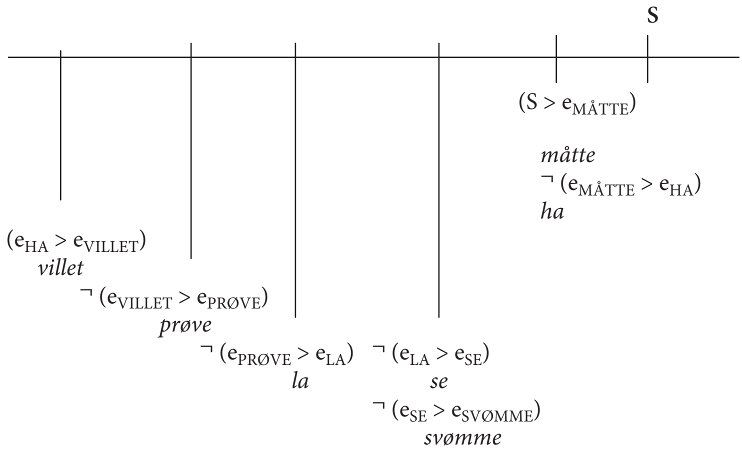


Figure 1.

Although I have not yet fully investigated to what extent the binding theory constructed for nominal reference pertains to all facets of tense elements,[11] it is suggestive that both nominal and temporal anaphors seem to be sensitive to the presence of certain 'barrier' categories, such as DPs. Thus, assuming that the English progressive, being non-finite, is a temporal anaphor might help explain the parallel behaviour of the progressive and the nominal anaphor *herself* in (10).

(10) a. *Mary*$_i$ *showed Lucy*$_j$ [$_{DP}$ *a picture of herself*$_{i/j}$].
b. *Mary*$_i$ *let Lucy*$_j$ *help herself*$_j$.
c. [$_{DP}$ *Those sitting*$_{T1/S}$ *on the floor*] *were*$_{T1}$ *expelled.*[12]
d. *Sitting*$_{T1}$ *on the floor, John was*$_{T1}$ *expelled.*

In (10a), *herself* may be bound by either *Mary* or *Lucy;* in (10b) the closest c-commanding antecedent *Lucy* is the only possible binder. Likewise, in (10c), the

11. An anonymous reviewer asks the following question, [A]s the author acknowledges… (fn. [10]) the dependence of tense in embedded contexts is highly sensitive to the particular embedding predicate chosen in each case. This seems to have no parallels in the domain of referential chains… [W]hy don't we find instances of (nominal) referential chains exhibiting sensitivity to the particular verbs that intervene?" This is an excellent question, and I think a possible candidate for such sensitivity in the nominal domain would be the "backward binding" phenomenon, where certain predicates (like e.g., psych verbs) give rise to unusual binding patters. Moreover, approaches such as Reinhart and Reuland (1993), Reuland and Reinhart (1995) ascribe the various binding possibilities of nominal referential chains exactly to the semantic properties of the selecting predicate.

12. Thanks to Renaat Declerk for the example in (10c).

event encoded by the progressive may be simultaneous with S ('those presently sitting on the floor') or (near-) simultaneous with the time of the expelling event. Without the DP barrier in (10d), the sitting event can no longer be bound by S, but must be (near-)simultaneous with the expelling event. Again, the anaphor must be bound by the closest c-commanding binder.[13]

A final point helps illustrate the similarities between temporal and nominal pronouns and anaphors. As illustrated by the Capeverdean Creole data in (3), in a language without a finiteness distinction (i.e., a language lacking overt markers of temporal pronominality), the same form can be used in finite and non-finite environments. This has a counterpart in the nominal domain: in certain Romance languages, e.g., European Portuguese, in 1st and 2nd person accusative the same clitic form can be used as an anaphor or a pronoun, depending on the context, as shown in (11).[14]

(11)	**Acc. pronoun (cl.)**	**Acc. anaphor (cl.)**		
	me	*me*		me/myself
	te	*te*		you/yourself
	o	*se*		him/himself
	a	*se*		her/herself
	nos	*nos*		us/ourselves
	(vos	*vos)*	(rare)	you/yourselves
	os	*se*		them (m)/themselves
	as	*se*		them (f)/themselves

Most main verbs in Present Day English are like the Capeverdean Creole verbs in (3): they do not encode an overt distinction between finite (pronominal) and non-finite (anaphoric) forms (except for 3rd person singular). This is similar to the ambiguous nominal forms in (11): only the context indicates whether the form is an anaphor or a pronoun.

In what follows, I will advance two major claims. Firstly, I will argue that the finiteness distinction amounts to a morphologically encoded distinction between absolute (pronominal) and relative (anaphoric) tense forms. Secondly, the loss of the morphologically encoded finiteness distinction in main verbs caused Present

---

13. Potentially related is the fact that in a DP with an embedded relative clause, as in (i), the finite non-past tense of the relative clause can be construed as simultaneous to the speech event S (on a *De Re* interpretation) or simultaneous to the future point in time at which the seeing event will take place (on a *De Dicto* interpretation).

(i) *John will see the unicorn that is walking on the meadow.*

14. Cf. also Schlenker (1999) for a discussion of logophoricity, pronouns, and tenses.

Day English syntax to be dramatically different from the syntax of Middle English and that of Mainland Scandinavian. Section 4 fleshes out the underpinnings and consequences of these claims, pinpointing the semantic, syntactic and morphological difference between MSc and English verb forms.

4. The finiteness distinction and its consequences

Unlike Creole languages and Present Day English, Mainland Scandinavian languages (MSc) productively encode finiteness with all types of verbs; any verb form in MSc is inherently specified for finiteness (i.e., 'pronominality'). A [+Finite] form like *spiser* 'eatPRES' or *spiste* 'ate' has distinct morphology to show that it is an absolute tense form; no context needs to be provided to determine this. Likewise, *spise* and *spist* are unambiguously non-finite, infinitive and past participle, respectively, regardless of the context. In contrast, an out-of-context main verb in English, e.g., *love* or *loved,* could be either an absolute or a relative tense form. *Loved* could be either a preterit or a participle, *love* either a present or an infinitive; in fact, *love* could even be a subjunctive or an imperative. Just like in the Creole languages discussed earlier, the context will disambiguate the English tense form; the tense element of the structurally highest verb takes on the function of absolute tense, even when the verb is not morphologically specified for finiteness. Thus, the key difference in the construal of temporal chains in MSc and English is that in the typical MSc temporal chain the structurally highest verb is obligatorily [+Finite] morphologically. In the corresponding English temporal chain, the structurally highest verb takes on the absolute interpretation, but there is no requirement for a morphologically [+Finite] verb, except in certain contexts which I will discuss later. Otherwise, the forming of the temporal chain proceeds similarly in MSc and English: each verb links up to the previous event in the chain, taking the previous verb (or event) as its point of temporal reference.

4.1 A crash-proof derivation of temporal chains

How can we model this semantically and syntactically relevant difference between morphologically finite verb forms, i.e., pronominal tense elements (typically) relating to S, and verbs forms that are indifferent to the type of event they link up to, like the English main verbs? Preferably, the semantic feature separating these two types of verb forms should be minimal and should play a part in the derivation of the clause. In what follows, I outline one possible implementation of this in a type of crash-proof grammar.

Stroik (2009) proposes a theory seeking to meet the requirements of a crash-proof grammar envisaged in Frampton and Gutmann (2002). Stroik's theory rejects the existence of uninterpretable features (cf. also Stroik & Putnam 2005) and advocates a grammar where syntactic operations are not designed to check lexical features for interface compatibility; rather they check for concatenative integrity (Stroik 2009: 31–32):

> A well-formed derivation will exhaustively build its form and meaning, as interpreted at the interfaces, out of the inter-relationships of all the features of all its lexical items. Each and every lexical feature contributes something to the concatenative integrity of a derivation by linking appropriately to the features of other lexical items, thereby generating a derivational string that is interpretable at the interface. Syntactic operations serve to link these features by placing them in feature matching domains where the feature can be cross-checked for compatibility. Hence, syntactic operations are designed to build crash-proof syntactic objects from lexical features by linking compatible features together. All the features of all the lexical items in a Numeration must contribute to the concatenative integrity and the interface-interpretation of a derivational string. If even one feature on a single item fails to match in a derivation, the derivation will be incomplete and will not be sent to the interfaces to be processed/interpreted.

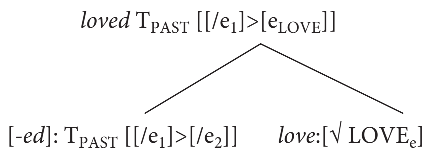
Another crucial ingredient of the theory is the Survive principle, described by Putnam (2007: 6) as follows (cp. Stroik 2009: 69 for a slightly different formulation):

> As an alterative to *Internal Merge/Move*, Stroik [1999…and 2009] proposes that immediately following the concatenation of X and Y to form K, Y is automatically repelled back to the *Numeration* to await another future opportunity to *Remerge* into the narrow syntax. This *Remerger* will take place when another head (Z) enters the derivation bearing a matching feature with Y, in this case, β. Under such an analysis, Y survives its initial concatenation with X and must participate in another concatenate union with another head (Z) bearing a matching feature in order to ensure derivational harmony. Stroik refers to this view of syntactic 'displacement' as the *Survive Principle*:
>
> *The revised Survive Principle* (based on Stroik 1999: 286):
> If Y is a syntactic object (SO) in an XP headed by X and Y has an unchecked feature [+F] that is incompatible with the feature X, Y remains active in the Numeration.
>
> Although at first blush the *Survive Principle* […] looks familiar to *Internal Merge/Move*, the key advantage to this principle is that syntactic objects do not remain active in the narrow syntax and are therefore not subjected to unwanted look-ahead or look-back mechanisms.

In the theory of temporal chains I am advocating, the relevant temporal features contributing to the concantenative integrity of a syntactic string are of four types: (1) tense elements, which are dyadic predicates; (2) events, the arguments of the temporal predicates; (3) the finiteness specification, requiring a particular kind of argument for certain verb forms; and 4) the selectional requirements of specific verbs, accepting certain complements and rejecting others. We will address all four types.

When a verb enters a syntactic derivation from Numeration, it is already equipped with a tense element. This tense element is a dyadic predicate, imposing a particular temporal relation on its two arguments (cf. Section 3). I will write this as $T_{PAST}[[/e_1]>[/e_2]]$, denoting that a tense element, in this case encoding past, requires two event arguments; [Past] orders these two arguments temporally such that e_1 is AFTER e_2. A tense element cannot help but take as one of its arguments the event variable of the verb of which it is an affix. The first concatenation, the union of the verb root with its tense marker, thus presumably happens in a morphological component prior to selection for the given Numeration.[15]

(12)



I will assume that this union of a root with a tense marker is exactly what creates a verb. Most approaches assume that a verb root becomes a verb when merging with small v; the present approach, on the other hand, assumes that the distinctive feature of a verb is tense (even in languages that do not employ overtly distinguishable tense markers).[16]

In the concatenation in (12), the root LOVE provides T with its first event argument. This syntactic object is a verb with a tense affix, where the T imposes a temporal relation between the loving event and some other event e_1 yet to be specified. T thus survives this morphological merger because it needs to check another argument. The tensed verb enters the Numeration with one active T-feature, the unchecked argument position. It can only check this second argument when it encounters another event in the derivation. Moreover, whenever T remerges to check its next argument, the verb will be pied-piped since T is not an independent lexical item. It needs the verb, itself an affix, as its vehicle.

---

15. Although it is also possible to assume that this is a syntactic requirement, as pointed out by the reviewer.

16. If language is a symbolic system allowing us to create and manipulate symbolic elements in a mental model of the world, any natural language, hence Universal Grammar, should at least capture the fact that we exist in time and space. In my opinion, verbs and linguistic elements denoting events and states exist along the axis of time. Although it is often said that events (and to a lesser extent, states) *take place*, this is in fact inaccurate. Events do not take *place*, they take *time*. I strongly suspect that what is sometimes referred to as the syntactic

Next, the affixed verb will need to check its Theta-features, i.e., merge with the appropriate arguments. It survives these concatenations because of its unchecked T-argument e_1 (I am not considering the case-checking operations of the object *John* or the possible workings of little v here).

(13) a. Merge {loved, John} → loved John
b. Survive {loved}
c. Merge {Mary, {loved, John}} → Mary loved John
d. Survive {Mary}
e. Merge {S,{Mary, {loved, John}}} → S Mary loved John
f. Remerge {loved{S,{Mary, {loved, John}}}} → loved S Mary loved John
g. Remerge {Mary{loved{S,{Mary,{loved, John}}}}} → **Mary loved** S Mary loved **John**

After the merge operation in (13c), if the Numeration is void of other verbs or auxiliaries, T pied-piping the verb *loved* must remerge because T needs another argument. If there is no other verbal event in the structure, T takes S to check its argument position e_1. I will assume (at least for now) that S is represented as a type of abstract verb in the clause (cf. Ross 1970 for a very early expression of this idea[17]). Thus, T waits for S to enter the derivation (13e) and immediately remerges

backbone of the clause, i.e., the set of operators belonging to the CP-IP-VP domains, in fact constitute an abstract representation of a time line. This would account for our intuition that all the major clausal heads are somehow T-related (cf. section 0). To partake in a time line, however abstract, a verb must have temporal properties. These temporal properties are provided by the tense element T, imposing an interpretation on the root along the time line.

A time line alone does not constitute a proposition, however. A proposition ensues when the (abstract) time line axis is crossed by the (abstract) space axis. The space axis is represented by entities and locations. Entities and locations exist and extend in space, unlike events. I suspect that this is the fundamental reason for the EPP (in its more classical version, not necessarily understood as a generalized dislocation trigger). The interfaces will not interpret a syntactic object as a proposition unless it concatenates a time line element with a spatial element in a particular fashion, thereby depicting a particular point in time and space. Thus, it is to be expected that there are various ways to fulfil the EPP in natural language: through location or person features (representing the participants in the speech event) or through the expletive subjects we know from well-studied languages (cf. Ritter & Wiltschko 2005). Also, we find the workings of the EPP even with predicates that do not express tense overtly since predicates extend along the time axis.

17. As pointed out to me by the anonymous reviewer, Ross' intuition encodes this performative component at the level of the sentence. This is not a problem for this particular derivation, but recall that my analysis needs to allow for even deeply embedded finite verbs to have access to this "performative syntactic object", i.e., the speech event S. Thus, it might be more consistent and empirically adequate to consider S in this derivation as a pronominal element referring to the speech event.

to check its remaining T-argument feature. According to Putnam (2007: 20), this new object is 'interface-interpretable' immediately upon concatenation because of the fusion of features. This is the operation ***Link!*** According to Putnam (2007: 20, n. 12), it is similar in design and function to Kobele's (2006) *Principle of Immediacy*, which requires features to be checked as soon as appropriate configurations arise in the narrow syntax.

> ***Link!*** (Putnam 2006: 6)
> If XP is a syntactic object (SO) bearing the interpretable feature α, and a syntactic head Y^0 enters the derivation bearing the interpretable feature α, XP will remerge into the structural position immediately dominating Y^0 [Spec YP], at which point the matching features will amalgamate and be interpreted at the interfaces.

Although this operation is formulated for remerging XPs, I will assume that the verb-T complex falls under the same definition in the relevant sense. Thus, *loved* remerges in (13f) and has its remaining T-argument feature checked by S. This gives rise to a T element with both event positions valued and checked: *loved* T_{PAST}[[S]>[e_{LOVE}]]; i.e., the speech event is temporally ordered as being **after** the loving event.

Stroik (2009: 89 ff) claims that the sensimotoric interface will interpret only overt morphophonetic features and it will interpret these features where their concatenation relations are checked. This usually means that the sensimotoric interface will interpret only the highest copy of a chain. In our derivation, we are assuming that the subject *Mary* has to remerge again in (13g) to check a surviving case feature since the subject must check its case by a (semantically) finite T. The verb *loved* does not count as finite and doesn't know it is finite until it has taken S to check its argument feature e_1. Thus, the subject cannot get its case feature checked until S is introduced in the derivation.

We see this more clearly if we let the partial derivation in (13d) take a different direction, one where there is an auxiliary in the Numeration. The auxiliary must also combine with its tense element before it is selected for Numeration. This combination results in an auxiliary with a tense affix. For simplicity, we assume that the tense affix denotes past for this auxiliary *had.* As a result of this union of the tense element and the auxiliary, the first event argument of the tense element T is set to be HAVE, i.e., the event of the auxiliary itself. It may seem controversial to assume that auxiliaries denote their own events instead of operating on other events, but the tense system I advocate does not care whether tense affixes attach to auxiliaries or main verbs. The mechanics of this temporal tongue-and-groove tense system are blind to the properties of the given verb (although the aspectual properties of the verb play a part in specifying certain relations, cf. Section 3). Moreover, it seems fair to say that the auxiliary *have* denotes a state (cf. Eide 2005: 301 ff and 367 ff for a detailed discussion of the stative properties of *have*). Thus, the auxiliary *have* should partake in temporal chains like other (stative) verbs and behave like any other verb regarding temporal ordering. When the tense element combines with the auxiliary HAVE, it cannot but take the HAVE-event as its argument e_2.

(14) *had* $T_{PAST}[[/e_1]>[e_{HAVE}]]$

[-*d*]: $T_{PAST}[[/e_1]>[/e_2]]$ *ha*-:[√ $HAVE_e$]

Now *had* is in the relevant Numeration with one event argument checked and another unchecked. It enters the derivation at the point where the subject has checked the Theta-features of the main verb. Up to (15d), the derivation is identical to the one in (13).

(15) a. Merge {loved, John} → loved John
b. Survive {loved}
c. Merge {Mary, {loved, John}} → Mary loved John
d. Survive {Mary}
e. Merge {had,{Mary, {loved, John}}} → had Mary loved John
f. Survive {had}
g. Remerge {loved,{had,{Mary, {loved, John}}}} → loved had Mary loved John
h. Merge {S, {loved,{had,{Mary, {loved, John}}}}} → S loved had Mary loved John
i. Remerge {had,{S,{loved{had{Mary,{loved, John}}}}}} → had S loved had Mary loved John
j. Remerge {Mary{had,{S,{loved{had{Mary,{loved, John}}}}}}} → **Mary had** S **loved** had Mary loved **John**

In (15e), the auxiliary *had* is introduced into the derivation. This allows for the T of the verb *loved* to remerge in (15g) and check its argument e_1 by the event variable HAVE of the auxiliary *had*. The T element of *loved* is rendered inactive, with both event positions valued and checked: *loved* $T_{PAST}[[e_{HAVE}]>[e_{LOVE}]]$; the having event is temporally ordered after the loving event and *loved* is 'past relative to *had*', which gives the right result. However, *had* does not get its event argument e_1 checked in this concatenation; presumably, it is introduced at this point to check its selectional requirements, e.g., that it must take a VP (or vP) complement. Moreover, although the checking relation is to some extent reciprocal (cf. the Link! operation), there is also a requirement that one of the elements is the checker and the other is being checked; cf. Stroik (2009: 48):

(16) a. Lexical Items (LIs) are Syntactic Objects (SOs)
b. If A [+F] and B[$+F_c$] are SOs, then Merge ⟨A,B⟩ is an SO.
c. Nothing else is an SO.

> [16b] states that A and B can combine only if they share at least one atomic feature together with its value [+F] and the feature of B is specified as a feature-checker $[+F_c]$. Requiring B to be marked with the $[+F_c]$ feature rather than just the [+F] feature is necessary to prevent an LI from licitly combining with itself—which could occur if merge merely required two LIs to share features, since an LI necessarily shares features with itself.

The same restriction ensures that *had* cannot check its remaining T-argument in (15e) or in (15g). In (15g), *had* is the checker and *loved* is being checked. Thus, *had* has yet to check its remaining T-argument feature and must wait for the introduction of another event. This event is S, merged in (15h). Now *had* can remerge and check its second T-argument. In doing so, *had* becomes finite, since its e_1 argument is S: *had* $T_{\text{PAST}}[[S]>[e_{\text{HAVE}}]]$; the speech event is ordered temporally after the having event. Via transitivity, the temporal chain in (15) gives rise to a time line like (17).

(17)

The subject *Mary* must once again wait for the verb to become finite so it can check its case features. This is the final remerger in (15j). Unlike in (13), the verb *loved* cannot check the case feature of the subject *Mary* since *loved* in (15) is not finite. It does not take S as its e_1 argument, but e_{HAVE}. The auxiliary *had* is the finite verb in (15): it takes S as its e_1 argument and can check the case feature of the subject *Mary.*

Our derivations so far serve to illustrate the construction of temporal chains in modern English. I argue that the distinctive feature of English main verbs is their indifference to the type of event their tense elements take as an e_1 argument. As mentioned earlier, this is a feature that English main verbs share with verbs in Creole languages. However, this feature makes English main verbs distinct from verbs in other Germanic languages, e.g., Mainland Scandinavian, which employ a productive finiteness distinction. Based on the system outlined so far, we can model this distinction quite simply by assuming that a finiteness distinction specifies the concatenation requirements, the requirements that need to be satisfied for the concatenative integrity of a syntactic string. The finiteness distinction instructs the computational system to put overtly finite forms only in checking domains where they can be checked by S, i.e., take S as their e_1 argument, and non-finite forms only in the checking domains of other verbs, where they can take a verbal event as their e_1 argument.

To repeat a crucial point of the analysis from Section 2, I argue that whereas all verbs in Mainland Scandinavian languages adhere to the paradigm in (18a), in

English this paradigm has collapsed into the paradigm in (18b), where the finiteness distinction is lost. Hence, Present Day English main verbs are (productively) inflected according to the paradigm in (18b).

(18) a.

	+Finite	–Finite
+Past	**Preterit**	**Participle**
–Past	**Present**	**Infinitive**

b.

+Past	**Preterit/Participle**
–Past	**Present/Infinitive**

This is the morphological expression of the indifference of English main verbs to the type of event they take as their e_1 argument. MSc verbs, on the other hand, require S as their e_1 argument if the form is finite, a verbal event as their e_1 argument if the form is non-finite. Thus, we can enrich the paradigms in (18) with this information, yielding the paradigms in (19), which expresses both the morphological finiteness and the semantic ramifications of this distinction, the specifications for possible e_1 argument.

(19) a.

	+Finite	–Finite
+Past	**Preterit:** *Likte* T_{PAST} [[/S]>[/e_2]]	**Participle:** *Likt* T_{PAST} [[/e_{1V}]>[/e_2]]
–Past	**Present:** *Liker* $T_{\neg PAST}$ [[/S]>[/e_2]]	**Infinitive** *Like* $T_{\neg PAST}$ [[/e_{1V}]>[/e_2]]

b.

+Past	**Preterit/Participle:** *Liked* T_{PAST} [[/e_1]>[/e_2]]
–Past	**Present/Infinitive:** *Like* $T_{\neg PAST}$ [[/e_1]>[/e_2]]

MSc verbs have retained their morphological finiteness distinction and are hence specified as to whether their e_1 argument is S or the preceding verbal event. English main verbs have lost their morphological finiteness distinction and are indifferent as to whether their first argument is S or the preceding verbal event. Thus, the forms of English main verbs are ambiguous between the preterit (absolute) and participle (relative) function and between present (absolute) and infinitive (relative) function; they have only one generalized past form and one generalized non-past form. In what follows, we will examine more closely some syntactic ramifications of the loss of this finiteness distinction in English main verbs. We will also see that there are verbal elements that have maintained a finiteness feature even in Present Day English. But first, let us turn to the fourth type of concatenation feature relevant for temporal chains: the selectional requirements of auxiliaries and verbs.

We know that the auxiliaries *have* and *be* require their complements to be VPs (or vPs) and the verbal head of the complement to be a [+Past]. Modals have similar requirements; as auxiliaries,[18] they require their complements to be VPs and the verbal heads of these complements to be [–Past], i.e., an infinitive.[19]

(20) a. We have known/*knew/*know him for years.
b. He must know/*knew/*known many languages.

Historically, modals were directive and hence future-projecting: their complements had to denote a situation that was 'temporally future with respect to the modal'. The epistemic readings of modals developed from the directive readings and did not "become common until quite late" (cf. Bybee et al. 1994: 195). With the possibility for an epistemic reading comes the possibility for a non-future reading of the complement of a modal, e.g., a simultaneous situation. In (21), both a future and a simultaneous reading of the complement (the *be*-situation) are possible because the infinitive allows for relative 'future' and relative 'present' readings. The epistemic modal facilitates a 'present' reading of the infinitive; the directive reading of the modal requires a 'future' reading.

(21)

Jon	*må*	*være*	*på*	*kontoret*	(*før*	*ni*).
Jon	must	be	in his	office	before	nine

'Jon must be in his office (by nine).'

Epistemic readings of modals also allow for modal statements about past situations: *It must be the case that John left yesterday.* However, we cannot construct a clause where the modal takes a past verb form as its complement because the modal will not accept anything but an infinitival verb as its verbal complement. Thus, instead of (22a), we need to use a periphrastic construction, where the infinitival form of the auxiliary *have* fulfils the selectional requirement imposed by the modal and the participle fulfils the semantic requirement for a past.

(22) a. **John must left yesterday.*
b. *John must have left yesterday.*

---

18. In most Germanic languages, modals can also take other types of complements, unlike in Standard English, where modals no longer have Thematic properties.

19. There are some counterfactual constructions in Norwegian and Faroese where the modal seemingly takes a participle as a complement, cf. (i), from Eide (2005). However, as argued quite convincingly by Julien (2003), these participles are irrealis infinitives. A strong argument for this is that these 'participles' take infinitival markers in certain contexts (cf. ii).

(i)

Dette	*kunne*	*vært*	*en*	*løsning.*
this	might	been	a	solution

'This might have been a solution.'

(ii)

Det	*hadde*	*vore*	*best*	*å*	*reist/*	*reise*	*no*.
It	had	been	best	to	leave.PTCP/	leave.INF	now

'It would have been best to leave now.'

This, I believe, is an example of a situation described by Frampton and Gutmann (2002: 96) as conflicting specifications by interfaces and narrow syntax:

> Plausibly, another requirement of optimality is that the derivational system generates a wide range of structures, though it is almost certainly too much to expect that it generates everything permitted by the constraints of the interfaces. There will likely be semantically perfect structures which cannot be generated by narrow syntax.

This is exemplified in (22a), which is semantically perfect, but cannot be generated by narrow syntax because of the selectional requirements of the modal.

Another requirement on the complement of auxiliaries is that they not show agreement. Traditionally, this effect has been attributed to the non-finiteness of complements of auxiliaries, since finiteness was believed to equal tense + agreement. In my theory, there is a looser relationship between agreement and finiteness. The fact that [+FIN] and [+AGR] tend to occur in the same contexts can be attributed to the Anaphor-agreement effect of Rizzi (1990b: 26):[20]

> *The anaphor agreement effect:*
> Anaphors do not occur in syntactic positions construed with agreement.

In my theory, verb forms (inherently encoding tense elements) are temporal anaphors when non-finite and temporal pronouns when finite. If temporal anaphors (non-finite verbs) are subject to Rizzi's restriction, we would expect non-finite verbs to be mutually exclusive with agreement, unless they show a particular type of anaphor-agreement (cf. Woolford 1999). Finite verbs, on the other hand, would be expected to show agreement. This would be true even for English main verbs in finite functions, i.e., when they take S as their first T-argument.

A crucial point in my analysis is that English main verbs have lost the finiteness distinction, whereas the verbs in Mainland Scandinavian have retained this distinction. This difference between MSc and English could be an accidental fact of homophony and need not point to any profound differences between the two languages. However, as I will demonstrate in the sections that follow, the ways in which finiteness is instantiated in MSc and English give rise to a number of syntactic differences between these languages as well as between Middle English (ME) and Present Day English (PDE). Once the morphological differences are

20. As noted by the reviewer, it would be interesting to discuss the agreeing infinitives in European Portuguese in this perspective. This is very true, especially if we could argue that the type of agreement in these constructions is in fact an instance of Woolford's (1999) particular anaphor-agreement type, e.g., because it gives rise to other semantic effects than ordinary "finite" agreement.

observed, a number of syntactic differences fall out. To a linguist, this is always a most welcome result.

4.2 Verb second and verbal inflection

All Germanic languages except Present Day English have generalized verb movement to the second position of the clause in declarative main clauses. According to McWhorter (2005: 287, quoting Hopper 1975: 82), the general consensus is that verb second, V2, was a Proto-Germanic feature. Until approximately the fifteenth century, i.e., throughout Old English (OE) and into Middle English (ME), English had generalized V2 (Roberts 2007: 58; among others). At some point, the V2 rule ceased to be obligatory; in Lightfoot's (2006) terms, the cue for V2 (which he formulates as $_{CP}$[XP $_{c}$V]) becomes an optional, not obligatory, grammatical trait for new native speakers. In Present Day English, verb second is ungrammatical with main verbs in most of the contexts where it was (nearly) obligatory in OE and ME. In certain structures, such as non-subject root *wh*-clauses, a 'residual V2' is still obligatory (cf. Rizzi 1990a), but it is instantiated by specific auxiliaries instead of main verbs. I will discuss some of these structures in sections 4.4 and 4.5.

There have been many attempts to explain what caused this syntactic change in English and relate it to other, preferably morphological, changes in the verbal paradigm. To explain the relationship between loss of agreement and loss of V2, Roberts (1985: 32) suggests that "obligatory [V2] movement of main verbs should apply to languages with rich agreement systems only." However, the data from MSc are at odds with this claim: MSc languages have obligatory V2 with main verbs, but less agreement than English; in fact, MSc has no agreement at all in the relevant contexts. Roberts (1993: 120) modifies his earlier claims, emphasizing the difference between English and MSc as regards infinitival endings: MSc has retained and English has lost distinct infinitival endings, so the infinitive in English (unlike in MSc) is non-distinct from the imperative and many other verb forms.[21]

Noting that the loss of V2 is generally linked to the erosion of verbal inflectional morphology, McWhorter (2005: 287 ff) offers a comprehensive overview of the arguments and concludes that "overall, the explanations…lack explanatory power or falsifiability" and "it seems clear that the link [between verb movement

21. In contrast, Solà (1996) suggests that English main verbs are inflected neither for tense nor for agreement and should be considered (present or past) participles. Thus, the lack of verb movement with main verbs stems from the presence of a null auxiliary that represents tense (and agreement) and blocks the movement of the main verb.

and inflectional morphology] is too weak in itself to offer a conclusive explanation for what happened in English in comparison to its sisters."

I believe to have isolated the relevant inflectional feature. It is not the erosion of mood or agreement or the differing properties of tense that caused the loss of V2 in English. Nor is it the expression of a subset of agreement markings (e.g., 1st and 2nd person), tense, or mood markings. The relevant feature is the (non-derivative) finiteness feature, which amounts to an explicit distinction between absolute and relative verb forms. This distinction is encoded as finiteness in Germanic languages. It is the loss of this feature that made English main verbs behave differently from auxiliaries and from main verbs in other Germanic languages.

4.3 More on the morphological expression of finiteness

Any analysis treating finiteness as a derivative composed of 'more primary' features, such as tense and agreement, fails to account for the finiteness-related differences between MSc and PDE. English has 'more agreement' than MSc, but neither 'more' nor 'less' tense. However, as mentioned in Section 2, English does have much more homophony in the inventory of verb forms than MSc. To appreciate the consequences of this, we need to understand that the infinitive and the past participle are genuine, full-fledged participants in the basic tense system, giving rise to a four-way tense paradigm with two finite and two non-finite forms; cf. the table in (4), repeated here as (23a). Over time, English main verbs developed into a system employing the collapsed paradigm in (23b); here the finiteness distinction, the distinction between absolute and relative tense forms, is lost. This results in a paradigm employing one generalized [+Past] and one generalized [–Past] form.

(23)

a.	+Finite	–Finite
+Past	**Preterite**	**Participle**
–Past	**Present**	**Infinitive**

⇨

b.

+Past	**Preterite/Participle**
–Past	**Present/Infinitive**

Old English was like present day Mainland Scandinavian: it had distinct forms for all four cells in this paradigm; this was the case for both strong verbs (where the preterit is formed with *ablaut,* i.e., vowel shift) like *singan* and weak verbs (where the preterit is formed with a dental suffix, e.g., *-de, -ed*) like *lufian*, cf. paradigm (24a). English, however, is changing before our eyes. It is becoming more like the Creole languages mentioned earlier. The finiteness distinction starts to disappear in weak verbs, which adhere to the collapsed paradigm. In (24b), one generalized [+Past] form covers the function of the preterit and the past participle, and one [–Past] form covers the present and the infinitive. Note that nothing has been said so far about the 3rd person singular, which has a distinct form for main verbs even in PDE. I am going to assume that the 3psg ending *-s* is an agreement marker that plays no role in the tense system.

English regular weak verbs have been inflected according to the collapsed paradigm for centuries, as shown in (24b), while irregular strong verbs have maintained a slightly more complex system. Although there is only one generalized [–Past] form, there have been distinct forms for the preterit and the past participle. Thus, instead of a four-way system, Standard English has a three-way system, as in (24c). Currently, the remains of the finiteness distinction in main verbs, hitherto maintained in strong irregular verbs, are disappearing from many variants and dialects of English,[22] resulting in alignment with the rest of the system, as in (24d). This gives rise to constructions like *Here's what I would have wrote* and *You should have went yesterday.*[23]

(24) a.

	+Finite	**–Finite**
+Past OE: MSc:	**Preterit:** *sang, lufode* *sang, likte*	**Participle:** *gesungen, lufod* *sunget, likt*
–Past OE: MSc:	**Present:** *singe, lufie* *synger, liker*	**Infinitive:** *singan, lufian* *synge, like*

b.

Weak verbs	
+Past PDE	**Preterit/Participle:** *liked, killed, smiled*
–Past PDE	**Present/Infinitive** *like, kill, smile*

c.

Standard strong verbs	**+Finite**	**–Finite**
+Past PDE	**Preterit:** *went, saw, drove*	**Participle:** *gone, seen, driven*
–Past PDE	**Present:** *go, see, drive*	**Infinitive:** *go, see, drive*

d.

Dialectal strong verbs	
+Past PDE	**Preterit/Participle** *gone, seen, drove*[24]
–Past PDE	**Present/Infinitive** *go, see, drive*

Standard English strong irregular main verbs maintain a distinction between the preterit and the past participle, as in (24c). However, these forms are in a sense relics,[25] like certain obligatory case endings in modern Norwegian. They are too

22. I.e., different varieties of English. Eisikovits's (1987) study is of inner-city Sydney, but as pointed out by Trudgill and Chalmers (1991: 52), her generalizations apply to other dialects of English. Sampson's (2002) study is mainly of British English dialects; the data are taken from the CHRISTINE/I corpus; cf. Sampson (2002) for details.

23. Cf. Eide (2009) for a discussion on the dialectal facts of English participles.

24. Not all dialects of English take the same path in simplifying the paradigm. In some cases, the preterit form is the one covering both functions, e.g., *should-a-went*; in other cases, a new form, neither the original preterit nor the past participle, is employed as the [+past] form.

few and do not provide enough evidence for a finiteness system for English main verbs. I argue that although there are a number of main verbs where the old finiteness distinction is still visible in the [+Past] forms, the finiteness system ceased to be productive for English main verbs many centuries ago. In fact, my analysis predicts that it must have disappeared at about the same time as the generalized V2 rule. I discuss the relationship between finiteness and raising more thoroughly elsewhere (cf. Eide 2007). For now, suffice it to say that the present approach encompasses the insight of Roberts (1993 and subsequent) that the loss of V2 in English is related to the loss of infinitival endings and Solà's (1996) insight that this loss is related to the non-distinctness of English preterits from past participles.

As finiteness is encoded in the verbal form, any main verb can serve as a cue for the child learning a language. The child simply has to observe the same verb in absolute and relative functions, note whether the forms are different, and set the finiteness parameter correctly. Eventually, more and more L1 learners of English set the parameter as *finiteness not encoded* for main verbs. Thus, the paradigm in (23b), with one generalized [+Past] form and one generalized [−Past] form, becomes the productive paradigm. This leads to a situation reminiscent of that in the Creoles discussed in Section 2, where the same verb form is used for both absolute (finite) and relative (non-finite) tenses.

4.4 Auxiliaries and main verbs

The literature discusses a number of differences between main verbs and auxiliaries in PDE with regards to sentential negation (Section 4.5), subcases of inversion,

25. F. Newmeyer and J. McWhorter commented on previous versions of this paper, and they both hesitate to accept the term *relics* for these forms. F. Newmeyer (p.c): "I am a bit uncomfortable with calling verb forms distinguished in the past and perfect 'relics'. Many of the most common verbs in English behave that way, so the distinction seems very much part of our competence. Also, the distinction is being newly-created. When American English created 'dove' as past tense of 'dive', it left 'dived' as the perfect." I use the term *relics* to indicate that this distinction is no longer productive in Present Day English, and new verbs entering the language are not inflected according to the paradigm in (24c), but the paradigms in (24b) and (24d). I am not at all implying that the ± finite distinction is not part of the linguistic competence of speakers of Present Day English (as seems to be implied in a comment by the anonymous reviewer). In fact, certain syntactic rules can only be accounted for if the ± finite distinction is observed (cf. section 4.4. and 4.5). On the other hand, although strong irregular verbs are frequent, they are few. Pinker (2000: 18) claims that are only about 150 to 180 irregular verbs in modern English, and "there have been no recent additions. The youngest one is probably *snuck*, which sneaked into the language over a century ago and is still not accepted by purists." My point here is simply that the ± finite distinction for main verbs is no longer syntactically relevant, even to the extent that it is in fact maintained in the paradigm for a given verb.

wh-question root clauses, etc. There are also differences within the category of auxiliaries that are much less discussed and still lack a robust explanatory analysis, a topic to which we return shortly. The similarities and differences between English main verbs and auxiliaries have been a recurring topic in syntactic theories at least since Chomsky (1955, 1957).

Although Chomsky (1957) lists auxiliaries, i.e., modals and *have* and *be,* as appearing under a common (albeit phrasal) node Aux (e.g., p. 111), he groups main verbs and auxiliaries together under a common 'v' in certain structural analyses (e.g., the rule of Auxiliary Transformation; p. 113, op. cit.). According to Lasnik (2000: 68), Chomsky's (1957) treatment of auxiliary verbs in fact renders them as "belonging to no syntactic category at all." This led Ross to question Chomsky's analysis of auxiliaries. Ross (1969) claims that all English 'auxiliaries' in are in fact part of the natural class of verbs. Furthermore, in *Syntactic Structures* (Chomsky 1957), the element *do* is described as the substitute "bearer of an unaffixed affix" (cf. also the "stranded affix" filter of Lasnik 1981), and Chomsky states that the *do* appearing in these contexts is, at least morphologically,[26] "the same element as the main verb in *John does his homework*" (Chomsky 1957: 62). I will dispute this claim shortly.

Roberts (1985) suggests that the syntactic differences between auxiliaries and main verbs in English are due to a semantic property—the fact that auxiliaries lack Theta-properties. As Lasnik (2000: 161) points out, however, "it is not obvious that auxiliary verbs (especially modals) don't have a theta-role to assign."[27] Roberts (1985) concedes that modals do contribute to the clause something resembling a Theta-role and, to account for this effect, adopts Zubizarreta's (1982) analysis of modals as modifiers analogous to Jackendoff's (1972) agent-oriented adverbs. Roberts' analysis is in relevant respects also the analysis adopted by Pollock (1989). Pollock's (1989) analysis is adopted by Chomsky in subsequent works even though instead of referring to the auxiliaries as lacking Theta-properties, Chomsky (1995: 198) characterizes them as semantically vacuous.[28] Although it seems rather uncontroversial to assume that

---

26. Lasnik (p.c) suggests that Chomsky means morphologically, not necessarily syntactically. However, I want to dispute this claim: main verb *do* has main verb inflection, i.e., no finiteness distinction, in some dialects, whereas auxiliary *do* has only finite versions, like the modals. Cf. Eide (2007) for more details.

27. Lasnik (1999: 103) on the alleged semantic vacuity for the auxiliary *be*: "First, it is not clear that *be* is always semantically vacuous, yet the syntactic behaviour of *be* in finite clauses is always the same. For example, it is reasonable to assume that in [(i)], *is* has the meaning of *exists.* Yet, as seen in (ii), it raises overtly nonetheless:

(i)	There is a solution.	(ii)	a.	There is not a solution.
			b.	Is there a solution?

the auxiliaries *have* and *be* are semantically light and perhaps not Theta-assigners,[29] modals pose a problem for this generalization. They seemingly contribute something resembling a Theta-role in many contexts, especially on their root readings (obligation, permission, and volition).[30] There are also a number of other differences between modals and the auxiliary *do,* on the one hand, and the auxiliaries *have* and *be* on the other that receive no explanation under the approaches outlined so far.

Schütze (2003) lists a number of contexts where there is a split between modals and the auxiliary *do* and *have* and *be*, e.g., subjunctives (25), *to*-infinitives (26), small clauses (27), *Why not* constructions (28), and others. All data and grammaticality judgements are Schütze's.

(25) Subjunctives
  a. It is vital that John be here on time.
  b. It is vital that John be smiling on the photograph.
  c. It is vital that Rover have eaten before we arrive.
  d. *It is vital that John do not be late.
  e. *It is vital that John will not come unprepared.[31]

---

28. "Raising of the auxiliaries reflects their semantic vacuity; they are placeholders for certain constructions, at most 'very light' verbs. Adopting the intuition (but not the accompanying technology), let us assume that such elements, lacking semantically relevant features, are not visible to LF rules. If they have not raised overtly, they will not be able to raise by LF rules and the derivation will crash."

29. Pollock (1989: 385, 386) "[*Have* and *be* and their French equivalents] have a unique status with respect to θ-theory: they arguably fail to assign any θ-role to the constituents they are subcategorized for.… That aspectual *be/être, have/avoir*, and 'passive' *be/être* are not θ-role assigners is not, I think, controversial."

30. This is particularly evident on the subject-oriented reading (term due to Barbiers 1995). As shown in Eide (2005), the root vs. epistemic distinction is not the relevant cut-off point between modals that assign an external Theta-role and those that do not. In fact, root modals are quite capable of taking expletive subjects. However, only subject-oriented modals allow for their complement to undergo pseudo-clefts; whether the modal is root or epistemic is not important (except for the fact that epistemic modals are never subject-oriented). Cf. the data from Eide (2005: 198), both modals are root modals, but only the one in (i) is subject-oriented (θ-assigning).

(i) *Det vi alle bør, er å tenke gode tanker.*
It we all should, is to think good thoughts
'What we all should (do), is think good thoughts'

(ii) **Det en kvinne bør, er å bli vår neste statsminister*
it a woman should, is to become our next prime minister
Intended: What should happen is that a woman becomes our next PM.

31. (25e) is grammatical on a non-subjunctive reading: 'The fact that John will not come unprepared is vital.'

(26) *To*-infinitives
   a. It is important (for everyone) to be on time.
   b. It is important (for a movie star) to be smiling whenever the paparazzi are nearby.
   c. It is important (for every applicant) to have finished high school.
   d. *It is important (for us) to do not leave her alone.
   e. *It is important (for us) to can be alone.

(27) Small clauses
   a. I made him be alone for a while.
   b. The director made us be dancing when the curtain opened.
   c. ?The coach made her not just have eaten when she came to practice.
   d. *The conductor made us do not sing the harmony line.
   e. *The therapy made her can/could walk again. (cf. The therapy made her be able to walk again).

(28) *Why (not)* constructions
   a. Why (not) be a responsible citizen?
   b. Why be working when you could be partying on the beach?
   c. ?Why not have made the appointment with her before she has a chance to make one with you?
   d. *Why do not go to the beach?
   e. *Why should/must stay home? (cf. ?Why be obliged to stay home?)

Schütze (op.cit. p. 406) notes that "For *do* and modals to pattern together against *be* and *have* in so many environments clearly should not be a coincidence." His explanation is that *do* and modals belong to the category Mood and that the contexts in (25) through (28) are either too small to contain a Mood projection at all (e.g., small clauses) or come with their own Ø Mood morpheme that blocks the insertion of any of the Mood heads, i.e., *do* or modals.

The present approach offers a unified account for the facts discussed in sections 4.2, 4.3, and 4.4. This finiteness-based analysis explains not only why main verbs stopped raising in English, but also why auxiliaries take part in the 'residual V2' constructions (some of which will be discussed more thoroughly in 4.5). It also explains why modals and the auxiliary *do* pattern together against the auxiliaries *have* and *be*. The reason is not Theta-properties, or semantic vacuity, or the presence of modal or mood-like properties in the auxiliary *do.* It is the morphosyntactic feature make-up of the verbs and auxiliaries, i.e., whether or not they encode the finiteness distinction.

Present Day English (PDE) main verbs inflect according to the collapsed paradigm in (5a), repeated here as (29a), with one generalized [+Past] and one generalized [–Past] form. They encode no finiteness distinction and have no slot for a finiteness feature. Modals and the auxiliary *do,* in contrast, have only finite forms, as in (29b). The auxiliaries *have* and *be* are the only PDE verbal elements with a morphologically encoded finiteness distinction; their paradigm is thus the four-way paradigm of all verbs in OE and present day Mainland Scandinavian, (29c).

(29)

+Past	**Preterit/Participle**
–Past	**Present/Infinitive**

a. PDE Main verbs

[+Past,+Finite]	**Preterit**
[–Past,+Finite]	**Present**

b. PDE modals, auxiliary *do*

	+Finite	–Finite
+Past	**Preterit**	**Participle**
–Past	**Present**	**Infinitive**

c. PDE *have/be,* MSc, OE

What do we gain by considering the morphosyntactic feature make-up of verbs and auxiliaries the crucial property explaining the facts discussed in this section? Firstly, we can maintain the assumption that main verbs, modals, and the auxiliaries *do, have,* and *be* form a 'natural class'. Secondly, the assumptions summarized in (29) are not a leap of faith; any traditional grammar of English will tell you that modals have no non-finite forms and we know that *have* and *be* may appear in finite as well as non-finite functions. The more radical assumptions about the lack of morphological finiteness in PDE main verbs also receive robust support from verbal paradigms, especially when we look at them diachronically. Thirdly, if we take seriously the idea expressed in Chomsky (1995: 169) that "variation must be determined by what is 'visible' to the child acquiring language, that is by the PLD [Primary Linguistic data]," the syntactic behaviour of main verbs vs. *have* and *be* vs. modals and *do* should be detectable on the basis of paradigmatic distinctions encoded in the morphosyntactic make-up, not a semantic quality that is gradable at best and questionable at worst.[32] What semantic distinction would make auxiliary *do* pattern like modals and unlike *have* and *be*? And how could the child find the right point on the semantic scale where 'be able to' and 'be obliged to' are on one end and *can* and *must* on the other, patterning with auxiliary *do*? The finiteness feature is easily detectable by the child. The verb has the same form in absolute and relative functions (main verbs), in which case there is no finiteness distinction; it occurs only in absolute functions (modals and auxiliary *do*), in which case it is always [+Finite];[33] or it has morphologically

32. Cf. Bouchard (1995: 41): "The very notions on which theta-roles are based are external to Grammar."

33. According to the anonymous reviewer, to acquire this generalization would require negative evidence. However, one could invoke something resembling Thráinsson's (1996: 261) *Real Minimalist Principle* guiding acquisition in such cases: "Assume only those functional categories that you have evidence for". One might assume that the child would not assign a syntactically relevant finiteness distinction to a given auxiliary unless there is positive evidence.

distinct forms in absolute and relative functions (*have* and *be*), in which case it encodes the finiteness distinction [±Finite] (cf. Eide 2009 for a discussion on the paradigms of *have* and *be*).[34,35]

Although PDE grammar no longer encodes finiteness in main verbs, it frequently refers to finiteness in grammatical rules. There is no conceptual necessity underlying this diachronic development; one may very well envision a line of events such that English grammar, after abandoning the finiteness system for main verbs, developed into a grammar where the finiteness distinction simply disappeared from the syntactic system. However, this is not what happened. Instead, many operators remained 'unaware' that main verbs could no longer provide the required finiteness distinction, like an amputee patient with a phantom limb. These operators, even in PDE grammar, keep selecting and requiring [+Finite], [–Finite], and *[+Finite] goals (or complements). *[+Finite] means that the operator rejects any complement positively specified for finiteness, but will accept a [–Finite] complement (i.e., a non-finite auxiliary) or a complement with no finite distinction [ØFinite] (i.e., a main verb). This is the case in Schütze's data in (25)–(28). In all these structures, the presence of a [+Finite] feature in the complement renders the construction ungrammatical. Since modals and the auxiliary *do* are always [+Finite], they are banned from these constructions. Main verbs are accepted because they encode no finiteness feature; the non-finite versions of *have* and *be* are acceptable because they are [–Finite].

Other PDE syntactic heads still ask for an active [+Finite] feature, like (Σ) sentential negation, cf. Section 4.5, and the question operator Q. The finiteness requirement of Q is a property that PDE still shares with its Germanic relatives. Again, main verbs and non-finite *have* and *be* cannot accommodate this requirement because main verbs have no finiteness feature (they are [ØFinite]) and non-finite *have* and *be* are [–Finite]. Only modals, auxiliary *do* and finite *have* and *be* encode the right morphosyntactic feature [+Finite]. In MSc, any main verb or auxiliary may satisfy Q since all verbs and auxiliaries encode the finiteness distinction. Of course, [–Finite] verbs would create a feature mismatch

34. It is conceivable that there are also elements in natural language that are always [–Finite], with no [+Finite] forms. The *-ing* form (gerunds, progressives) is a good candidate for an obligatorily [–Finite] form in PDE.

35. As the observant reader will notice, I am glossing over a huge big problem for the present analysis represented by the paradigm of *have.* Firstly, there is no evidence for a finiteness distinction in the paradigm of *have,* since it inflects like a main verb (like 29a). This suggests that there is also a category distinction at stake here, between auxiliaries and main verbs (cf. Eide 2009 for discussion). Secondly, why does main verb *have* behave like auxiliary *have* in some

and are thus banned from raising to Q even in MSc, as in (30a). However, all verbs have a [+Finite] variant to meet the finiteness requirement of Q, as in (30b). In PDE, non-finite *have* and *be* are excluded since they are [–Finite], as in (30c); finite *have* and *be* are allowed, however, as in (30d). Unlike MSc main verbs (30e), PDE main verbs would also be insufficient as goals for the probe Q since they have no finiteness feature (30f). Modals and the auxiliary *do,* in contrast, have the right feature: they are always [+Finite] and fit the job description perfectly (30g).

(30) Inversion and finiteness

a. **Ha John gått?* **Norwegian**
b. *Har Jon gått?* **Norwegian**
c. **Have John left?* (Cf. example (25c) It is vital that Rover have…)
d. *Has John left?*
e. *Gikk Jon?* **Norwegian**
f. **Left John?*
g. *Must/Did John leave?*
h. *Må/*Gjorde Jon gå?* **Norwegian**

Note that MSc languages have no verb corresponding to the auxiliary *do* in PDE (30h). This could be an accident, but since PDE auxiliary *do* is a substitute employed to fulfil the requirements of specific goals requiring finiteness, and since all verbs have their own finiteness distinction in MSc, this verb would be between jobs most of the time.[36]

---

variants of English? Thirdly, as pointed out by the anonymous reviewer, *have* and *be* pattern with main verbs and not with modals and *do* in e.g. imperatives. This represents another intriguing set of facts which I discuss at length in Eide (2009). The pattern of imperatives can also be accounted for in my analysis, with the added generalization that imperatives do not accept agreeing verb forms in English. This disqualifies *have* and *be* from negative imperatives, since the Σ negation requires a [+FIN] form (cf. section 4.5), but the imperative operator rejects agreeing forms. For *have* and *be,* all their finite forms are also (incidentally) agreeing forms. This means that although *have* and *be* are fine in positive imperatives, they must be replaced in negative imperatives by auxiliary *do.*

36. There is a verb *gjøre* 'do' that occurs in tag-questions (i) and in VP-fronted constructions (ii), with some properties of PDE *do.* However, it is never found in true *do*-support constructions such as residual V2 contexts.

(i) *Marit kommer ikke, gjør hun?*
Marit comes not, does she?

(ii) *[Drikke/Drakk seg full] gjorde han aldri.*
[Drink/Drank himself wasted] did he never

Note that it is finiteness, not agreement (encoded by the marking -*s*), that allows the auxiliary in inversion, as in (30d), and bans it from the constructions in (25)–(28). Modals never show agreement; neither does the preterit of the auxiliary *do.* We could stipulate that these forms have covert agreement, but this would provide poor cues for the child acquiring the system. Instead, I argue that the agreement marking (visible in 3PSG -*s*) is partly independent of the presence of a morphologically encoded finiteness feature; this marker shows up with PDE main verbs (with no finiteness distinction), with *have* and *be* (where finiteness is encoded), and with the auxiliary *do* (where the finiteness feature is obligatorily [+Finite]). Likewise, MSc languages have the finiteness distinction encoded in all verbs, although there is no agreement; there are also agreeing infinitives in many languages (e.g., European Portuguese). Thus, agreement is neither a necessary nor a sufficient condition for finiteness. Instead, I assume that the relationship between agreement and finiteness is looser than usually assumed, although they often show up in the same contexts. I have assumed that this is because of the Anaphor-agreement effect of Rizzi (1990b: 26); cf. Section 4.1 above.

To sum up, instead of proposing two widely different explanations for the data in (25)–(28) and (30), I propose a unified account based on the morphosyntactic feature of finiteness, (usually) easily detectable in verbal forms.

4.5 English sentential negation

The analysis of the behaviour of PDE main verbs and auxiliaries with respect to sentential negation can also be simplified by applying the non-derivative finiteness distinction. Like the other syntactic differences between main verbs and auxiliaries in PDE, *do*-support in English negated clauses is a topic that has occupied generative linguists for fifty years. In a sense, all the analyses are refinements of Chomsky (1955, 1957); cf. Lasnik (1999: 98). The basic facts to be explained are quite familiar, as pointed out in Cormack and Smith (1998: 1; 2000: 50) and many others:

(31) Verbs and negation in English
a. John often snores.
b. *John not snores.
c. John did/will/must not snore.
d. *John snores not.

The data in (31) give rise to the following three questions (from Cormack and Smith 2000: 50):

Q1: Why is (31b) ungrammatical? In particular,
Q2: What accounts for the difference between *often* and *not* in (31a) vs. (31b)?
Q3: Why can an Aux or modal precede *not,* while a V cannot (31c) vs. (31d)?

One common way of answering these questions is to assume that there are two types of negation markers in Germanic languages. Type A is a head, projecting its own functional projection in the IP domain of the clause (cf. e.g., Zanuttini 1996). Type B is an adverb, with a distribution similar to other sentence adverbs. Whereas sentential negation in German, Dutch, and MSc is believed to be of type B (Zanuttini 1996: 191), English sentential negation is assumed to be type A, giving rise to a designated NEGP projection. This partially answers question Q2 since *not* is not an adverb. This type of approach would also typically answer question Q1 by means of the Head Movement Constraint, HMC (Travis 1984: 131), which restricts movement of a head to the nearest head position. A syntactic head X^0 cannot move over a head Y^0 to reach a head position Z^0 (this is a subcase of Relativized Minimality; cf. Rizzi 1990a). Under the assumption that the inflection (or the relevant head checking the inflection) of *snores* is generated in a position above negation, whereas the (bare) verb is merged below negation, the negation head would intervene between the main verb and (the head checking) its inflection. Thus, the HMC accounts for the ungrammaticality of (31b) and (31d). In HMC-based approaches, auxiliaries are generated above negation (for some reason) and are thus not restricted by it the same way verbs are; this answers question Q3.

In contrast to these and other analyses where elaborate mechanisms are needed to account for the facts, the present approach can account for the facts in (31) with a single assumption:

> Sentential negation in Germanic requires a [+Finite] verb.

Admittedly, this requires a restrictive definition of sentential negation, excluding the type of negation we find in infinitives, ECM constructions, and subjunctives. I will adopt the definition used in Schütze (2004: 11), who, following "quite common practice" refers to "the head in which sentential negation *not* is generated" as Σ (Gleitman 1965; Laka 1990).

> In addition to *not/n't,* Σ can host overt expressions of positive polarity, as in *John does TOO/SO know Arabic,* and a segmentally empty morpheme that induces prosodic emphasis, call it $Ø_{emph}$, as in *Mary DOES like pineapple!* All these items trigger *do*-support in English, the only value of Σ that does not is nonemphatic positive polarity $Ø_{pos}$.

Note, however, that *do*-support is attested even with non-emphatic, positive polarity Σ in certain "modern dialects and registers and in child English"; cf. Schütze (2004), and, most famously, in the English of the 1500s. According to Denison (1993: 264–5), *do*-support is attested in positive declaratives from the thirteenth century and shows up in negative declaratives a little later (starting in the fourteenth century).[37]

No matter what mechanism we choose, observing the non-derivative finiteness distinction substantially simplifies our analysis of sentential negation. For concreteness, let us assume within the crash-proof/survive framework that sentential negation in English (i.e., negative Σ) has as a concatenative requirement an active [+Finite] feature. Given the discussion in the previous sections, my assumption that sentential negation in Germanic requires a [+Finite] verb rules out the possibility that PDE main verbs or non-finite *have* and *be* can satisfy the concatenative requirement of sentential negation. It also explains why modals and finite *have* and *be* occur with sentential negation in English, explains *do*-support in sentential negation structures (in a manner reminiscent of the solution in Syntactic Structures and Lasnik's 1981 "stranded affix filter"), and explains why any verb in MSc can fulfil the requirements of sentential negation. Even in MSc, Σ negation is subject to compatibility restrictions, as suggested for various adverbs by Ernst (2002). And one of these restrictions is the generalization that sentential negation in Germanic requires a [+Finite] verb.

5. Conclusion

I have attempted to show how temporal chains are construed in a syntactic structure, as implemented in a crash-proof syntax of the type envisaged in Frampton and Gutmann (2002), Putnam (2007), and Stroik (2009). I argued that the links in T-chains are local T-heads, where every main verb and auxiliary brings its own tense package to the chain. As I explained, the semantic difference between finite and non-finite T-element consists in the choice of first argument; the speech event S (finite tense) or a preceding verbal event e (non-finite tense).

I pointed out that the finiteness distinction is not a derivative made up of tense or agreement features, but a primitive distinction in its own right expressed

37. If English has retained some traces of the old four-way verbal paradigm in the form of the auxiliaries *have* and *be* and if there are some verbs (the modals and auxiliary *do*) that are inherently [+Finite], wouldn't we expect the obligatory insertion of auxiliary *do* in all positive declarative sentences rather than a change to a completely different system in which finiteness is just a matter of structural hierarchy? The insertion of *do* in positive declaratives was the first step in this process; cf. Denison (1993: 265) and McWhorter (2007: 6). Whereas *do*-support spread to other constructions requiring [+Finite] verbs, it eventually dropped out of positive declaratives, in which it was first introduced, and is now found only in certain dialects, registers, and child language, cf. Schütze (2004).

in verbal paradigms independently of tense and agreement. This helps explain the syntactic behaviour of main verbs and auxiliaries in English and differences between types of English auxiliaries. Furthermore, the crucial difference between Modern English and Mainland Scandinavian languages (MSc) is in overt inflectional markings encoding finiteness. Middle English, like MSc, encoded finiteness with all types of verbs. MSc has retained the finiteness distinction, but English main verbs have lost it. However, English main verbs have retained their tense and agreement markings. This different development in English and MSc fuelled many syntactic differences, e.g., *do*-support versus verb-second. The loss of the finiteness distinction was also argued to be responsible for many major syntactic changes from Old English to Present Day English. Previous analyses had not isolated the relevant feature; in my analysis, this feature is the non-derivative finiteness distinction.

References

Adger, D. 2007. Three domains of finiteness: A minimalist perspective. In *Finiteness,* I. Nikolaeva (Ed.), 23–58. Oxford: OUP.

Akmajian, A. 1984. Sentence types and form-function fit. *Natural Language and Linguistic Theory* 2: 1–23.

Baptista, M. 1997. The morphosyntax of nominal and verbal categories in Capeverdean creole. Ph.D. dissertation, Harvard University.

Barbiers, S. 1995. The syntax of interpretation. Ph.D. dissertation, University of Leiden (Distributed by HAG (HIL), The Hague).

Barron, J. 2000. The morphosyntactic correlates of finiteness. In *Proceedings of the LFG00 conference*, M. Butt & T. Holloway King (Eds). http://csli-publications.stanford.edu/

Bisang, W. 2001. Finite vs. non finite languages. In *Language typology and language universals,* Vol. 2, M. Haspelmath, E. König, W. Oesterreicher & W. Raible (Eds), 1400–1413. Berlin: Mouton de Gruyter.

Blom, W.B.T 2003. *From root infinitive to finite sentence. The acquisition of verbal inflections and auxiliaries.* Utrecht: LOT Publications.

Bouchard, D. 1995. *The semantics of syntax*. Chicago IL: The University of Chicago Press.

Büring, D. 2007. Pronouns. To appear. In *Semantics: An international handbook of natural language meaning* [Handbücher zür Sprach- und Kommunikationswissenschaft/Handbooks of Linguistics and Communication Science], K. von Heusinger, C. Maienborn & P. Portner (Eds). Berlin: Mouton De Gruyter.

Bybee, J., Perkins, R. & Pagliuca, W. 1994. *The evolution of grammar-tense, aspect and modality in the languages of the world.* Chicago IL: The University of Chicago press.

Chomsky, N. 1955. The logical structure of linguistic theory. Ms, Harvard University (Revised 1956 version published in part by Plenum, New York NY, 1975; University of Chicago Press, Chicago IL, 1985).

Chomsky, N. 1957. *Syntactic structures.* The Hague: Mouton.

Chomsky, N. 1995. *The minimalist program.* Cambridge MA: The MIT Press.

Chomsky, N. 2001. Derivation by phase. In *Ken Hale. A life in language*, M. Kenstowicz (Ed), 1–52. Cambridge MA: The MIT Press.
Comrie, B. 1985. *Tense.* Cambridge: CUP.
Cormack, A. & Smith, N. 1998. Negation, polarity, and verb movement. *Working Papers in Linguistics* 10: 285–322. University College London.
Cormack, A. & Smith, N. 2000. Head movement and negation in English. *Transactions of The Philological Society* 98(1): 49–85.
Cowper, E. 2002. Finiteness. Ms, University of Toronto.
Davidson, D. 1967. The logical form of action sentences. In *The logic of decision and action*, N. Rescer (Ed.), 81–95. Pittsburgh PA: University of Pittsburgh Press.
Denison, D. 1993. *English historical syntax: Verbal constructions.* London: Longmans.
Eide, K.M. 2005. *Norwegian modals* [Studies in Generative Grammar 74]. Berlin: De Gruyter.
Eide, K.M. 2007. Finiteness and inflection: The syntax your morphology can afford. http://ling.auf.net/lingbuzz
Eide, K.M. 2009. Finiteness: The *haves* and the *have-nots*. In *Advances in Comparative Germanic Syntax* [Linguistik Aktuell/Linguistics Today 141], A. Alexiadou, J. Hankamer, T. McFadden, J. Nuger & F. Schäfer (Eds). Amsterdam: John Benjamins.
Eiskovits, E. 1987. Variation in the lexical verb in inner-Sydney English. *Australian Journal of Linguistics* 7: 1–24.
Enç, M. 1987. Anchoring conditions for tense. *Linguistic Inquiry* 18: 633–657.
Ernst, T. 2002. *The syntax of adjuncts.* Cambridge: CUP.
Faarlund, J.T., Lie, S. & Vannebo, K. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
Frampton, J. & Gutmann, S. 2002. Crash-proof syntax. In *Derivation and explanation in the minimalist program,* S.D. Epstein & T.D. Seely (Eds), 90–105. Oxford: Blackwell.
Giorgi, A. & Pianesi, F. 1997. *Tense and aspect: From semantics to morphosyntax.* Oxford: OUP.
Givón, T. 1990. *Syntax. A functional-typological introduction,* Vol. II. Amsterdam: John Benjamins.
Gleitman, L.R. 1965. Coordinating conjunctions in English. *Language* 41: 260– 293.
Gretsch, P. & Perdue, C. 2007. Finiteness in first and second language acquisition. In *Finiteness,* I. Nikolaeva (Ed.), 432–484. Oxford: OUP.
Grohmann, K. 2000. Null modals in Germanic (and Romance): Infinitival exclamatives. *Belgian Journal of Linguistics* 14: 43–61.
Guéron, J. & Hoekstra, T. 1995. The temporal interpretation of predication. In *Small clauses. Syntax and semantics,* A. Cardinaletti & M.T. Guasti (Eds), 77–107. New York NY: Academic Press.
Guéron, J. & Lecarme, J. 2004. *The syntax of time*. Cambridge, MA: The MIT Press.
von Heusinger, K. 2002. Reference and representation of pronouns. In *Pronouns: Representation and grammar* [Linguistik Aktuell/Linguistics Today 52], H. Simon & H. Wiese (Eds), 109–135. Amsterdam: John Benjamins.
Holmberg, A. & Platzack, C. 1995. *The role of inflection in Scandinavian syntax.* Oxford: OUP.
Hopper, P.J. 1975. *The syntax of the simple sentence in Proto-Germanic.* The Hague: Mouton.
Hornstein, N. 1990. *As time goes by-Tense and universal grammar.* Cambridge MA: The MIT Press.
Jackendoff, R. 1972. *Semantic interpretation in generative grammar.* Cambridge MA: The MIT Press.
Julien, M. 2001. The syntax of complex tenses. *The Linguistic Review* 18: 125– 167.
Julien, M. 2003. Dobbelsupinum og irreal modus. *Norsk Lingvistisk Tidsskrift* 21: 135–161.
Kobele, G. 2006. Generating copies: An investigation into structural identity in language and grammar. Ph.D. dissertation, University of California Los Angeles.

Koptjevskaja-Tamm, M. 1994. Finiteness. In *The encyclopedia of language and linguistics,* R.E. Asher (Ed.), 1245–1248. Oxford: Pergamon.

Laka, I. 1990. Negation in syntax: On the nature of functional categories and projections. Ph.D. dissertation, MIT (published as *On the syntax of negation*, 1994, New York NY: Garland).

Landau, I. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22: 811–877.

Lasnik, H. 1981. Restricting the theory of transformations: A case study. In *Explanations in Linguistics,* N. Hornstein & D. Lightfoot (Eds), 152–173. London: Longman.

Lasnik, H. 1999. *Minimalist analysis.* Oxford: Blackwell.

Lasnik, H. 2000. *Syntactic structures revisited. Contemporary lectures on classic transformational theory.* Cambridge MA: The MIT press.

Lasser, I. 1997. *Finiteness in adult and child German* [MPI Series in Psycholinguistics 8]. Wageningen: Ponsen and Looijen.

Lightfoot, D. 2006. *How new languages emerge.* Cambridge: CUP.

Luhtala, A. 2000. *On the origin of syntactical description in stoic logic.* Münster: Nordus.

McWhorter, J. 2005. *Defining creole.* Oxford: OUP.

McWhorter, J. 2007. What else happened to English? A brief for the Celtic hypothesis. Ms, to be published in *English Language and Linguistics.*

Muysken, P. & Law, P. 2001. Creole studies. A theoretical linguist's field guide. *GLOT International* 5(2): 47–57.

Nikolaeva, I. 2007. Introduction. In *Finiteness,* I. Nikolaeva (Ed.), 1–19. Oxford: OUP.

Partee, B.H. 1973. Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy* 70: 601–609.

Partee, B.H. 1984. Nominal and temporal anaphora. *Linguistics and Philosophy* 7: 243–286.

Pinker, S. 2000. *Words and rules.* London: Phoenix.

Platzack, C. 1998. The initial hypothesis of syntax: A minimalist perspective on language acquisition and attrition. In *Generative perspectives on language acquisition* [Language Acquisition & Language Disorders 10], H. Clahsen (Ed.), 369–414. Amsterdam: John Benjamins.

Pollock, J.Y. 1989. Verb movement, UG and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Putnam, M.T. 2006. Eliminating agree. Ms, Michigan State University.

Putnam, M.T. 2007. *Scrambling and the survive principle.* Amsterdam: John Benjamins.

Reinhart, T. & Reuland. E. 1993. Reflexivity. *Linguistic Inquiry* 24(4): 657– 720.

Reuland. E. & Reinhart, T. 1995. Pronouns, anaphors and case. In *Studies in Comparative Germanic Syntax,* H. Haider, S. Olsen & S. Vikner (Eds), 241–268. Dordrecht: Kluwer.

Ritter, E. & Wiltschko, M. 2005. Anchoring events to utterances without tense. WCFLL 2005, http://www.linguistics.ubc.ca/People/martina/papers/wccfl_ritter_wiltschko.pdf

Rizzi, L. 1990a. *Relativized minimality*. Cambridge MA: The MIT Press.

Rizzi, L. 1990b. On the anaphor agreement effect. *Rivista di Linguistica* 2: 27–42.

Roberts, I. 1985. Agreement parameters and the development of English modal auxiliaries. *Natural Language and Linguistic Theory* 3: 21–58.

Roberts, I. 1993. *Verbs and diachronic syntax.* Dordrecht: Kluwer.

Roberts, I. 2007. *Diachronic syntax.* Oxford: OUP.

Romaine, S. 1993. *Pidgin and creole languages*. London: Longman.

Ross, J.R. 1969. Auxiliaries as main verbs. In *Studies in philosophical linguistics*, W. Todd (Ed.), 77–102. Evanston IL: Great Expectations.

Ross, J.R. 1970. On declarative sentences. In *Readings in English transformational grammar*, R.A. Jacobs & P.S. Rosenbaum (Eds), 222–277. Waltham MA: Ginn and Company.

Sampson, G. 2002. Regional variation in the English verb qualifier system. *English Language and Linguistics* 6: 17–30.

Schlenker, P. 1999. Propositional attitudes and indexicality. A cross- categorial approach. Ph.D. dissertation, MIT.

Schütze, C.T. 2003. When is a verb not a verb? In *Proceedings of the 19th Scandinavian Conference on Linguistics, Nordlyd,* A. Dahl, K. Bentzen & P. Svenonius (Eds), 312: 400–415. Tromsø: University of Tromsø.

Schütze, C.T. 2004. Synchronic and diachronic microvariation in English *do*. *Lingua* 114(4): 495–516.

Siegel, M. 1987. Compositionality, case, and the scope of auxiliaries. *Linguistics and Philosophy* 10: 53–75.

Solà, J. 1996. Morphology and word order in Germanic languages. In *Minimal ideas, Syntactic studies in the minimalist framework* [Linguistik Aktuell/Linguistics Today 12], W. Abraham, S. Epstein, H. Thráinsson & C.J.-W. Zwart (Eds), 217–251. Amsterdam: John Benjamins.

Stowell, T. 1995. What do the present and past tenses mean? In *Temporal reference, aspect and actionality,* Vol. 1*, Semantic and syntactic perspectives,* P.M. Bertinetto, V. Bianchi, J. Higginbotham & M. Squartini (Eds), 381–396. Torino: Rosenberg and Sellier.

Stowell, T. 2007. The syntactic expression of tense. *Lingua* 117: 437–463.

Stroik, T. 1999. The Survive principle. *Linguistic Analysis* 29: 278–303.

Stroik, T. 2009. *Locality in minimalist syntax.* Cambridge MA: The MIT Press.

Stroik, T. & Putnam M.T. 2005. The lexicon at the interfaces. Paper presented at LASSO 35. Lubbock, Texas.

Thráinsson, H. 1996. On the non-universality of functional categories. In *Minimal ideas. Syntactic studies in the minimalist framework* [Linguistik Aktuell/Linguistics Today 12]. Amsterdam: John Benjamins.

Travis, L. 1984. Parameters and effects of word order variation, Ph.D. dissertation, MIT.

Trudgill P. & Chalmers, J.K. (Eds). 1991. *Dialects of English: Studies in grammatical variation.* London: Longman.

Vikner, S. 1985. Reichenbach revisited: One, two, or three temporal relations. *Acta Linguistica Hafniensia* 19: 81–98.

Woolford, E. 1999. More on the anaphor agreement effect. *Linguistic Inquiry* 30: 257–287.

Wurmbrand, S. 2007. Infinitives are tenseless. *U. Penn Working Papers in Linguistics* 13(1).

Zanuttini, R. 1996. On the relevance of tense for sentential negation. In *Parameters and functional heads*: *Essays in comparative syntax*, A. Belletti & L. Rizzi (Eds), 181–207. Oxford: OUP.

Zubizarreta, M.L. 1982. On the relationship of the lexicon to syntax. Ph.D. dissertation, MIT.

When grammars collide

Code-switching in Survive-minimalism*

Michael T. Putnam & M. Carmen Parafita Couto
Carson-Newman College/ESRC Centre for Research on Bilingualism, Bangor University

This study provides a Survive-minimalist analysis of two issues related to DP-structures in code-switching (CS) grammars: (i) the relationship between determiners and nouns in a DP where each respective lexical item originates from a separate language and (ii) the linearization of **Det**(erminers)-**Adj**(ectives)-**N**(ouns) in CS-grammars where each respective language contributing a surface order contrasts with the other. Violable constraints that filter the selection possibilities (i.e., the operation ***Select***) of determiners are posited. We contend that a formal feature, definiteness [+ Def], triggers the re-configuration of lexical items to conform to structural requires of a given CS-grammar. That same feature motivates both the DET-ADJ-N and the DET-N-ADJ orderings. The advantages to pursuing this analysis of DP linearity in CS-grammars are that it is: (i) consistent with the desiderata of Survive-minimalism and (ii) does not require features similar to the EPP to exist in the system.

1. Introduction

Since the focus of generative grammar has shifted to the mental processes which underlie human behavior in the context of the "cognitive revolution", the purpose of linguistic analysis has been to explain the knowledge or competence that a native speaker has (Chan 2003). This competence, which is argued to be an innate faculty, can be found in any human being regardless of his/her native language.

As Leung (2001) puts it, generative studies in code-switching (hereafter CS) investigate the strategy bilinguals adopt in order to attain a higher level of

*We are thankful to constructive and critical comments from Vicki Carstens, Margaret Deuchar, John Hale, Dorian Roehrs, Tom Stroik, John te Velde, and Elly van Gelderen. Our paper also benefited from the comments from an anonymous reviewer who undoubtedly strengthened the arguments in this paper as well as helped in clarifying key passages. All remaining shortcomings are our own.

optimality, and which tactics should be operative across languages and individuals. The switch of lexical items could obviate numerous algorithmic complexities which speakers might otherwise come across. Leung (2001) claims that CS provides a fresh way to solve the puzzle of determining the level of optimality that bilinguals maintain in CS-structures.

Leung (2001) contends that the simple and minimal computation involved in the formation of CS as an instance of synchronic variation shows clearly what is really indispensable to language, i.e., language with fewer imperfections. CS is in some sense parallel with individual natural languages in terms of language competence and design. Thus, a complete theory of language should be able to put CS at it disposal. This study is a test of the feasibility of one formal framework that adheres to the main principles and tenets of the Minimalist Program: Survive-minimalism. We want to see, to what extent, Survive-minimalism improves our understanding of Minimalist grammars and their relationship to understanding/analyzing code-switching data.

In this study we will examine the syntactic properties of the mixed determiner phrases, i.e., DPs that contain lexical items from more than one source language. In particular we will look at the variation regarding syntactic derivations that relate to the determiner and the adjective, and the associated feature checking operations (Moro in press, Deuchar et al. 2008) that derive these DPs, including the linear order of the constituents that belong to these mixed DPs. We will see that there are advantages as well as potential problems in the theory of formal feature strength, and these will be discussed along with some implications in the context of the principles of Survive-minimalism. We will then consider the main theoretical foundations of Survive-minimalism, which are critical for the analysis that we develop, and an overview of the core data of mixed DPs in this article. In what follows, therefore, we will sketch the main theoretical foundations on which this study is based.

The paper is organized as follows. Section 2 presents a review of existing work on intra-sentential language mixing by bilingual speakers, noting that there have been two main theoretical approaches to which researchers adhere: a constraint-oriented approach and a constraint-free approach (MacSwan 2008). In Section 3 we shift into the main focus of the paper (the DP) and look at recent proposals regarding its internal structure. In Section 4 we explore the Survive Principle in its historical dimension and discuss what new challenges arise from adopting Survive-minimalism in our treatment of code-switching. This leads us to our exploration of lexicon-syntax mapping issues discussed in sections 5 and 6. Following our treatment of lexicon-syntax mapping issues, we turn to Moro's (in press) analysis of code-switched DPs and formulate a Survive-based analysis of the data. Section 7 summarizes our main claims and acknowledges that although Survive-minimalism, as one of the many instantiations of the Minimalist Program (MP), may face some of the same problems that mainstream Minimalism does, it also simultaneously has distinct advantages in delivering a more parsimonious analysis of the CS-data highlighted in this study.

2. Minimalist grammars and CS

Currently, there are relatively few studies dedicated to the study and analysis of CS from a generative perspective (cf. Chan (2003), and MacSwan (1999, 2008) to name a few). In this section we introduce some of the more widely-held assumptions about CS in the MP. As a point of departure, let's clarify what we mean by CS: CS is the alternating use of two (or more) languages within the same utterance, as illustrated in (1):[1]

(1) a. I'm fine. *¿Tú* *cómo* *estás*?
PRON.2S how be.PRES.2S
I'm fine. How are you?

b. I visit *mi* *abuelo* on the weekends. (Ghirardini 2006)
POSS grandfather
I visit my grandfather on the weekends.

The type of CS shown in (1b), in which an alternation happens within sentential boundaries, is known as ***intra-sentential CS***, whereas switching between sentences, as in (1a) is known as ***inter-sentential CS***. Because grammatical theory is primarily focused on relations below the sentence level, this paper will focus on intra-sentential CS.

CS research has developed into two distinct research programs (MacSwan 2008): the constraint-oriented program and the constraint-free program. Following MacSwan (2008), code-switching is constrained in the descriptive sense since some patterns are well-formed and others are not:

(2) My big brother *va* *a* *comer* *con* *nosotros.*
go.PRES.3S to eat.NFIN with PRON.2PL
My big brother is going to eat with us.

(3) *He *va* *a* *comer* *con* *nosotros.*
go.PRES.3S to eat.NFIN with PRON.2PL
He is going to eat with us.

1. Key to glosses: PRON=pronoun; POSS=possessive pronoun; DET=determiner; NFIN=non-finite; PRES=present; 1S/1PL=1st person singular/plural; 2S/PL=2nd person singular/plural; 3S/3PL=3rd person singular/plural.

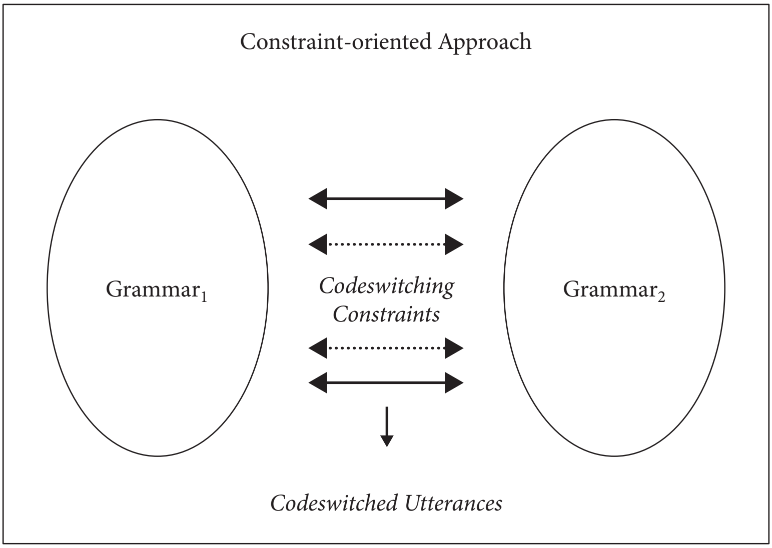


Diagram 1. Constraint-oriented Approach (MacSwan 2008)

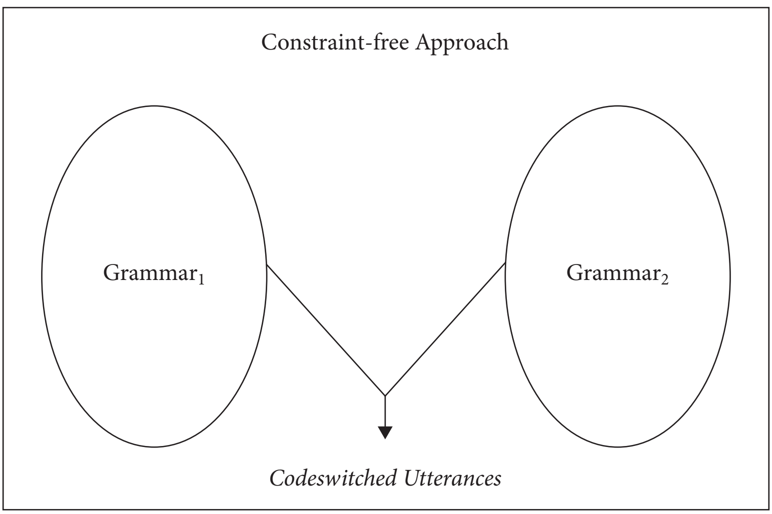


Diagram 2. Constraint-free Approach (MacSwan 2008)

On the other hand, in the theoretical sense, a constraint is a principle or rule of grammar that bars one or another structure.

The constraint-free approach claims that nothing constrains CS apart from the requirements of the mixed grammars. Thus, CS is the union of two grammars, with no regulatory grammatical mechanisms permitted (MacSwan 2008). This search for a constraint-free explanation can be traced back to seminal research by scholars such as Pfaff (1979), Poplack and Sankoff (1984), Woolford (1983), Lipski (1985), Di Sciullo, Muysken and Singh (1986), Clyne (1987), Belazi, Rubin and Toribio (1994), Santorini and Mahootion (1995), MacSwan (1997), Chan (2003). Among those, a variety of non-linear approaches, especially those based on Chomsky's Principles-and-Parameters system were proposed, drawing on, for example, the Government-and-Binding (GB hereafter) framework (Woolford, 1983; Disciullo, Musyken, and Singh, 1986; Halmari, 1997), the Functional Head Constraint (Belazi, Rubin, & Toribio 1994), the Null Hypothesis (Mahootian, 1993), the Minimalist approach (MacSwan 1997).

Researchers in code-switching have acknowledged the preference for a constraint-free solution. The central problem in such approaches, as pointed out by MacSwan (2008), is late lexical insertion. In the MP, lexical insertion occurs at the onset, and language-particular properties are carried along with each lexical item into the syntax. Thus, the problem of late lexical insertion disappears and there is no need to match languages up with their respective phrase structure rules later on. In the MP (Chomsky 1995 et seq.), all parameters are encoded in the lexicon, with the consequence that linguistic variation falls out from the morphological properties of the lexical items (Borer 1984). In this model, there are two central components: A computational system for human language, which is presumed to be invariant across languages, and a lexicon, to which the idiosyncratic differences observed across languages are attributed.

In the MP phrase structure is derived from the lexicon, projecting a Universal Base. Several operations are in place: ***Select*** picks items from the lexicon and introduces them into the ***Lexical Array***, which is a subset of the lexicon that is then used to build a derivation. ***Merge*** takes items from this array and produces hierarchically organized branching substructures. ***Move*** (also known as ***Internal Merge*** (Chomsky 2001)) raises elements to higher positions within the tree. At the point of Spell Out, the representation splits. At this point features relevant to PF (Phonetic Form) are handed over to the phonological component, and features relevant to LF (Logical Form) are handed over to semantic interpretation. The derivation will converge if all the appropriate features are valued and deleted before the interface levels. If all syntactic variation is associated with the lexicon, as in the MP, then CS may be seen as the simple consequence of mixing items from multiple lexicons in the course of a derivation. MacSwan (1999, 2000, 2005a, 2005b, and in press) develops a model of CS in which items may be drawn from the lexicon of multiple (discretely represented) languages (two, in the simple case) to introduce features into the ***Lexical Array***, which must then be checked for convergence in the same way as monolingual features must be checked, with no CS-specific mechanisms permitted:

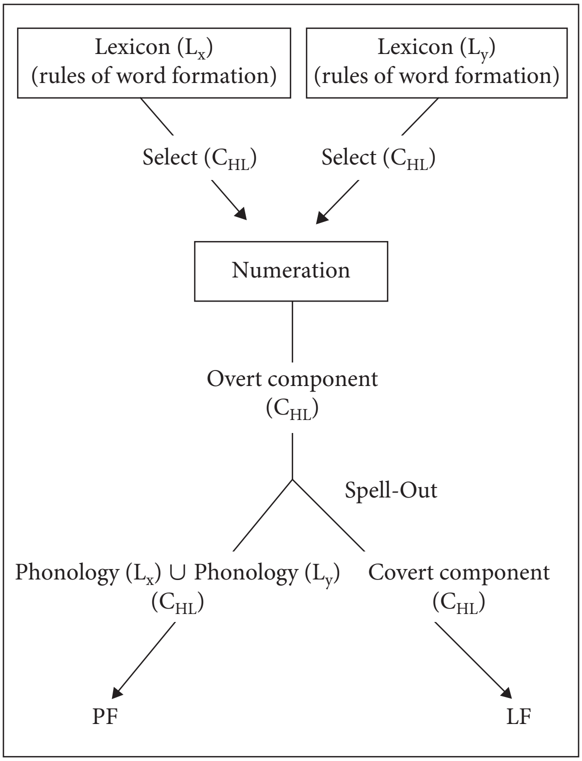


Diagram 3. The bilingual language faculty (MacSwan 2000)

In MacSwan's model, the grammatical requirements are simply carried along with the lexical items of the respective systems. On this view, the relevant theory of CS is stated as in (4):

(4) Nothing constrains CS apart from the requirements of the mixed grammars (MacSwan, 1999).

Note that (4) is a theory of CS, and not part of the theory of grammar. Indeed, (4) implies that no rule or principle of grammar may make explicit reference to CS, nor to the identities of particular languages. This approach to CS is explicated in greater detail in MacSwan (1999, 2005a, 2005b, in press), where CS-specific mechanisms and constraint-oriented approaches to CS are rejected.

2.1 Enter survive

Recapitulating the discussion in the introduction of this volume on the MP (Putnam and Stroik, this volume), it is important to recall that Minimalism represents an entire family of theoretical approaches that share a finite number of core assumptions and guiding intuitions rather than a rigid set of pre-established desiderata. Concomitant with the core principles of Minimalism, any variant of the MP that embraces these aforementioned criteria can be classified as a "minimalist" formalism. As such, our theoretical discussions in this article focus on a version of the MP that employs the Survive Principle (Stroik 1999, 2000, 2008 and further developed in Putnam 2007) as its chief mechanism of constituent "movement"[2] in the narrow syntax and the "pros" and "cons" that a Minimalist grammar encounters in traditional analyses of CS. For the sake of space and to eliminate redundancy, we will spare the reader an in depth discussion of these previously discussed issues. Recall that the Survive Principle essentially functions, in the words of Preminger (2008, this volume) and Lechner (this volume), as a "push-chain" theory of syntactic movement. Stroik (1999: 286) formulates the original Survive Principle as such (but see Lechner (this volume) and Putnam (2007) for adjustments to this principle):

(4) **The Survive Principle**
If Y is a syntactic object (SO) in an XP headed by X, and Y has an unchecked feature [+F] that is incompatible with the feature X, Y remains active in the Numeration.

Accordingly, Survive-minimalism does not advocate the existence of Derivation-to-Derivation mapping licensed by ***Internal Merge/Move*** as found in other stripes of Minimalist theory; Survive-minimalism only utilizes the operations ***Merge*** and ***Remerge*** (which in actuality can be reduced to the same operation) that enact Numeration-to-Derivation concatenations, thus forming syntactic objects that are interpretable at the external interfaces from individual lexical items. As a result, all instances of ***(Re)Merge*** in Survive-minimalism are what we would classify as ***External Merge*** in most standard applications of the MP. Putting aside the similarities that Survive-minimalism shares with other flavors of the MP, there are nonetheless significant ways where Survive-minimalism radically departs from standard Minimalist assumptions that will ultimately have a profound impact on our treatment of conceptual issues and empirical data surrounding CS. The first, and perhaps most important area, where the Survive-minimalist system distances itself from orthodox Minimalism is in its treatment of the relation between the ***Lexicon*** and the narrow syntax. The standard/traditional approach, as mentioned above, is that an operation ***Select*** probes the ***Lexicon***, selects those and only those elements that will be relevant for an upcoming syntactic derivation, and finally

2. It bears repeating that in the Survive-minimalist framework developed by Putnam (2007), Stroik (2009) and Putnam and Stroik (to appear), all instances of what is commonly referred to as "movement" or "displacement" in transformational generative grammar, is regulated by ***External Merge***, i.e., ***(Re)Merge***.

places these few lexical items into the ***Lexical Array***, where they can be accessed in the course of the derivation. The operation ***Select*** is, under these assumptions, quite powerful and mathematically complex. As a result, the iterative application of ***Select*** within this framework has the potential to generate excessively complex, i.e., "wacky", **Lexical Arrays** (a.k.a. ***Numerations***). Stroik (2009, this volume) argues for a direct line of access between the ***Lexicon*** and the narrow syntax without the assistance of an intermediary ***Select*** operation. In eliminating ***Select***, a clear distinction between multiple (but separate) lexicons (MacSwan 1997 and later works) can no longer be upheld. From a preliminary standpoint, it is unclear how the syntax can correctly access separate lexicons – or even a singular lexicon with separate sub-components – in the absence of some sort of ***Select*** operation, a situation that we will address in the following section. Second, Survive-minimalism maintains a strict mapping algorithm between the derivation and phonological realization at PF (in some ways, quite similar in design and function to Kayne's (1994) Linear Correspondence Axiom). This is a particularly poignant issue in any discussion of CS data, especially when certain linear orders are preferred over others. For example, in either a Minimalist formalism that makes use of ***Select*** and one that does not, these issues are not easily avoidable.

(5) $\Gamma = \{\alpha, \{\alpha, \beta \dots\}\}$

Here α and β represent syntactic objects (SOs). Assume that α is the label in this string as well. Linear ordering – being a representational condition imposed on a grammatical string by the phonological component (PF) external to the narrow syntax – properly orders, or perhaps more appropriately *linearizes*, the syntactic objects in a successfully derived grammatical string. This operation, for lack of a better term, which we label as ***Linearize***, maps hierarchical relations onto linear ones. The same is true if we merge the independently constructed SO γ as well, creating the following more complex SO to serve as the input producing a properly-linearized string (note that in the following SO α is once again functioning as the label):

(6) $\Delta = \{\alpha, \{\gamma, \{\alpha, \{\beta \dots\}\}\}\}$

Successive applications of ***Merge*** and ***Remerge*** produce a derivationally-construed c-command (in the sense of Epstein (2001)) that holds among all and only the syntactic objects merged. As mentioned by Hinzen (2006: 247), a natural suggestion would then seem to be that the linear ordering will somehow *mirror* the way that the command relations came about derivationally. Hence, tracing the derivational order backwards, in the examples above, the ordering will be that γ precedes α, α precedes β, and β will precede whatever terms it contains: that is, $\langle\gamma, \alpha, \beta \dots\rangle$. A welcome outcome is when/if ***Linearize*** can exploit the hierarchical syntactic structure already produced during the course of the derivation. The game changes dramatically, however, once we attempt to look at CS data and apply some version of ***Linearize***. Consider the following data from Bokamba (1989), Stenson (1991), and Woolford (1983), as cited in Moro (in press):

(7) He presented a **paper exceptionnel** ('exceptional').
(English-French; Bokamba, 1989:282, (16a))

(8) Ta **carr light green** aige.
he car at-him
He has a light green car.
(English-Irish; Stenson, 1991:171, (7a))

(9) The big casa
DET big house
The big house.
(Quintero, cited in Woolford 1983)

(10) *El hombre old
DET man
The old man.
(Gingrás, cited in Woolford 1983)

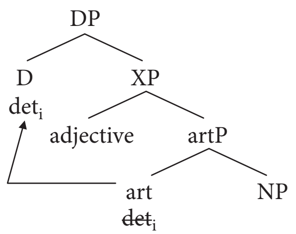
As demonstrated in the data above (7) – (10), the proper order of DET-ADJ-N inside the DP of CS-grammars presents a daunting challenge to any grammar formalism. Further suggesting that CS-grammars are natural grammars is the fact that certain combinations (cf. (10)) are judged to be ungrammatical. From a preliminary standpoint, when comparing examples (9) and (10), what appears to determine the linear order of DP-internal units is the language from which the adjective hails. As a matter of fact, Moro (in press) makes this claim explicitly; however, as will be discussed below in Section 3, evidence suggests that adjectives are base-generated in the specifier positions of AgrPs and are thus not the heads of a functional phrase. This jeopardizes the ability of adjectives to license/determine the linear alignment of DP-internal constituents. This, of course, poses a challenge for not only Survive-minimalism, but also for most variations of the MP, which is a point we address in detail in **Section 5**.

3. Internal structure of DP

Similar to clause-level structure, there is some disagreement with regard to the kinds, number, and sequence of phrases founded in the DP-middle field. Following, Julien (2002:287; 2005) and Roehrs (2006:16), we adopt the following hierarchical representation, moving top-down, which consists of a Determiner Phrase (DP), a Cardinal Number Phrase (CardP), an Agreement Phrase (AgrP for Roehrs/αP for Julien), an Article Phrase (as proposed by Roehrs (2006)), a Number Phrase (NumP), a light noun Phrase (*n*P) and NP.

(11) $[_{DP}$ D $[_{CardP}$ Card $[_{AgrP}$ Agr $[_{artP}$ art $[_{NumP}$ Num $[_{nP}$ n $[_{NP}$ N]]]]]]

Roehrs (2006) – based on a long tradition of similar proposals (cf. Brugé 1996, 2002; Campbell 1996; Carstens 1991; Giusti 1997, 2002; Grohmann and Panagiotidis 2004; Panagiotidis 2002; Vangsnes 1999:119–20 among others) – champions an analysis where the relationship between demonstratives and (in-) definite articles and adjectives determiners (i.e., demonstratives and (in-) definite articles) are base-generated in a position below all adjectives but above the head noun and its complements prior to being moved to the DP. Simplifying things for a moment, Roehrs' general claim can be represented as follows (adapted from Roehrs 2006:10):

(12)


Returning to the motivation of the existence of artP, Roehrs turns to the close relationship that exists between adjectives and determiners. To set the stage, note that semantically vacuous determiners can be optional in unmodified predicate noun phrases or with proper name:

(13) a. Sie ist (eine) Lehrerin. **German**
she is a teacher
'She is a teacher.'

b. (Der) Hans ist gestern hingefallen.
the Hans is yesterday down-fallen
'Hans fell down yesterday.'

However, when modified, the determiner becomes obligatory (the adjective *arm* 'poor' in (14b) must be interpreted with a "non-restricted" meaning, cf. Gallmann 1997:75):

(14) a. Sie ist *(eine) gute Lehrerin. **German**
she is a good teacher
'She is a good teacher.'

b. {Der arme Hans/ *Armer Hans} ist gestern hingefallen
the poor Hans/ poor Hans is yesterday down-fallen
'Poor Hans fell down yesterday.'

The data in (13) and (14) clarify the relationship between determiners and adjectives: while the presence of the adjective is parasitic on the position of the determiner, the determiner can appear independently of whether an adjective is present or not. Bearing in mind that these determiners are semantically vacuous, their obligatory presence with an adjective (or more generally, a modifier, cf. Roehrs 2006: Chapter 2) should be captured in a syntactic way. This represents the core of Roehrs' claim where he suggests how this unidirectional relationship between determiners and adjectives can be captured in a straightforward manner according to Minimalist assumptions. As represented in the structure (12) above, Roehrs assumes the local obligatory selection of artP by Agr.[3] Before moving on to a discussion of how this will be implemented into our current study, a few words need to be said about the syntactic status of adjectives. Following Cinque (1994) and Roehrs (2006) among others, we assume that AgrP hosts an adjective phrase in its Specifier position and that it does not exist as the head of AgrP. Taking AgrPs to be recurring (in the sense of Cinque 1999, Scott 2002), each phrase contains one adjective in its Specifier. To provide evidence for this assertion, consider the following data from Haider (1992:320) (cited in Roehrs 2006:20–21) who notes that modifiers such as *sehr* 'very' cannot be stranded in split-NPs:

(15) a. [Kostbaren Vasen] besitzt er nur [drei *(sehr)]. **German**
precious vases owns he only three very

b. [Sehr kostbaren Vasen] besitzt er nur [drei].
very precious vases owns he only three
'He only owns three very precious vases.'

The data in (15) illustrates why the argument that the modifier itself is a head of the extended projection of the noun cannot be upheld. The fact that the modifier cannot be stranded as in (15a) follows directly from the assumption that the modifier is in the Specifier position of AP and the intermediate A'-position may not be moved.[4]

Roehrs' analysis of determiners being base-generated in the Specifier position of artP also provides a straightforward explanation of the parametric variation of

3. Roehrs also notes that if determiners are generated just below adjectives, as suggested by the internal structure of DPs that we also adopt here (cf. (6)), it is not unexpected that demonstratives are similar to adjectives in some languages.

4. In fact, Corver (1991, 1997) and Zamparelli (1993, 2000: chap. 7) argue that the modifier is part of the extended projection of the adjective.

the placement of determiners as either pre- and/or post-nominal forms. Consider the following Icelandic data (Roehrs 2006:60):

(16) gamli maður-inn **Icelandic**
old man-the
'the old man'

In languages like English and German that display the **Det**(miner)-**Adj**(ective)-**N**(oun) ordering, the determiner, base-generated in art, raises to D. In languages that display the DET-N-ADJ order such as Icelandic (cf. (16)), the aforementioned linear order is easy to derive if we assume that in languages like Icelandic the determiner does not move on its own, but rather the entire artP raises to Spec,DP, thus resulting in the DET-N-ADJ linear order. The derivational steps in (17) (taken from Roehrs 2006:60) illustrate the critical steps in this derivational history:

(17) a. $[_{DP}$ $[_{AgrP}$ **gmali** $[_{artP}$ **–inn** $[_{NP}$ **maður**$]]]_j$ D … $t_j]$
b. $[_{DP}$ $[_{AgrP}$ **gamli** $[_{artP}$ **maður**$_k$**+inn** $[_{NP}$ maður$_k]]]_j$ D … $t_j]$

As will be elaborated on further in **Sections 4** and **5**, Roehrs' internal structure of the DP provides an ideal platform from which we can launch our analyses of DP-internal code-switching data from a Minimalist perspective.

4. Lexicon-syntax mapping issues

As stated above, Survive-minimalism significantly departs from mainstream approaches to Minimalist linguistic inquiry in arguing for a direct route of access between the ***Lexicon*** and the narrow syntax. At first glance, this results in the elimination of the operation ***Select*** as it is found in mainstream versions of the MP. In a Survive-minimalist analysis of CS data; however, we must re-examine whether or not ***Select*** must remain in the theory based on the principle of virtual conceptual necessity. Much of this depends on our interpretation of the ***Lexicon*** in this model, i.e., is there one united lexicon in the grammar of CS speakers or, as suggested by MacSwan in subsequent works, are there two separate lexicons that exist in the mind of the speaker (e.g., **Lexicon X** and **Lexicon Y**) that remain distinct from one another? To begin, let's assume the null hypothesis; namely, that there exists only one unified lexicon for both languages. An immediate problem arises from this hypothesis in determining syntactic objects that will serve as specifiers and complements to a related head. For example, in English-Spanish CS, if a bilingual speaker wanted to say "the red balloon", what would prohibit her from inserting either *red* or *rojo* into the larger DP structure? Cross-linguistically it is probable, if not highly anticipated, that a universal constraint that would require CS-grammars to select a lexical item in a unified lexicon from one language (L_1) or another (L_2)

cannot be constructed. If we explore the alternative, namely, if separate lexicons exist for separate both languages in a bilingual grammar, we also fail to receive a parsimonious solution to our inquiry. Similar to generative research on second language acquisition, White (1993) among a host of others claims that aspects of UG are ever-present in the grammar of L2-grammars; however, regardless if this is true or not, Gass and Selinker (2007) explicitly show how difficult – if not impossible – it is to determine the underlying representation, i.e., the state of the L2-grammar, of a non-native speaker. Generative studies in CS face a very similar challenge in that rather than speaking, in most cases, in terms of a clear binary grammatical vs. ungrammatical distinction, bilingual grammars exhibiting CS forms must be analysed in probabilistic models. This relates not only to any discussion of the nature of the CS-grammar, but also to the structure and relationship of the lexicon(s). Survive-minimalism addresses this problematic issue in a novel way that has the potential of avoiding some of the pitfalls of mainstream Minimalism. The primary function of the narrow syntax in this model is to ensure the concatenate integrity of individual lexical items as they participate in the structure-building operations ***Merge*** and ***Remerge*** in the narrow syntax. Following these core assumptions, syntactic objects produce licit, interface-interpretable units as concatenated forms.

Data elicited for CS-studies support the principle of concatenate integrity championed in Survive-minimalism, thus illustrating that the framework may indeed make interesting claims about such natural data while improving the elegance and explanatory adequacy of the conceptual claims involved. Here we will investigate the structure of DPs, with a particular focus on the agreement properties of Spanish-English (i.e., *Spanglish)* and Spanish-German (i.e., *Esplugish*) CS-data. In our treatment of these sets of data according to Survive-minimalist desiderata, we in particular seek to provide answers to the following questions:

– If both participating languages have a grammatical gender system (like Spanish and German), how does gender agreement work in CS-DPs?
– If one participating language has a grammatical gender system (like Spanish) and the other one does not (like English), how does gender agreement work in CS-data?
– How does this CS-language deal with neuter gender in German, a gender which Spanish does not have?

In the discussion below, we attempt to provide insight on how a Survive-minimalist account of natural CS-data is both empirically and conceptually superior to other Minimalist models with regard to clarifying the lexicon(s) – narrow syntax relations. As will be discussed and elaborated on below, we advocate the existence of violable constraints to mediate the selection process from the lexicon(s). Selecting the correct lexical items (or, more appropriately, the statistically preferred units) is contingent upon features that exist on the head noun (N) and the determiner (D) and the ranked preference (based on the harmonic alignment of violable constraints) of the CS-grammar with regard to agreement phenomenon. Within such an approach, some sort of ***Select*** operation (in the form of constraint-ranking and evaluation) must be included into a Survive-minimalist account of CS. This, of course, comes with the advantage of obviating the debate of a single, unified lexicon or multiple lexicons that can be simultaneously accessed. Such argumentation is deemed to be a non-sequitur in the framework adopted here.

Looking first at the Esplugish data, we will illustrate that our version of ***Select*** must exist in order to properly account for CS-data in a Survive-minimalist framework. In particular, we will center our discussion on gender agreement in the DP and how this subsequently affects the configurational possibilities of DPs in Esplugish (data in this section is cited from González-Vilbazo & López[5] (2008) unless otherwise noted). The data elicited by González came from informants at the German School in Barcelona, Spain. Most students/informants are fluent in both German and Spanish, while some of them are also fluent in other languages, e.g., Catalan. Data was elicited in two recorded interview sessions (1996 and 2003) and also via grammaticality judgment tasks. As an introduction to our discussion and analysis of González's data, let us first consider gender marking possibilities in natural languages:

Table 1. Cross-linguistic Gender-Marking Distinctions

1 Gender	Basque, Chinese and English
2 Genders	Arabic, Hebrew and Spanish
3 Genders	Ancient Greek, Czech and German
More than 3	Swahili, Zulu and Dyirbal

Whereas Spanish only displays a 2-gender system (e.g., masculine and feminine). German employs a 3-gender marking system (e.g., masculine, feminine and neuter) on its nouns (note also that the assignment of gender to nouns is arbitrary in both languages):

Spanish:

(18) a. tenedor
fork-MASC

b. cuchara
spoon-FEM

5. Note that these data may come from late bilinguals which might have a different codeswitching behavior from early bilinguals, as shown by Liceras et al. (2008).

German:

(19) a. Löffel
spoon-MASC

b. Gabel
fork-FEM

c. Messer
knife-NEUT

As a result of the differences in the gender marking systems in both Spanish and German respectively, it stands to reason that the gender-marking system in Esplugish must adopt one of the three possible strategies:

- Esplugish will adopt the Spanish gender-marking system.
- Esplugish will adopt the German gender-marking system.
- Esplugish will create a new gender-marking system.

Both of the first two hypotheses encounter difficulties, as will be illustrated with the data below in determiner + N constructions where there is a language mismatch between the determiner and the noun evidenced in the DP. González suggests that the mapping of gender assignment between Spanish determiners-German nouns and German determiners-Spanish nouns is markedly distinct. First, consider the Spanish determiner-German noun data:

(20) a. el Gürtel der Gürtel el cinturón
ART-MASC belt-MASC masculine masculine

b. *la Gürtel
ART-DEF-FEM belt-MASC

c. la Tüte die Tüte la bolsa
ART-FEM (plastic) bag-FEM feminine feminine

d. *el Tüte
ART-MASC (plastic) bag-FEM

e. el Brötchen das Brötchen el panecillo
ART-MASC bun-NEUT neuter masculine

f. *la Brötchen
ART-FEM bun-NEUT

With German and Spanish displaying both masculine and feminine gender in the form of definite articles (i.e., determiners), it would stand to reason that the simplest mapping algorithm would see masculine nouns in German requiring masculine Spanish articles and the like with the feminine nouns and feminine Spanish articles. This fact is borne out in the data: In comparing (20a) with (20b), we see that a masculine German noun requires a masculine Spanish determiner. Example (20b) is ungrammatical due to the inability of the German masculine noun to license a DP with a feminine Spanish article. In similar fashion, (20c) and (20d) exhibit the inverse behavior; namely, that feminine German nouns require feminine Spanish articles in Esplugish. An interesting situation develops in the case of German nouns that are inherently neuter; as noted above, Spanish does not possess a neuter determiner in its morphological inventory. All that is at the disposal of the Esplugish grammar is the masculine Spanish determiner *el* and the feminine determiner *la*. The data in (20e) and (20f) confirm that the masculine determiner *el* (cf. (20e)) can combine with a neuter German noun, while the feminine article cannot (cf. (20f)).

The above data involving Spanish determiner-German noun mapping deal with an under-specification of gender with regard to the inventory of Spanish determiners available in the Esplungish grammar. Inversely, when we are confronted with the opposite scenario; namely, with German determiner-Spanish noun the Esplungish grammar must determine (pun intended) the fate of neuter German determiners. Furthermore, we cannot take for granted the parsimonious feminine Spanish article-German feminine noun/masculine Spanish article-German masculine noun will naturally hold inversely for the German determiner-Spanish noun data in Esplungish. As a matter of fact, the data below refute such a claim:

(21) a. die torre der Turm la torre
ART-FEM tower-FEM masculine feminine

b. *der torre
ART-MASC tower-MASC

c. *das torre
ART-NEUT tower-MASC

d. *die cuaderno das Heft el cuaderno
ART-FEM notebook-MASC neuter masculine

e. *der cuaderno
ART-MASC notebook-MASC

Although the feminine article-feminine noun mapping still holds (cf. (21a)), it appears that neither the masculine or neuter German articles can co-exist in a DP with Spanish masculine nouns (cf. (21b) – (21e)). The situation improves slightly; however, when we consider the potential combinatory capabilities of indefinite German articles with Spanish masculine nouns:

(22) a. *das cuaderno
ART-NEUT notebook-MASC

b. ein cuaderno
ART-INDEF notebook-MASC
MASC/NEUT

An ideal situation would see a universal mapping algorithm of N-gender and the selection of a relevant determiner in the 'other' language in a bilingual grammar. Unfortunately, such a solution is untenable based on the limited corpus of CS-data provided above. Esplugish does not adopt the German or the Spanish gender-marking system. As for developing its own, separate grammar in relation to CS-data, the data suggest that a parsimonious, ubiquitous explanation is not possible. The natural gender of the noun and its language of origin control the selection of the determiner. The question that remains to be answered is to what degree this is a "problem" for the narrow syntax? In almost all flavors of Minimalism – including Survive-minimalism – the valuing of gender within a DP takes place through the checking of an agreement (i.e., phi-) feature, most likely (but not obligatorily required) in a local Spec-Head relationship. In a Survive-minimalist approach, such as the one most recently sketched out by Putnam and Stroik (2009), the "survival" of the noun for ***Merge*** and ***Remerge*** opportunities and the subsequent selection of a proper determiner is dependent on the Interface Features (IFs) of the head noun. Survive-minimalism has the potential to avoid some of the perceived shortcomings that more mainstream versions of the MP run into with regard to gender-mismatches in CS-data. In most instantiations of the MP, lexical items are pre-selected by means of an operation labeled ***Select***, which probes the ***Lexicon*** and selects only those lexical items that are deemed necessary to insure derivational convergence at the external interfaces. Once selected, these lexical items are placed in a ***Lexical Array***, awaiting application and valuing in the narrow syntax. Since agreement features are often unvalued and/or uninterpretable prior their existence in the narrow syntax according to Chomsky (2000 and later subsequent work), ***Select*** should by default be unable to correctly probe the ***Lexicon*** and select the correct determiner. Quite simply, the C_{HL} has the opportunity to "crash" prior to the initial operation in the narrow syntax; in other words, it is likely doomed from the start. Survive-minimalism sidesteps many of these possible pitfalls with two of its seminal desiderata: First, following Stroik and Putnam (2005) and Stroik (2009) (and contra Svenonius 2007), Survive-minimalism does not advocate the existence of uninterpretable features in syntactic operations. The rationale for this conjecture is straightforward: If we assume that the all lexical and functional items are borne interface, i.e., LF and PF, interpretable, the theory of syntax is a theory that, according to Putnam (2007:38) "begins in lexical items and ends with concatenated features." With the elimination of uninterpretable/unvalued features at all of levels of representation – including in the ***Lexicon*** and

Numeration – the selection process will be more straightforward; the operation ***Select***, regardless of what degree we argue for this formal operation's existence in Survive-minimalism, will not be as impaired as in other forms of the MP when engaging and its search and selection procedures. Second, Survive-minimalism significantly reduces the distance between the ***Lexicon*** and the narrow syntax. Following Stroik (2000, 2009 this volume), Survive-minimalism obviates the possibility of what Hornstein et al. (2005) refer to as "crazy numerations", i.e., numerations that entail a massive over-generation of lexical items for potential participation in syntactic operations. At face value, such a view of the selection process between the ***Lexicon*** and the ***Lexical Array***/narrow syntax is more reminiscent of the GEN function in Optimality-Theoretical Syntax rather than a minimalist model. Arguing for a more direct mapping relation between the ***Lexicon*** and the narrow syntax, the system developed in Stroik (2000, 2009 this volume) and Putnam and Stroik (2009) has the advantage of not producing a ***Lexical Array*** that is essentially doomed for failure once the determiner must be inserted. With this being said, Survive-minimalism also faces the challenge of correctly selecting the correct determiner to conform to the gender agreement requirements of Esplugish – or any other CS-grammar for that matter – with regard to the head-noun of the DP. Based on the data outlined above, one-to-one correspondences between Spanish determiners-German nouns and German determiners-Spanish nouns sets are not possible. As a result, we observe that the selection process of determiners in CS-settings presents a serious challenge to that which currently exists in Survive-minimalism.

Unlike in monolingual grammars, the correct selection of a determiner in some CS-grammars is not as direct as the one-to-one algorithms found in the former. To circumvent this issue we propose the implementation of soft, violable constraints that act between the ***Lexicon*** and narrow syntax. Note again that this option is only available in a generative framework that neutralizes the contrast between interpretable and uninterpretable features (cf. Stroik & Putnam (2005)). In principle, the search and selection process enacted by ***Select*** in a more traditional view of the MP, which places lexical items into a ***Lexical Array*** prior to their participation in a derivation, could also be constrained by the same violable constraints that we propose below. The crucial difference; however, is the stance one takes with regard to (un)interpretable features: Even if ***Select*** was subject to some sort of filtering mechanism in the form of violable constraints, the eventual head-noun and determiner would have no viable means of communicating with one another prior to their insertion into the narrow syntax. As a result, once again, the system would have created a ***Numeration*** doomed for failure. Below we provide a novel sketch of how such constraints on ***Select***, which, of course, must exist to some extent in Survive-minimalism based on the Esplungish data above, would look and function in a Survive-minimalist approach to CS-grammars.

First, let's consider the situation when a German head noun appears with a Spanish determiner (cf. (20a) – (20f)). As borne out in the data above, the gender of the German noun accepts masculine and feminine Spanish determiners that match the gender value of the head-Noun. We capture this generalization with the following constraint:

Match
The gender of the determiner must be identical to the gender of the head-Noun.

The only difficult issue arises with German nouns with neuter gender; since Spanish does not exhibit neuter gender marking, the grammar of Esplungish must find a new mapping strategy. As demonstrated by the examples in (20e) and (20f), whereas masculine Spanish articles can co-occur with German neuter nouns, the presence of the feminine article results in an ungrammatical result. This observation is encapsulated in the constraint [det-masc] > [n-neut] (introduced below). Note that the language of the respective determiners and nouns do not have to be specified in the constraint due to Spanish's lack of neuter gender.

[det-masc] > [n-neut]
Masculine determiners may co-occur with neuter nouns.

The ranking for these constraints is not crucial; however, the constraint Match will apply to most cases (i.e., except when the German head-Noun is neuter). The following tableau demonstrate these constraints in action.

Tableau 1.

Det: [x] / Noun: Ger[$_{\text{MASC}}$]	Match	[det-masc] > [n-neut]
☞ $_{\text{DP}}$[el Gürtel]		
* $_{\text{DP}}$[la Gürtel]	*!	

The input consists of the language and gender of the head-noun (in this case, we are dealing with a masculine German noun *Gürtel* 'belt') and an open variable for the Spanish determiner that will be selected from the ***Numeration***. The unspecified input value is not a violation of the Richness of the Base hypothesis that is central to Optimality-Theoretical accounts; rather, the unspecified input is reminiscent of OT-LFG formalisms (cf. Bresnan 2001) where certain aspects/values of the f-structure remain incomplete prior to their valuation by means of algorithmic functions. Unlike most frameworks that employ violable constraints at some level of linguistic knowledge, the possible candidate set generated by GEN – at least for Esplungish – is a paltry two candidates. It is crucial to keep in mind that the violable constraints that we advance above in **Tableau 1** do not play a role in the narrow syntax; Survive-minimalism in its current instantiation does not advocate the employment of soft constraints that act upon derivational operations. Rather, **Tableau 1** represents the selection of the head of artP from the ***Numeration***. To demonstrate, consider example (20a) repeated as (23) below for the sake of the reader:

(23) el Gürtel
ART-MASC belt-MASC

Following the successful initial merger of the noun *Gürtel* 'belt' into the narrow syntax, this noun projects (among other features) its natural gender feature [+ MASC]. Following Putnam and Stroik (this volume), phi-features are regarded as ***Interface Features*** (IFs) that will force the remerger of lexical item that possesses them at a later stage of the derivation. Following Roehrs' (2006) partition of the internal structure of DPs, the phi-features, most notably the gender feature, plays the decisive role in selecting a possible article from a language other than that of the noun (N). Setting aside projections and derivational steps other than those immediately necessary for the time being, the sub-tree in (24) represents the stage of the derivation where the constraints in **Tableau 1** apply:

(24) a. $[_{NP}$ Gürtel$_{[+\ MASC]}]$ (initial Merge of N)
b. [+ MASC]-$[_{NP}$ Gürtel$_{[+\ MASC]}]$ (projection of [+ MASC]-feature)
c. **Tableau 1** (selection of head of artP)
d. $[_{artP}$ el$_{[+MASC]}$ $[_{NP}$ Gürtel$_{[+\ MASC]}]]$ (initial Merge of head-artP)

Paramount for our analysis are the derivational steps (24b–d): Following the successful creation of the NP in (24b), the gender feature on N projects onward throughout the derivation until it is time to merge the lexical head of artP into the narrow syntax. At this point, the operation ***Select*** – but crucially not ***Merge*** – is at a crossroads: If a Spanish determiner is selected from the ***Numeration***, what options are available (if any) that will ***Match*** and ***Merge*** with the [+ MASC] feature of N? The constraints listed in **Tableau 1** allow ***Select*** to probe into the Spanish component of the ***Lexicon*** (and ***Numeration***) and extract the determiners that will be optimal choices based on the constraint ranking that restricts ***Select***.[6] For the remainder of this section the tableaux that follow operate in identical fashion to the first in that they: (i) constraint ***Select*** rather than ***Merge***, and (ii) apply at the

6. As mentioned to us by an anonymous reviewer, it is most likely more economical to have the two search fields (e.g., the ***Lexicon*** and the ***Numeration***), ordered in a series with the ***Numeration*** proceeding the ***Lexicon*** in any search procedure. Although at first glimpse this suggestion may lead to a more economical design of the CS-grammars, it completely side steps the discussion of multiple vs. singular lexicons in bilingual grammars.

derivational stage of merging the lexical item (i.e., the determiner) that will occupy the head-artP position. **Tableau 2** illustrates the constraint ranking when the input contains a German neuter noun.

Tableau 2.

Det: [x] / Noun: German[$_{\text{NEUT}}$]	MATCH	[DET-MASC] > [N-NEUT]
☞ $_{DP}$[el Brötchen]	*	
* $_{DP}$[la Brötchen]	*	*!

Both candidates violate MATCH due to the lack of a neuter determiner in Spanish. The optimal candidate is the first candidate, i.e., *el Brötchen* 'the roll', since it does not violate [DET-MASC] > [N-NEUT], which requires that Spanish masculine determiners be permitted to co-occur with German neuter nouns.

As evidenced in the possible combinations of German determiners with Spanish nouns (cf. (21)–(22)), we see a pattern that is much more restrictive. While the German feminine article can appear with a Spanish feminine noun (cf. (21a)) neither the masculine or neuter German determiner can appear in the same DP with a Spanish masculine noun (cf. (21b), (21c), and (21e)). Furthermore, as we observed in the Spanish determiner-German head-Noun DPs, feminine articles cannot modify masculine nouns (cf. (21d)). Initially we see that our MATCH constraint in its current form is too general and must be split to account specifically for masculine and feminine nouns:

MATCH: F-F
Feminine determiners can appear with feminine nouns

MATCH: M-M
Masculine determiners can appear with masculine nouns

By parsing the original MATCH constraint into **MATCH: F-F** and **MATCH: M-M** respectively, we now have proper constraints that will accept example (22a) with the feminine article-feminine noun pairing. From the data presented above; however, there are no instances of masculine article-masculine head-noun mappings with regard to German determiners and Spanish nouns. Neither neuter nor masculine German determiners may co-occur with Spanish masculine nouns. There is; however, an apparent loophole: Whereas *definite* masculine and neuter German determiners are banned from occurring with masculine Spanish nouns, it appears that the restriction is lifted when the articles are indefinite; i.e., under-specified for gender (both the masculine and neuter Nominative indefinite article share the homophone *ein* ‘a’).[7] To account for this preference, we suggest the presence of a constraint that favors gender-ambiguous articles:

AMBI-GENDER
Select articles that are ambiguous in their gender distinction

Tableau 3 is provided for the article selection of example (23b), *ein cuaderno* ‘$a_{\text{Masc/Neut}}$ notebook.’

Tableau 3.

Det: [x]/Noun: Span[$_{\text{Masc}}$]	MATCH: F-F	MATCH: M-M	AMBI-GENDER
☞ $_{DP}$[ein cuaderno]	*		
* $_{DP}$[das cuaderno]	*	*!	
* $_{DP}$[der cuaderno]	*		*!
* $_{DP}$[die cuaderno]	*	*!	

All candidates violate **MATCH: F-F** due to the fact that the head-Noun in the input is masculine. The motivation for the high-ranking of this constraint ensures; however, that examples such as (21a) are evaluated as grammatical. **MATCH: M-M** eliminates the candidates that exhibit non-masculine articles, i.e., the second and fourth candidates. The eventual winner of the competition – due to its underspecified gender marking as being potentially either neuter or masculine – as well as the candidate with a masculine determiner does not violate this constraint. The optimal candidate of this competition is determined by the constraint **AMBI-GENDER**, which selects the DP with the gender-ambiguous indefinite article. In conclusion of this section, it is crucial to understand exactly when this competition takes place. Returning to Roehrs’ suggested internal DP-structure that we adopt in this pilot study, these tableaux represent the selection restriction between the ***Numeration*** and the narrow syntax at the point in the derivation when the head of artP enters the narrow syntax. The upside to this analysis is that the soft, violable constraints do not act upon syntactic operations per se, but rather the selection process of lexical

7. It is interesting to note that this under-specification for gender indefinite articles in German only exists for Nominative and Accusative (excluding masculine nouns) Case-marked nouns. As a matter of fact, the distinction between singular neuter and masucline nouns in the accusative case would render a situation where the indefinite article is no longer ‘gender-ambiguous’ (cf. NEUT. ACC. Singular: [ein] / MASC. ACC. Singular: [einen]). Future studies on German-Spanish CS-data must include a focus on Case distinction in order to arrive at a better understanding at the role that Case would play in possible determiner-noun combinations.

items from the ***Numeration***. Again, this is possible due to Survive-minimalism's view of the relationship of the more direct relationship of ***Numeration*** in relation to syntactic operations. A potential downside that we see to this line of reasoning is that through the introduction of violable constraints that filter selections from the ***Numeration*** in CS-grammars could be perceived as an attempt to introduce an internal-level of syntactic structure with filtering operations between the ***Numeration*** and narrow syntax. With that being said, the concept we develop here is that in a bilingual grammar where two sets of determiners from Language 1 and Language 2 are both viable options in the construction of a DP, the gender information present on the head-Noun of the DP is able to communicate with lexical items present in the ***Numeration*** to ensure optimal congruence between the head and determiner in a given DP. Although this notion of constraint interaction filtering the selection process between the ***Numeration*** and the narrow syntax must be further developed to include other facets of DP-internal and -external properties (e.g., Case, referentiality, theta-role interpretation, etc.), we hope to have shown here that this idea has promise due to its adherence to core Minimalist principles and is indeed worthy of further pursuit.

5. Linearization issues

The linear order of constituents in CS data also presents a significant challenge to any derivational formalism of syntax. In many mainstream approaches to the MP, linear ordering is often dealt with by means of Kayne's (1994) **Linear Correspondence Axiom** (LCA), which regulates the linear order of constituents by means of the anti-symmetry of syntactic heads prior to the presentation of derivationally compiled representations to PF. As alluded to by Kobele (this volume), Survive-minimalism adopts a strict, unviolable view of the connection between the hierarchical order derived from the narrow syntax and the final phonological representation delivered to PF. Survive-minimalism significantly departs from other Minimalist approaches to linearization in its treatment of occasional "disconnections" between PF and LF with regard to which previous instantiation of ***Merge/Remerge*** an individual constituent has undergone in a derivational history should/must be valued independently at PF and LF. For example, in his discussion of reconstruction effects connected with middle field scrambling in Russian and Japanese, Putnam (2007:196–199) comes to terms with the mechanics of how the scope (LF) and phonological (PF) aspects of a grammatical string are sketched out in Survive-minimalism. The contrast in scope readings between Russian and Japanese (taken from Bailyn 2001) is illustrated in examples (25) and (26) below. Agreeing with Putnam (2007), the contrast in scope readings between the

Russian example (25a) and the Japanese example (26a) supports the argument against LF-lowering as a viable means of scope reconstruction (contra Bošković and Takahashi (1998) and Bošković (2004)). At the core of the mechanism known as LF-lowering is the prediction that there is neither surface interpretive facts nor locality constraints at play. Example (27) shows that Russian scrambling over the subject is still subject to subjacency violations. Therefore, these counter-examples to LF-lowering clearly illustrate that scrambling is a syntactic operation that is sensitive to island effects.

Russian

(25)	a.	[Každogo malčika]$_i$	kto-to	xo čet,	čtoby	Boris	uvidel **t**$_i$.
		[every boy]$_{ACC}$	someone$_{NOM}$	wants	that	Boris	saw

‘Everyone boy someone wants Boris to see.’

b. (i) *∃x ∀y (ii) ∀y ∃x
(high scope for the scrambled NP$_{ACC}$)

Japanese

(26)	a.	[Daremo-ni]$_i$	dareka-ga	[Mary-ga **t**$_i$	atta	to]	ometteiru.
		everyone$_{DAT}$	someone$_{NOM}$	Mary$_{NOM}$	met	that	thinks

‘Somone thinks that Mary met everyone.’

b. (i) ∃x ∀y (ii) *∀y ∃x
(low scope for the scrambled NP$_{DAT}$)

(27)	*Borisa$_i$	ty	pozvonil	[agentu	[kotoryj	ljubit **t**$_i$]]
	Boris$_{ACC}$	you$_{NOM}$	phone	spy$_{DAT}$	who	loves

Intended: ‘It’s Boris who you phoned a spy who loves!’

Assuming that scrambling is driven by some sort of grammatical feature that is active in the narrow syntax – referenced by the mnemonic Σ – Putnam makes suggestions on how a Minimalist approach to XP-displacement based on the Survive Principle could further simplify the differences in scope reconstruction in Russian and Japanese. Constituent displacement enacted by the Survive Principle offers a unified account of scrambling in Japanese and Russian in spite of the different scope readings evidenced in (25) and (26). If both Russian (R) and Japanese (J) possess the scrambling-feature Σ, then the different scope readings in R and J respectively are the result of different realizations of the scrambled objects at both LF and PF. The proposed derivational histories, taken from Putnam (2007:1999), will elucidate this point:

(28) **Russian**

a. PF: [CP XP$_Σ$ Subj … [*v*P … ~~XP~~$_Σ$ …]

b. LF: [CP XP$_Σ$ Subj … [*v*P … ~~XP~~$_Σ$ …]

c. Scope interpretation: *∃x ∀y > ∀x ∃y

(29) **Japanese**

a. PF: [CP XP_{Σ} Subj … [*v*P … ~~XP~~$_{\Sigma}$ …]

b. LF: [CP ~~XP~~$_{\Sigma}$ Subj … [*v*P … XP_{Σ} …]

c. Scope interpretation: $\exists x \forall y > {}^{*}\forall x \exists y$

Following Putnam (2007:1999), "Note the key difference between the LF derivational history of R and J in steps (28b) and (29b) respectively. Although both scrambled constituents are recognized at PF in their respective scrambled positions, the LF interpretation of the scrambled items differ considerably. Based on the key tenets of Survive-minimalism, a version of 'reconstruction' is no longer necessary in syntactic theory: All 'reconstructions' involve non-covert concatenations. In other words, nothing is actually 'reconstructed.' The scope reconstruction that merged in (29b) exists because the dislocated, i.e., remerged, DPs had previous instances of ***Merge*** that are interpretable at LF." As mentioned to us by John Hale (p.c.), such a view of PF renders this interface's participation in evaluating and compiling/linearizing complete representations very powerful from a mathematical point of view. Following Putnam (2007), such a view is not only expected, but perhaps also welcome, if we hold to the conjecture that LF remains constant cross-linguistically, whereas language variation resides in operations found at the PF-interface.

Although Putnam's approach to PF/LF disconnects opens the door to future studies centering on linearization issues in Survive-minimalism, these ideas are not fully developed and focus (thus far) solely on monolingual grammars. As could be expected, CS data challenge many hypotheses under standard Minimalist assumptions regarding how phonologically myopic the narrow syntax can/must remain with regard to information necessary for proper linearization. As we shall discuss below, the analysis that we will expand upon in this paper for DET-ADJ-N ordering sequences finds a natural home within a version of the MP that adopts the Survive Principle. In particular, our discussion will center on the linearization of DET-ADJ-N, illustrated below in (30) and (31).

(30) The big casa
The big house
(Quintero, cited in Woolford 1983)

(31) *El hombre old
DET man
The old man
(Gingrás, cited in Woolford 1983)

5.1 Determiner position

According to Moro's (in press) and Deuchar et al.'s (2008) analyses of mixed Spanish-English nominal constructions consisting of a determiner followed by a noun phrase, the determiner will always come from the language which has grammatical gender, i.e., Spanish. Moro explains that Chomsky (2001) proposes a MINIMALIST architecture in which features (such as the phi-features of D or Agr, but not N) enter the derivation without values specified. This phi-features set is a probe that seeks a goal ("matching features" on the Noun) establishing agreement. The operation ***Agree*** values and deletes these features. Only a probe or functional category with a full set of phi-features is capable of deleting and valuing its own features. Moro follows Chomsky (2000:124) and "takes deletion to be a 'one fell swoop' operation, dealing with the phi-set as a unit. Its features cannot selectively delete: either all delete or none." This analysis was extended to Welsh-English code-switched constructions by Deuchar et al. (2008):

(31) a. Spanish D, phi = {person, number, gender}
English N, phi = {person, number}

*b. English D, phi = {person, number}
Spanish N, phi = {person, number, gender}

(32) a. Welsh D, phi = {person, number, gender}
English N, phi = {person, number}

*b. English D, phi = {person, number}
Welsh N, phi = {person, number, gender}

According to Deuchar et al. (2008) in the grammatical examples above, the Spanish and Welsh Probes (the AgrP of the Spanish or Welsh DP containing a complete set of phi-features) delete and value their unvalued phi-set. However, in the ungrammatical examples, the deficient English Probe, i.e., the AgrP of the English DP, lacks gender. Since it only has two phi-features, but is entering into a construction with a noun from Welsh or Spanish that has the additional gender phi-feature, it is incomplete and is unable to delete its uninterpretable features all at the same time. Consequently, the derivation does not converge.

Deuchar et al. (2008) point out that "the definition of the English probe as 'deficient' only arises in code-switching environments". Thus, when an English determiner is used monolingually, with an English noun, the presence of only two phi-features is not an obstacle to convergence, because gender features do not enter into the monolingual English derivation.

5.2 Adjective position

Regarding the position of adjectives in a different language from the noun they modify, Deuchar et al. (2008) claim that in nominal constructions consisting of an adjective plus noun, the grammar of the language of the adjective dictates the position of the adjective relative to the noun. Thus if the adjective is only used pre-nominally in its own language, it will occur before the noun in switched nominal constructions. If it is a post-nominal adjective in its own language, then it will only occur after the noun in mixed constructions. If the adjective can occur in either position in its own language, then it can also appear either before or after in nominal constructions containing a switch. If the two languages participating in the code-switched nominal construction have the same default position for adjectives with respect to the noun they modify, then the switch will have no effect on adjective placement.

In their treatment of Welsh, English, and Spanish DPs, Deuchar et al. (2008) assume that the base structure is identical for all three languages. According to Deuchar et al., in English no overt movement occurs and this base structure corresponds to the linear order of constituents at the end of the derivation. In Spanish and Welsh, overt movement takes place because the adjective is inserted into the derivation with uninterpretable and unvalued inflectional morphemes of gender and number that must be checked. In English, movement also takes place to allow the noun to check its features with those of the leftmost D head. However the movement is covert, i.e., it takes place at Logical Form (LF) after Spell Out and is not reflected in the phonological form (PF). They argue that the movement is overt in Spanish and Welsh by force, because the adjectives' features are uninterpretable.

After explaining their position on the base structure of DPs, Deuchar et al. (2008) return to code-switched DPs. They claim that since the agreement features of Spanish and Welsh adjectives are uninterpretable, they attract the nominal head (even if it is in English) ***before*** Spell Out. Any unchecked and consequently undeleted uninterpretable features will cause the derivation to crash if they remain at the point at which PF and LF diverge. Therefore the resulting word order is one with an English noun and a Spanish or Welsh postnominal adjective. On the other hand, because English adjectives do not carry uninterpretable gender features, they do not force the noun (even if it is in Spanish or Welsh) to move. The adjective remains pre-nominal in mixed nominals, and the agreement features can be checked, covertly at LF. Conversely, they claim an English adjective used post-nominally with a Spanish or Welsh noun is not predicted (* "el hombre ***old***") under their analysis.

Although their approach merits further explanation in some facets, Moro and Deuchar et al.'s treatment of DP-internal linearization issues in its current instantiation are in some crucial respects incompatible with a Survive-minimalist model. More concretely, there are four key issues where we believe Moro and Deuchar et al.'s (2008) analysis falls short. First, they continue to make use of mainstream Minimalist notions such as the uninterpretable/interpretable feature distinction that hold no place in Survive-minimalism. Secondly and closely related to the first point, their analysis advocates the existence of non-local, i.e., non-sisterhood, checking relations, e.g., Probe-Goal (***Agree***) relations established through derivational c-command (cf. in the sense of Epstein (2001)). The Survive Principle functioning as, in the words of Preminger (2008) and Lechner (this volume) as a push-chain version of syntactic displacement, envisages a syntactic system that requires local evaluations at every iterative step of the derivation. Accordingly, any instantiation of the MP that makes use of Probe-Goal relations in the narrow syntax requires – at least to some degree – a form of look-back or look-ahead machinery that is undesirable in a crash-proof system of syntactic theory following Frampton and Gutmann (2002). Third, their observation, although potentially correct, that the language of the adjective in the DP determining the correct **Det**(erminer)-**Adj**(ective)-**N**(oun) linear order, is at odds with most standard analyzes of the licensing relationship that exists between determiners and adjectives (cf. Roehrs (2006) as outlined in Section 3). Lastly, the assertion that the specification for gender on D in English-Spanish and English-Welsh CS-data is responsible for the grammaticality of agreement (and to some degree also responsible for linearization of items within the DP) is stipulative and language specific. Furthermore, the following example poses significant difficulty for the ***Matrix Language Framework*** provided that the example in question does not provide us with enough information to determine which language functions as the "matrix language." However, in this case, the Spanish adjective *pequeño* "little/small" appears to follow the English word order, contra to Moro's predictions.[8]

(33) Un pequeño pocket
DET.M.SG little.M.SG pocket
"A little pocket"
(data from Sastre 1, Bangor corpus in progress)

The approach that we would like to tentatively adopt seeks to advance a solution that is (potentially) able to encompass more CS-data with regard to its coverage rather than reducing our analysis to language-specific constraints that make it difficult, if not impossible, to maintain the explanatory adequacy expected of generative grammar formalisms.

As demonstrated effectively in **Section 4**, bilingual grammars sometimes significantly deviate from either of the parent languages that contribute to the composition of the former. The challenged issued by these data is to see if there

8. However, Moro has pointed out that certain adjectives (e.g., *pequeño* 'little/small') do appear pronominal position in Spanish therefore example (33) does not violate Spanish syntactic constraints. Further research into the different classes of adjectives and how they can potentially affect the linearization in CS-situations is needed.

exists a derivational solution to these constituent orderings that does not rely upon EPP-features. As discussed by Epstein and Seely (2002, 2006), Stroik (2009) and a host of others, EPP-features are spurious and contribute nothing to the interfaces. From a descriptive point of view, we agree with Moro's suggestion that the language of the adjective determines the DP-internal constituent ordering; however, since the adjective is not the head of a functional projection, our analysis of DET-ADJ-N orderings in bilingual grammars in a Survive-minimalist framework is in need of some minor adjustments.

Returning to Roehrs' (2006) underlying DP structure that we adopt in this paper, we adopt the view that determiners are base-generated as the head of artP and the adjectives in Spec,AgrP. According to this view, the DET-ADJ-N and DET-N-ADJ orders would be derived according to the following derivational histories:

(34) DET-ADJ-N

a. Base: $[_{DP}$ D^0 $[_{AgrP}$ Adj Agr^0 $[_{artP}$ Det $[_{NP}$ $N^0]]]]$

b. Representation: $[_{DP}$ D^0-Det^0 $[_{AgrP}$ Adj Agr^0 $[_{artP}$ ~~Det~~ $[_{NP}$ $N^0]]]]$

(35) DET-N-ADJ

a. Base: $[_{DP}$ D^0 $[_{AgrP}$ Adj Agr^0 $[_{artP}$ Det $[_{NP}$ $N^0]]]]$

b. Representation: $[_{DP}$ $[_{artP}$ Det $[_{NP}$ N^0 $]_i$ D^0 $[_{AgrP}$ Adj $Agr^0]$ $\mathbf{t}_i$ $]$

Simply stated, the DET-ADJ-N order is generated by the head-movement of DET (base-generated as art^0) to D, whereas the DET-N-ADJ order involves the movement of the entire artP to the Spec,DP position. Taking an agnostic view on the motivation for these two word orders in monolingual grammars for the time being, we adopt the view that the native language of the adjective in CS-data from English-Spanish and English-Welsh motivates the leftward movement artP to Spec,DP, thus creating the DET-N-ADJ order.

(36) **Derivation for DET-ADJ-N**

a. the big casa (Quintero, cited in Woolford 1983)
the big house
'the big house'

b. Base: $[_{DP}$ the $[_{AgrP}$ big Agr^0 $[_{artP}$ the $[_{NP}$ casa$]]]]$

c. Representation: $[_{DP}$ the $[_{AgrP}$ big Agr^0 $[_{artP}$ ~~Det~~ $[_{NP}$ $\underline{\text{casa}}^0]]]]$

(37) **Derivation for DET-N-ADJ**

a. He presented a ***paper exceptionnel***
paper exceptional
'He presented an exceptional paper.'
(English-French; Bokamba, 1989:282, (16a))

b. Base: $[_{DP}$ a^0 $[_{AgrP}$ expectionnel Agr^0 $[_{artP}$ a $[_{NP}$ paper$]]]]$

c. Representation: $[_{DP}$ $[_{artP}$ a $[_{NP}$ paper $]_i$ a $[_{AgrP}$ expcептionnel $Agr^0]$ $\mathbf{t}_i]$

Consonant with Moro's analysis that the adjective "calls the shots", the challenge in a Survive-minimalist approach is what sort of feature is responsible for motivating the different linear orders in (34) and (35). At this point we would like to explore the option that both resulting orders of DET-ADJ-N are inherently related to one another and can be derived from the same base structure. Rather than calling upon a feature similar in scope and function to the EPP, we maintain that a formal syntactic feature – most likely definiteness [+ Def] – motivates is responsible for both movement types.[9] The novel sketch of this proposal starts by looking at the two distinct, yet related movements that yield the orders in (34) and (36) and (35) and (37) respectively. Our analysis of DPs requires two types of dislocation: art-dislocation (where N remains in situ) which yields the DET-ADJ-N order and art-N-dislocation (where N dislocates together with the article) which yields the DET-N-ADJ ordering. If the [+ Def] feature resides on the article alone, then the article in artP (36c) will dislocate on its own; however, if this feature resides on (is inherited by) the $[_{artP}$ α $[_{NP}$ N] unit, then the entire complex phrase would pied-pipe and dislocate together (37c). Hence, the differences found within DP linearity would involve the same features, but different inheritance relations. Although this theory is conceptually appealing and consistent with the main tenets of Survive-minimalism, it comes at the price of attempting to explain how feature inheritance works in this version of the MP.[10] Although the analysis of pied-piping and feature inheritance would be quite detailed and somewhat complex in Survive-minimalism, it would most likely take place within the ***WorkBench*** (i.e., as sub-derivations) rather than the narrow syntax. Be that as it may, we leave the specific details of this analysis for future studies. As the analysis now stands, we have shown how we can unite both movement types under one feature, i.e., [+ Def] without resorting to spurious EPP-features that do not contribute any interpretive material to either the LF or PF interface.

This section has endeavored to show in detail the derivation of code-switched DPs with respect to DET-ADJ-N combinatory and linearization factors. Moro's theory has been examined against the context of Survive-minimalism principles, which embodies the notion of less effort in language design. Despite some problems,

9. Thanks to Tom Stroik (p.c.) for discussing the feasibility of this analysis with us.

10. As pointed out to us by Vicki Carstens (p.c.), the relationship between the Numeration, WorkBench and the narrow syntax in Survive-minimalism faces a significant challenge when dealing with feature inheritance and pied-piping, due to its abandonment of ***Internal Merge***, i.e., the possibility of active features remaining in the narrow syntax motivating Derivation-to-Derivation movement chains. As a result, the challenges mentioned in this section are similar to issues such as remnant movement at the clausal-level.

it can be safely asserted that the proposed derivations give a principled account of the code-switching data considered in the different language pairs. Although significant ground has been covered, it is evident that much work is still open for further detailed investigation, both from an empirical and conceptual standpoint.

6. Conclusions and future research

In this paper we provided a novel analysis of two issues related to DP-structures in CS-grammars, namely, the relationship between determiners and nouns in a DP where each respective lexical item originates from a separate language and the linearization of **Det**(erminers)-**Adj**(ectives)-**N**(ouns) in CS-grammars where each respective language contributes a surface order that contrasts with the other. We approached these related issues from the perspective of Survive-minimalism, demonstrating both the "pros" and "cons" that such a system offers in better understanding the grammar properties (and limitations) of bilinguals. With regard to determiner selection, Survive-minimalism – due primarily to its argument for a more direct relation between the ***Numeration*** and the narrow syntax – made it possible to introduce violable constraints that filter the selection possibilities of determiners in the grammars under investigation. Regarding the questions surrounding the linear order of constituents in bilingual DPs, we introduced the notion of a formal feature; namely, definiteness [+ Def], which triggers the re-configuration of lexical items to conform to structural requires within a given CS-grammar. As a result of this analysis, we can argue for the same feature motivating both the DET-ADJ-N and the DET-ADJ-N orderings, with the former ordering being a result of artP undergoing displacement to a higher syntactic position in the DP and the latter requiring the complex concatenated unit $[_{artP}\ \alpha\ [_{NP}\ N]]$ to pied-pipe and move to Spec,DP. Admittedly, explaining pied-piping and feature inheritance in Survive-minimalism is a topic worthy of its own exploration, hence we leave a detailed discussion of these topics for future research. The advantages of pursuing this analysis of DP linearity in CS-grammars is that it is: (i) consistent with the desiderata of Survive-minimalism and (ii) does not require features similar to the EPP to exist in the system. Returning briefly to Moro's (in press) observation that the language of the adjective in bilingual DPs determines the linear order of its internal elements, her proposal would seem to predict that DPs with multiple adjectives from the same source language bearing identical phi-features will require the adjectives to be either all left-of-N (because N does not dislocate) or all right-of-N (because $[_{artP}\ \alpha\ [_{NP}\ N]]$ dislocates to the D-head). To the best of our knowledge, these constructions have not been researched and are a top priority of our future research in this area.

The formal study of code-switching offers us new perspectives on how speakers design a new form of language. As claimed by Leung (2001), CS exhibits the main trait of human language competence-the capability and courage of human beings to create and even take risks to make infinite use of language by finite means (Chomsky 1995). Thus far we have only looked at very basic DP-level phenomena; in the future we want to expand our coverage/discussion of this model in analyzing clausal-level data as well as more complex DP-structures and puzzles.

We have contended that there are syntactic reasons for code-switching but we don't deny that there are also metalinguistic ones, in line with Gardner-Chloros and Edwards' (2004:126) claim that "…although ***syntax*** plays an important role in CS, it cannot be assumed *a priori* that the constructs of syntacticians are the best means for characterising the processes of performance data such as CS". We should then look beyond mere linguistic behavior and examine the socio/political environments of the code-switchers (Nartey 1982). Furthermore, another difficulty that studies of code-switching run into is the use of the term "ungrammatical" for nothing more than a tendency (Clyne 2000:278). All of these factors taken together make CS-data a serious challenge for any generative model of grammar.

In conclusion, we recognize that there remains a great deal to discover concerning code-switching not only at the DP-level, but also at the sentential level. Perhaps the major realm of inquiry to be explored concerns the patterns of code-switching in a typologically very wide range of language pairs, such as the Kinyarwanda–French alternation explored by Gafaranga (2007). Otherwise, we may run the risk of developing and supporting powerful universal models which cannot be substantiated universally (Nartey 1982:188). At the very least, all these attempts to explain CS-data will lead us (following Boeckx 2007) to determine the boundaries of optimal properties of grammar, telling us how much language can be considered optimal, and to what extent other properties of this natural phenomenon have to be justified in currently unintelligible ways. What cannot be denied is that CS-data leads us to question different principles of grammatical formalisms, and test which of these frameworks can best explain (if this is possible in its strongest hypothesis) CS-data. Here we have presented a short, yet concise example of how Survive-minimalism presents a new approach to some old problems.

References

Bailyn, J. 2001. On scrambling: A response to Bošković & Takahashi. *Formal Approaches to Slavic Linguistics* 14(1): 1–22.

Belazi, H.M., Rubin, E.J. & Toribio, A. J. 1994. Code switching and X Bar theory: The functional head constraint. *Linguistic Inquiry* 23(2):221–237.

Boeckx, C. 2007. *Understanding minimalist syntax.* Malden MA: Blackwell.

Bokamba, E.G. 1989. Code-mixing, language variation and linguistic theory: Evidence from Bantu languages. *Lingua* 76: 21–62.

Borer, H. 1984. *Parametric syntax*. Dordrecht: Foris.

Bošković, Ž. 2004. Topicalization, focalization, lexical insertion, and scrambling. *Linguistic Inquiry* 35: 613–638.

Bošković, Ž. & Takahashi, M. 1998. Scrambling and last resort. *Linguistic Inquiry* 29: 347–366.

Bresnan, J. 2001. *Lexical functional grammar.* Malden MA: Blackwell.

Brugé, L. 1996. Demonstrative movement in Spanish: A comparative approach. *University of Venice Working Papers in Linguistics* 6(1): 1–61.

Brugé, L. 2002. The positions of demonstratives in the extended nominal projection. In *Functional structure in DP and IP: The cartography of syntactic structures,* Vol. 1, G. Cinque (Ed.), 15–53, Oxford: OUP.

Campbell, R. 1996. Specificity operators in Spec, DP. *Studia Linguistica* 50(2): 161–188.

Carstens, V. 1991. The morphology and syntax of determiner phrases in Kiswahili. Ph.D. dissertation, University of California, Los Angeles.

Chan, B.H. 2003. *Aspects of the syntax, the pragmatics and the production of code-switching: Cantonese and English.* Frankfurt: Peter Lang.

Chomsky, N. 1995. *The minimalist program.* Cambridge MA: The MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, R. Martin, D. Michaels & J. Uriagereka (Eds), 89–115. Cambridge MA: The MIT Press.

Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A life in language*, M. Kenstowicz (Ed.), 1–52. Cambridge MA: The MIT Press.

Cinque, G. 1994. On the evidence for partial N-movement in the Romance DP. In *Paths towards universal grammar: Studies in honor of Richard S. Kayne*, G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi & R. Zanutini (Eds), 85–110. Washington DC: Georgetown University Press.

Cinque, G. 1999. *Adverbs and functional heads.* Cambridge: CUP.

Clyne, M. 1987. Constraints on code–switching – How universal are they? *Linguistics* 25: 739–64.

Clyne, M. 2000. Contraints on code-switching. How universal are they? In *The bilingualism reader*, Li Wei (Ed.), 241–264. London: Routledge.

Corver, N. 1991. Evidence for DegP. *Proceedings of NELS* 21: 33–47.

Corver, N. 1997. The internal structure of the Dutch extended adjectival projection. *Natural Language and Linguistic Theory* 15: 289–368.

Deuchar, M. , Herring, J. & Moro, M. 2008. Evaluating theories of intrasentential code-switching across language pairs: Nominal constructions in Spanish-English and Welsh-English data. Paper presented at ICCLING 2008. Stellenboch, South Africa.

Di Sciullo, A.M., Muysken, P. & Singh, R. 1986. Government and code mixing. *Journal of Linguistics* 22: 1–24.

Frampton, J. & Gutmann, S. 2002. Crash-proof syntax. In *Derivation and explanation in the minimalist program*, S. Epstein & T.D. Seely (Eds), 90–105. Oxford: Blackwell.

Epstein, S. 2001. Un-principled syntax: The derivation of syntactic relations. In *Working minimalism*, S. Epstein & N. Hornstein (Eds), 317–346. Cambridge MA: The MIT Press.

Epstein, S. & Seely, T.D.. 2002. Rule applications as cycles in a level-free syntax. In *Derivation and explanation in the minimalist program*, S. Epstein & T.D. Seely (Eds), 65–89. Oxford: Blackwell.

Epstein, S. & Seely, T.D. 2006. *Derivationals in minimalism.* Cambridge: CUP.

Gafaranga, J. 2007. *Talk in two languages*. New York NY: Palgrave.

Gallmann, P. 1997. Zur Morphosyntax der Eigennamen im Deutschen. In *Lexikalische Kategorien und Merkmale*, E. Löbel & G. Rauh (Eds), 73–86. Tübingen: Niemeyer.

Gardner-Chloros, P. & Edwards, M. 2004. Assumptions behind grammatical approaches to code-switching: When the blueprint is a red herring. *Transactions of the Philological Society* 102(1): 103–129.

Gass, S. & Selinker, L. 2007. *Second language acquisition: An introductory course,* 3rd edn. New York NY: Routledge.

Ghirardini, F. 2006. Morphological constraints on intra-sentential code-Switching: Toward a standardization of Spanglish grammar. *Annali Online di Ferrara – Lettere* 1: 17–30.

Giusti, G. 1997. The categorial status of determiners. In *The new comparative syntax*, L. Haegeman (Ed.), 95–123, London: Longman.

Giusti, G. 2002. The functional structure of noun phrases: A bare phrase structure approach. In *Functional structure in DP and IP: The cartography of syntactic structure,* Vol. 1, G. Cinque (Ed.), 54–90. Oxford: OUP.

González-Vilbazo, K.E. & López, L. 2008. Syntactic phases and code switching. Presented on May 1st at University of Chicago Linguistics Colloquium.

Grohmann, K. & Panagiotidis, P. 2004. Demonstrative doubling. Ms, University of Cyprus and Cyprus College.

Haider, H. 1992. Die Struktur der Nominalphrase: Lexikalische und Funktionale Strukturen. In *Deutsche Syntax: Ansichten und Aussichten. Institut für deustche Sprache Jahrbuch 1991*, L. Hoffmann (Ed.), 304–333. Berlin: Walter de Gruyter.

Halmari, H. 1997. *Government and codeswitching: Explaining American Finnish* [Studies in Bilingualism 12]. Amsterdam: John Benjamins.

Hinzen, W. 2006. *Mind design and minimal syntax.* Oxford: OUP.

Hornstein, N., Nunes, J. & Grohmann, K. 2005. *Understanding minimalism*. Cambridge: CUP.

Julien, M. 2002. Determiners and word order in Scandinavian nominal phrases. *Studia Linguistica* 56: 264–314.

Julien, M. 2005. *Nominal phrases from a Scandinavian perspective.* [Linguistik Aktuell/Linguistics Today 87]. Amsterdam: John Benjamins.

Kayne, R. 1994. *The antisymmetry of syntax.* Cambridge MA: The MIT Press.

Leung, T. 2001. An optimality-theoretic approach to Cantonese/English code-switching. MA thesis, University of Hong Kong.

Liceras, J.M., Fernández Fuertes, R. & Perales, S. 2008. Gender and gender agreement in the bilingual native and the non-native grammar: A view from child and adult functional-lexical mixings. *Lingua* 118(6): 827–851.

Lipski, J. 1985. *Linguistic aspects of Spanichs-English language switching.* Tempe AZ: Arizona State University, Center for Latin American Studies.

MacSwan, J. 1997. A minimalist approach to intra-sentential code-switching: Spanish Nahuatl bilingualism in central Mexico. PhD dissertation, University of California, Los Angeles.

MacSwan, J. 1999. *A minimalist approach to intra-sentential code-switching*. New York NY: Garland.

MacSwan, J. 2000. The architecture of the bilingual language faculty: Evidence from intrasentential code switching. *Bilingualism: Language & Cognition* 3(1): 37–54.

MacSwan, J. 2005a. Codeswitching and generative grammar: A critique of the MLF model and some remarks on "modified minimalism". *Bilingualism: Language & Cognition* 8(1): 1–22.

MacSwan, J. 2005b. Remarks on Jake, Myers-Scotton and Gross's response: There is no "Matrix Language". *Bilingualism: Language & Cognition* 8(3): 277–284.

MacSwan, J. 2008. Programs and proposals in codeswitching research: Unconstraining theories of bilingual language mixing. Paper presented at the CLS Annual Meeting. Para-session on Codeswitching.

MacSwan, J. In press. *Grammatical theory and bilingual codeswitching*. Cambridge MA: The MIT Press

Mahootian, S. 1993. A null theory of code-switching. Ph.D. dissertation, Northwestern University.

Moro, M. In press. The semantic interpretation and syntactic distribution of determiner phrases in Spanish/English codeswitching. In *Grammatical theory and bilingual codeswitching*, J. MacSwan. Cambridge MA: The MIT Press.

Nartey, Jonas. 1982. Code-switching, interference, or faddism? Language use among educated Ghanaians. *Anthropological linguistics* 24(2): 183–192.

Panagiotidis, P. 2002. Pronominal nouns. In *Pronouns-Grammar and Representation*, H.J. Simon & H.Wiese (Eds), 183–203. Amsterdam: John Benjamins.

Pfaff, C.W. 1979. Constraints on language mixing: Intrasentential code- switching and borrowing in Spanish/English. *Language Volume* 55(2): 291–318.

Poplack, S. & Sankoff, D. 1984. Borrowing: The synchrony of integration. *Linguistics* 22: 99–135

Preminger, O. 2008. (Im)perfect domains: Yet another theory of syntactic movement. In *Proceedings of the 26th West Coast Conference on Formal Linguistics* (WCCFL 26), C.B. Chang. & H.J. Haynie (Eds), 402–410. Somerville, MA: Cascadilla Press.

Putnam, M. 2007. *Scrambling and the Survive principle* [Linguistik Aktuell/Linguistics Today 115]. Amsterdam: John Benjamins.

Putnam, M. & Stroik, T. 2009. Syntactic relations in Survive-minimalism. Ms, Carson-Newman College and the University of Missouri-Kansas City.

Roehrs, D. 2006. The Germanic noun phrase: Determiners MOVE into the determiner phrase. Ph.D. dissertation, Indiana University.

Santorini, B. & Mahootian, S. 1995. Codeswitching and the syntactic status of adnominal adjectives. *Lingua* 96: 1–27.

Scott, G-J. 2002. Stacked adjectival modification and the structure of nominal phrases. In *Functional structure in DP and IP: The cartography of syntactic structures,* Vol. 1, G. Cinque (Ed.), 91–120. Oxford: OUP.

Stenson, N. 1991. Code-switching vs. borrowing in modem Irish. In *Language contact in the British Isles. Proceedings of the Eighth International Symposium on Language Contact in Europe, Douglas, Isle of Man, 1988*, P. Sture Ureland & G. Broderick (Eds), *Linguistische Arbeiten* 238: 559–579. Tubingen: Niemeyer.

Stroik, T. 1999. The Survive principle. *Linguistic Analysis* 29: 278–303.

Stroik, T. 2000. *Syntactic controversies.* Munich: Lincom.

Stroik, T. 2009. *Locality in minimalist syntax.* Cambridge MA: The MIT Press.

Stroik, T. This volume. The numeration in survive minimalism.

Stroik, T. & Putnam, M. 2005. The lexicon at the interfaces. Paper presented at LASSO 35. Lubbock TX.

Svenonius, P. 2007. Interpreting uninterpretable features. *Linguistic Analysis* 33(3–4): 375–413.

te Velde, J. To appear. Using the survive principle to derive coordinate (a)symmetries. In *Towards of derivational syntax: Survive-minimalism,* M.T. Putnam (Ed.), Amsterdam: John Benjamins.

Vangsnes, Ø. 1999. The identification of functional architecture. PhD dissertation, University of Bergen.

White, L. 1993. *Second language acquisition and universal grammar.* Cambridge: CUP.

Woolford, E. 1983. Bilingual code-switching and syntactic theory. *Linguistic Inquiry* 14(3): 520–536.

Zamparelli, R. 1993. Pre-nominal modifiers, degree phrases and the structure of AP. *University of Venice Working Papers in Linguistics* 3(1): 138–163.

Zamparelli, R. 2000. *Layers in the determiner phrase*. New York NY: Garland.

Using the *Survive* principle for deriving coordinate (a)symmetries*

John R. te Velde
Oklahoma State University

This analysis examines the symmetries of coordinate structures, specifically how they can be generated in a minimalist, crash-proof grammar. I show that a phase-based model with selection of lexical items (LIs) before merge must have a matching operation across conjuncts, but this operation is prohibited by this model's own constraints. An alternative is presented that uses the *Survive* principle by which LIs are selected as needed for the merge operations of a coordinate structure. This selection process is guided and assisted by algorithms that map certain features from a leading conjunct to the next conjunct undergoing concatenation. With selection on demand and the mapping of features, coordinate symmetries can be generated that otherwise require global operations spanning all the conjuncts such as across-the-board movement. Additionally, the asymmetries that occur in coordinate structures are accounted for as consequences of additional mergers that do not require coordinate matching across conjuncts. Issues related to the limits of working memory can also be addressed.

1. Introduction

The development of generative syntactic theories within the minimalist framework has shown considerable progress in the last decade. Yet, some fundamental properties of coordinate structures remain a challenge. The most central property of coordination, symmetry, and its inverse, asymmetry, will be investigated here from the perspective of a minimalist, crash-proof grammar presented in Stroik (2009) and Putnam & Stroik (2008) called *Survive*-Minimalism. The derivational model they propose is outlined in Section 4. In Section 2 we review some important background to minimalist theory and coordinate structures, and in Section 3

*Without the enthusiastic support of Mike Putnam and Tom Stroik, especially Mike's Crash-Proof Grammar conference, this paper would not have gotten off the ground. I am indebted to the conference participants, especially John Hale, Daniel Seely, Tom Stroik and Jan-Wouter Zwart for their comments, but most of all, again, to Mike for his many helpful suggestions. The remaining errors and problems stem from me only.

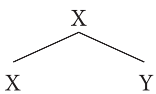
we consider a phase-based model as a point of reference for *Survive*-minimalism. In Section 4 we derive some coordinate structures using this model, with special focus on how it handles symmetry and asymmetry. In Section 5 we sum up the outcomes of Section 4 and list areas that need further research.

2. Some background

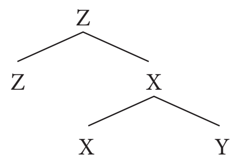
2.1 Assumptions about Merge, structure building and coordinate (a)symmetry

In the Minimalist Program of Chomsky (1995, 2000, 2005) the syntactic structure of a construction does not exist when lexical items (LIs) are selected from the lexicon. Rather, it is created when LIs merge in parallel, beginning at "the bottom"; further merge operations build upward from this initial structure:

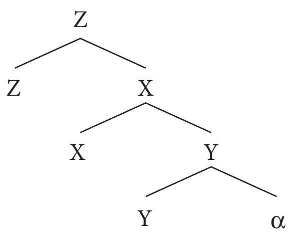
(1) a. after one merge operation



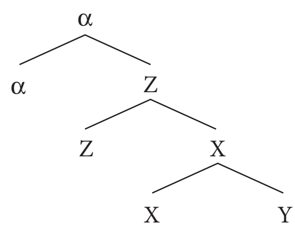
b. after a second merge operation[1]



c. conjuncts Z and Y sharing α



d. the same sharing relation, reverse ordering



This type of structure building and the syntactic relations that result from it directly predict certain properties of coordinate structures, for instance that two conjoined XPs, such as Z and Y in (1c,d), are not symmetric with respect to an element external to the coordinate structure, such as α in (1c,d), with X representing the coordinating conjunction (cf. work by Munn 1987; Johannessen 1998, among

1. The merge operation that assembles the structures in (1) has another property of asymmetry in that either one or the other of the categories can project when merge occurs. Here I have chosen the left one for projection.

others). Support for the assumption that relations between conjuncts is asymmetric like those in (1c,d) comes from the binding relations in (2), or the subject-verb agreement in (3):[2]

(2) a. Sam_i and his_i dog went for a walk in the park (cf. (1c))
b. *His_i dog and Sam_i went for a walk in the park

(3) a. There is a man and his dog in the kitchen (cf. (1d))
b. *There are a man and his dog in the kitchen

Other properties of coordinate structures, most notably their symmetries, are not predicted by this kind of structure building, if we assume that these symmetries are related to the structure in which they are found (cf. Camacho 2000 for a structure-based account of symmetric tense features):

(4) a. Sam and his dog walked.PAST in the park and visited.PAST their friends
a′. *Sam and his dog walked.PAST in the park and visit.PRES their friends
b. The park and the street both offer good walking areas
b′. #The park and the bench both offer good walking areas
c. *Der Park und* *(*die*) *Bank bieten beide gute Sitzplätze* (German)
the.MASC park and the.FEM bench offer both good sitting-places
c′. *Der Park und* (*der*) *Rasen bieten beide gute Sitzplätze* (*Rasen*=lawn)
d. The boys and (the) girls like to play in the park too
d′. The man and (the) woman enjoy sitting in the park

As (4a,a′) indicate, tense features must be symmetric; in (4b,b′) the semantic features of the DPs sharing a verb+complement must be symmetric. The comparison between (4c) and (4c′) indicates that the sharing of a determiner in German requires a gender symmetry. In (4d,d′) the determiner may or may not be omitted, depending on the intended reading; thus, the interpretation depends on the presence or absence of sharing, which in turn depends on the presence of symmetric features.

Adding to the complexity of coordinate symmetry vs. asymmetry, there are constructions in V2 languages in which an asymmetric agreement relation is tolerated when the verb has fronted ahead of the two conjoined DPs that it agrees with, as in (5a); this asymmetric agreement is ungrammatical, however, in the SVO configuration in (5a′). Asymmetric agreement is required with Comp-agreement

---

2. This observation goes back at least as far as early work of Munn (1987) and forms the core thesis of Johannessen's (1998) study of coordination. Alternate phrase structures can also produce the same results w.r.t. the asymmetry.

in (5b), while symmetric agreement is required when verbal agreement occurs in (5b′), both Frisian, from van Koppen (2003):

(5) a. *So verrauschte Scherz und Kuss*[3] (German)
so dissipated.3SG joke and kiss
'In this way joking and kissing dissipated'

a′. *Scherz und Kuss verrauschten/ *verrauschte so* (German)
joke and kiss dissipated.3PL/ dissipated.3SG so

b. *Ik tink dat-st [do en Marie] dit wykein yn R west ha*
I think that.2SG [you.2SG and M].PL this wkend in R been have.PL
'I think that you and Marie were in Rome this weekend'

b′. *Ha /*Ha-st [do en Marie] dit wykein yn Rome west?*
Have.PL /Have.2SG [you.2SG and M].PL this weekend in R. been
'Were you and Marie in Rome this weekend?'

As the data indicate, both symmetric and asymmetric agreement relations are possible in coordinate structures, and the configurations in which each occurs is not 100% consistent, such that the one always occurs in one type of configuration and the other in the inverse configuration. These findings undermine the hypothesis that symmetry and asymmetry are always structurally determined.

2.2 Further data indicating coordinate symmetry

In addition to the lack of a reliable correlation between structure and (a)symmetry, a number of symmetries in coordinate structures are clearly not related to structure, and thus not to the derivational principles and operations of the narrow syntax per se. These include the symmetries resulting from the matching or pairing of lexical features in corresponding conjuncts, as in (4b,d,d′). Further examples are given in (6):

(6) a. #The woman and the shopping bag went to the mall
a′. The woman and her shopping bag always go to the mall together
b. #The car sped down the street and leaped over the railing
b′. The boy sped down the street and leaped over the railing

In (6a) the conjunction is ill-formed because the two subjects do not both have a semantic feature, possibly [+volitional], required of the subject of 'went'. In (6b) the two verbs cannot be conjoined and share the subject 'the car' because this subject is not capable of leaping.

The features in (6) that create semantic symmetry are part of the feature matrix of the respective LIs when selected from the lexicon and are thus not created or valued

3. J. W. Goethe, from "An den Mond."

in the narrow syntax. Any derivational grammar, minimalist or otherwise, must account for the matching of the lexical – especially the semantic – features required for well-formed conjunction, as often determined by whether they can share a lexical item, that is, co-exist in parallel agreement relations between two or more conjuncts and a single element. This parallel co-existence in matching coordinate structures requires what will be referred to here as coordinate feature matching.[4]

In the next section we turn to some minimalist accounts of coordinate symmetry that have Phase Theory as a basis. We will use these as a reference point for illustrating the advantages that the *Survive* model outlined in Section 4 offers for deriving coordinate (a)symmetries.

3. Phase-based accounts of coordinate symmetry

3.1 Some assumptions of Phase Theory

We consider first an assumption about the selection of LIs from the lexicon in the phase models. In the previous section we noted that lexical features come with the LI when it is selected, regardless of the model used. Presumably the lexical features of LIs selected for conjuncts must be scanned at selection in order to determine whether they have the features needed for coordinate symmetry, as required by an agreement relation they enter into with a shared element, as illustrated in (1) – (6).

The exact procedure by which LIs are selected can make a significant difference as to how they are matched for coordinate symmetry. Chomsky (2000: 100–101) makes the following assumption about selection in the context of considering ways to address computational complexity and optimal design:

> Is it also possible to reduce access to Lex [as with [F], features], the second component of the domain of L? The obvious proposal is that derivations make a one-time selection of a *lexical array* LA from Lex, then map LA to expressions, dispensing with further access to Lex. That simplifies computation far more than the preceding steps. If the derivation accesses the lexicon at every point, it must carry along this huge beast, rather like cars that constantly have to replenish their fuel supply. Derivations that map LA to expressions require lexical access only once, thus reducing operative complexity in a way that might well matter for optimal design.

4. The symmetry requirements of coordinate relations – whether syntactic or semantic, but especially the syntactic symmetries – suggest that the coordinate structure itself is also symmetric. This point has been discussed extensively in the literature, with studies generally leaning one way or the other, i.e., either arguing that coordinate symmetries are all semantic and advocating syntactic asymmetry (the leading example Johannessen 1998), or attempting to accommodate them in a syntactic account that explains the symmetries in terms of syntactic relations or operations (Camacho 1997, 2000; Johnson 2002, among others).

If access to the lexicon is restricted in the way Chomsky proposes, then we must assume for coordinate structures that this selection process involves a matching mechanism so that the LIs of paired conjuncts that enter into an agreement relation with a shared element have the required symmetries, as in (6b′) 'the boy [sped… and leaped…]'.[5] A certain complexity, or at the very least a division of labor, ensues with this one-time selection: it can only guarantee matching *lexical* features; the symmetry of the verb tense in (6b′) must be guaranteed in the narrow syntax. Thus, Chomsky's strategy for constraining the grammar results in a two-layered coordinate feature matching procedure. In Section 4 we will consider whether this is the best strategy.

Let us turn now to an assumption about phases themselves. When a phase is complete, it is transferred to the interfaces, thus relieving the narrow syntax of the burden of holding it in working memory until the rest of the derivation is complete. This "multiple spell-out" approach to handling derivations, as advocated particularly by Uriagereka (1999) and adopted by Chomsky (2005), raises a number of questions, particularly with regard to coordinate structures. Boeckx (2007) addresses several of them; he points out that the computational system must be able to retrieve previously spelled-out material to provide a complete, coherent surface string and thus can't ignore or delete spelled-out elements. Boeckx (2007: 418) gives Norbert Hornstein credit for pointing out that interfaces appear to examine the internal content of full representations for specific processes. For example, the semantic component needs to see multiply spelled-out chunks for pronominal binding, and "PF quite possibly needs full clauses to determine intonational patterns, such as falling intonation ending up with a declarative (statement) or rising intonation yielding an interrogative …" (Ibid.).

Accounts using phases and multiple spell-out also leave unanswered how certain coordinate structures can be derived. In a phase-based approach coordinate TPs need to be derived and spelled out individually and sequentially, unlike e.g., conjoined DPs or PPs which, since they are not phases, lend themselves to parallel derivation and possibly to matching as well (but see discussion in §3.2 of Citko's (2005) proposal for conjoined TPs sharing a *wh*-element). Structures made up of CP conjuncts have an additional phase; they first require the spell out of the *v*P phase, creating a TP, and then the spell-out of the CP phase. At this point LIs for the second CP conjunct are extracted as subarrays, first the elements for the *v*P phase (which is then spelled out) and then those for the CP phase. After spell-out, however, the first *v*P phase of the first CP conjunct is no longer available for

5. Stroik (2008) discusses at length – apart from any considerations regarding coordinate symmetry – the implications of "blind" selection from the lexicon. He argues that it results in an extremely high number of merge possibilities that are neither manageable by C_{HL}, nor desirable in a minimalist grammar.

matching with the second *v*P phase. Matching is the basis for valuing the syntactic features so that the required coordinate symmetries are generated. Evidence that conjoined TPs must meet certain syntactic symmetry requirements comes from constructions like those in (7):

(7) a. First Sam walked.PAST the dog, then he cooked.PAST dinner
a′. *First Sam walked.PAST the dog, then he cooks.PRES dinner
b. Sam prepared$_i$ the main dish, and Sue e_i the dessert
b′. Sam prepared the main dish, and Sue *e (=baked) the pie
c. Their guests generally prefer e_i, and Sam and Sue also enjoy [wine with the meal]$_i$
c′. #Their guests prefer e_i, but Sam and Sue never offer [wine with the meal]$_i$

The symmetry requirement in (7a) can be narrowed down to the tense inflection and the ϕ-features, both of which must be valued in the narrow syntax. In (7b) the symmetry involves not only tense and ϕ-features, but also the lexical features of the two finite verbs. In (7c) the match arguably involves more than one LI, i.e., the two finite verbs 'prefer' and 'enjoy' as well as the adverb 'also', and the shared elements in brackets must also be part of matching. As the ill-formed (7c′) makes quite clear, matching must be global, i.e., essentially all of the LIs of the first conjunct must be matched with those in the second.

Earlier we noted that the matching of lexical features must precede the narrow syntax in a phase-based model, if selection of the LIs proceeds as Chomsky (op. cit.) proposes. This type of selection eliminates the possibility of matching lexical and syntactic elements globally. Syntactic features must be matched separately in the narrow syntax, and for this matching, the relevant features of the conjuncts must be present in the narrow syntax. But if the conjuncts are TPs or CPs, then multiple spell-out has eliminated the first TP or CP conjunct from narrow syntax before the next one is derived and thus makes it impossible to match the required syntactic features before transfer to the interfaces, at which point unfulfilled feature matching for coordinate symmetry causes a crash.[6]

Conjuncts unaffected in this way by multiple spell-out, for instance DP conjuncts such as those in (4), also present a problem for a phase-based grammar without either the adoption of certain assumptions about phrase structure and syntactic categories (Camacho 2000) or the operation Copy (Frazier & Clifton 2001). In the next section we turn to a proposal for an alternative phrase structure resulting

6. Matching in the narrow syntax in a derivational grammar may not be possible without major modification of assumptions on phrase structure and representation. In essence, the grammar would need to accommodate 3D or parallel representations along the lines of proposals by Grootveld (1994), Moltmann (1992), Muadz (1991) and Wesche (1995), or it would need to assume a form of Copy, such as proposed by Frazier & Clifton (2001) and Frazier (2008). The question remains: How derivational and crash-proof are such grammars?

from an assumption about merge as an example of a strategy within a phase-based model to address a type of coordinate symmetry.

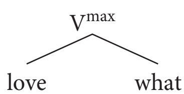
3.2 Citko (2005): Alternative way to merge SOs

Addressing a construction type that has inspired a long history of syntactic analysis, Citko (2005) proposes "Parallel Merge" to generate the relations that exist in across-the-board (ATB) *wh*-questions of the sort in (8), which require coordinate sharing – and therefore coordinate symmetry – not unlike the constructions considered earlier:

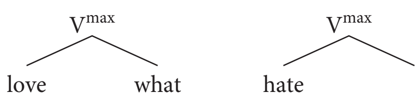
(8) What did Sam love *t* and Sally hate *t*?

Citko's Merge and Parallel Merge applied in sequence results in the multiple-dominance and sharing relation in (9b) so that *what* is dominated by two nodes and shared by two verbs:

(9) a. Merge *love* and *what*; project *love*:

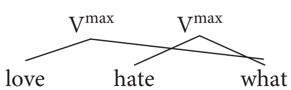


b. Parallel Merge *hate* and *what*; project *hate*



The linearization of the elements in (9b) requires crossing branches:[7]

(10)



The merit of Citko's proposal is that Parallel Merge requires no new operations or principles; it simply combines the properties of Internal and External Merge (IM and EM) proposed by Chomsky (2005). It is like EM in that it involves two distinct rooted objects; it is like IM in that it combines the two by taking a subpart of one of them.

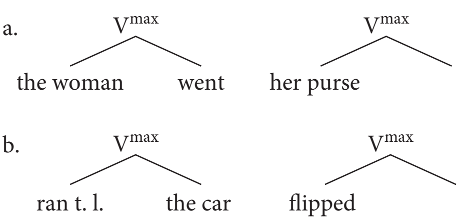
7. Citko points out that the relations resulting from Parallel Merge do not cause problems for the linearization requirements of syntactic derivation, as formulated by Kayne (1994) in his Linear Correspondence Axiom (LCA), because the LCA, following Chomsky (1995), does not apply until the derivation reaches the SM interface, i.e., its affects remain invisible. As long as Internal Merge (IM) repositions the shared object *what* before the SM interface, the derivation will not crash.

Thus, Citko's proposal can account for ATB *wh*-questions, if we assume that "taking a subpart" and Parallel Merge in general are operations that conform to assumptions about Merge and Copy.[8] Even if we assume this is the case, Citko's proposal would still require an additional linearization rule for constructions such as those seen earlier with conjoined DPs that do not have IM available to them for lining up the LIs at the PF interface. These constructions do not require any overt displacement like *wh*-movement to linearize the LIs; therefore, the only option for linearizing them after Parallel Merge has generated the sharing relations would be to propose some kind of NP-movement operation that accomplishes this. However, the generative literature provides no arguments or evidence that a realignment operation is necessary in the derivation of, for instance, the constructions in (11):

(11) a. The woman and her purse went to the mall
b. The car ran the light and flipped over the railing

In other words, what is transferred to the PF interface follows directly from the Merge operations that construct (11a,b); that is, neither the subject-verb relation in (11a) nor the verb-object relation in (11b) requires an overt realignment that is comparable to the *wh*-movement required in (8). Yet this is what would be required if Merge and Parallel Merge constructed (12a,b); the result would be:[9]

(12) a. b.



A point left unaddressed by Citko is how Parallel Merge is unified with transfer to the interfaces. Presumably (8) requires more than one phase. Thus, Parallel Merge must somehow "cancel" transfer until the merge operation is complete. This appears to be an ad hoc mechanism.

8. Chomsky (2007) states (n. 10): "Citko argues that parallel Merge is 'predicted' as IM is, but that is not quite accurate. It requires new operations and conditions on what counts as a copy, hence additional properties of UG."

9. An objection to this argument might be that in the narrow syntax there is no linear relation between the SOs. We can counter this objection, however, following this line of reasoning: Merge holds to the head parameter of a particular language. If Merge were not sensitive to this parameter, i.e., were completely non-linear, then overt movement (IM) would be required every time orderings like *went – her purse* and *ran the light – the car* were transferred to the PF interface. Presumably such orderings are just as likely as their reverse in a completely non-linear narrow syntax that is insensitive to the head parameter of the language for which it is being used.

These structures would require further linearization before they could meet the PF interface; yet, unlike the linearization of the structure (9b) in (10), there would be no syntactic motivation for it.

Beyond these limitations of Citko's proposal, there is the more general one that it has no operation or mechanism that matches the features of conjoined SOs; thus, it is not able to assure the symmetry of syntactic inflections as required, for instance, with the finite verbs in (11b), or the symmetry of lexical features in (4) and other constructions we have seen. Thus, limitations and problems remain with even this "enhanced" phase-based approach.

We turn in the next section to the *Survive* model of Stroik (2009) and Putnam & Stroik (2008) to investigate whether the problems of the Citko proposal and other phase-based models can be avoided.

4. The *Survive* principle and its application to coordinate (a)symmetry

4.1 The *Survive* model and its advantages

A single principle guides the concatenation of SOs in the *Survive* model without resort to IM or phases, which Stroik and Putnam argue do not make a minimalist grammar more strictly derivational or crash-proof. For these and a number of other reasons, they employ a version of Merge that may target a given LI more than once, depending on its feature matrix. For instance, in (11), *who* is targeted by Merge three times (i.e., is remerged twice) because it has three features, each of which must be checked by a different head. Each time a SO has a remaining unchecked feature, it "survives" and remains active in the numeration and remerges for checking (from Putnam 2007: 14 – 15):

(13) Who cares?
  a. Merge {who, cares} → who cares
  b. Survive {who}
  c. Merge {T, {**who**, cares}} → T **who** cares
  d. Remerge {who, {T, {**who**, cares}}} → who T **who** cares
  e. Survive {who}
  f. Merge {C, {**who**, {T, {**who**, cares}}}} → C **who** T **who** cares
  g. Remerge {who, {C, {**who**, {T, {**who**, cares}}}}} → who C **who** T **who** cares

In (13) *who* first merges with *cares* (in a) for checking its θ-feature. It survives and remerges (in d) to check its ϕ-features with T and survives again. Finally, it remerges a second time (in g) to check its Q-feature with C. Note that other merge operations intervene between the remergers of *who* to make available the heads that check the remaining features of *who*. In this way a syntactic structure is built from the bottom up, as are structures using Phase Theory. Note, however, that the entire numeration remains active until all features have been checked, i.e., nothing is transferred to the interfaces until the entire derivation is completed in the narrow syntax.

What advantages does this type of merge operation offer for the derivation of the coordinate structures we have seen? We recall that a major shortcoming of phase-based approaches for deriving coordinate structures was the lack of a clear derivational or checking mechanism for coordinate symmetries, either at selection (for lexical features) or in the narrow syntax (syntactic features); in fact, the one-time selection of LIs for merge and the multiple spell-out of merged structures create major obstacles to the matching necessary for coordinate symmetries, as we have seen, to the detriment, in my estimation, of fundamental minimalist objectives. The *Survive* model's method of selection and feature checking provides a possible solution to these problems. For considering this method, some additional aspects of a typical derivation in the *Survive* model that aren't apparent in (11) must be outlined.

In contrast to the form of selection described by Chomsky (2000: 100) cited above, the *Survive*-model uses an "on-demand" or dynamic selection process, in contrast to one-time selection, that does not require extraction of LIs from a lexical array or subarray when they are merged. Rather, the LIs to be merged are selected directly from the lexicon. Furthermore, when merged, mapping algorithms are established for pronoun-antecedent relations and other long-distance relations that cannot be captured in a local feature-checking relation. Such mapping algorithms, along with dynamic selection, offer a means to first generate and then check coordinate symmetries. Consider the *Survive* derivation of the German construction *Frank ruinierte den Laptop, aber er behielt das Zubehör* 'Frank ruined the laptop, but he kept the accessories' in (14) (each → is a selection from the lexicon; s.m. = semantic matrix; WB = WorkBench,[10]

10. "WorkBench" is where DPs are assembled in the Survive model. Putnam (2007: 99) explains the need for the WB this way: "...if a *Derivation* D is restricted to the computation of an event (expressed in the V-*v* projections) and its temporal and speech event coordinates (expressed in the T-C projections), then D must be built from two different sources of syntactic objects – the *Numeration* N and a workspace (WS) [= WB] not contained in D."

C = conjunct; c/c = compare/contrast):

(14) Deriving *Frank ruinierte den Laptop, aber er behielt das Zubehör*

LEXICON

a. Merge {Frank, ruinierte} → Frank ruinierte - - $[_{F}$ g, n], $[_{r}$ s.m.] - > - -

a′. Merge in WB {den, Laptop} → den Lt[11]

b. Survive {Frank} b′. Survive in WB {den}

c. Merge {Fr. ruinierte {den Lt}} → Fr. r-te den Lt - - $[_{DP}\ \theta_{th}]$[12]- > - - -

c′. Merge in WB {D {**den** Laptop}} → D den Laptop

d. Merge {*v*, {Fr. r-te den Lt}} → Fr. r-te den Lt- $[_{DP}$ ACC][13]- - - - - - -

e. Merge {T, {Frank r-te den Lt, }} → T Frank r-te den Lt

f. Remerge {**Fr**, **r-te**[14] {T Frk r-te den Lt}} - [ϕ, T] - > - - - - - - -

→ Frank ruinierte den Laptop

(C-2 begins; Merge as in C-1, with algorithmic input.)

g. Merge {er, behielt} → er behielt - - $[_{er}$ g, n], $[_{beh}$ (c/c Vs)]- - <- - - - -

g′. Merge in WB {das Zubehör}} → das Zubehör

h. Survive {er} h′. Survive in WB {das}

i. Merge {er behielt {das Zubehör}} → er behielt das Z-hör - < - - - -

i′. Merge in WB {D {**das** Zubehör}} → D das Zubehör

j. Merge {*v*, {er behielt das Zubehör}} → er bht das Z-hör - < - - - -

k. Merge {T, {er bht das Z-hör}} → T er bht das Z-hör

l. Remerge {**er**, **behielt** {T er bht das Z-hör}} → Er bht das Z <-

m. Merge {**aber**, {er behielt das Z-hör}}-(c/c C-1 & C-2[15])- - - - -

→ aber er behielt das Zubehör

n. Merge (conjoin): {Frank, r-te, den Lt, {aber, er, behielt, das Z-hör}}

→ Frank ruinierte den Laptop aber er behielt das Zubehör

11. The parsing formalisms differ somewhat from those in Putnam (2007). Boldface is used to single out an LI that (re)merges; thereafter it occurs in normal typeface; category labels are not repeated once they have been introduced into the numeration. In like manner, the "copies" of LIs remerged are not repeated after a remerger.

12. I leave open if and when the V2 structures have a clause-final [V].

13. In a more detailed outline, *Frank ruinierte den Laptop* would remerge in the *v*P after [*v*] has merged; the same with *er behielt das Zubehör* in C-2 (assuming the *v*P is required in the *Survive* framework).

14. [T] triggers the remerger of both *Frank* and *ruinierte* in C-1 and *er behielt* in C-2 because it has features for both agreement and tense. In a more detailed outlined, each LI would remerge separately.

15. Comparison and contrast are presumably required for the selection of *aber* 'but' as the coordinating conjunction.

The mapping algorithms indicated on the right with broken horizontal and solid vertical lines accomplish two things: (i) they guide selection from the lexicon by mapping the features of LIs in the numeration with the features of potential LIs for the second conjunct, and (ii) they map a particular feature from one conjunct that has completed derivation to another that is being assembled. Thus, algorithms assure certain symmetries and eliminate the randomness of selection that would otherwise generate numerations that crash at the interfaces.[16]

In the example construction, first the number and gender features of *Frank* and the semantic feature matrix of *ruinierte* 'ruined' must be matched with the same features of corresponding LIs selected for the second conjunct. Later, the ϕ-features of *Frank* and the tense features of *ruinierte* are mapped to their equivalents in C-2. The features of the LIs *er behielt* 'he kept' must match those of their counterparts in the first conjunct for grammaticality and semantic well-formedness; they represent the minimum required by coordinate symmetry.[17]

Likewise, the Case features of the DPs *den Laptop* and *das Zubehör* 'the accessories' must match for similar coordinate symmetry requirements. If one or the other DP had a Case inflection other than [ACC], the construction would be ungrammatical even without conjunction, if the verbs remained the same. Selecting another verb for the second conjunct that requires a dative object, for instance, results in non-matching Case features that render the construction ill-formed:[18]

(15)	*#Frank*	*ruinierte*	*den*	*Laptop,*	*aber*	*er*	*dankte*	*dem*	*Besitzer*
	F	ruined	the	laptop,	but	he	thanked	the	owner

Other verbs that require a dative object, such as *ähneln* 'resemble', *begegnen* 'encounter', *dienen* 'serve', *drohen* 'threaten', *fehlen* 'lack', etc. produce even worse results in this construction (some completely ungrammatical). Accordingly, both

16. Putnam and Parafita (this volume) outline how the *Survive* principle can be applied to the selection and merger of LIs in a grammar that involves code switching. Their proposal addresses simplex constructions and thus presents a picture of derivational operations that precede the assembly of coordinate constructions, explored here.

17. We note that in the *Survive* model there is no distinction between [+interpretable] and [–interpretable] features as to where what features are checked. This distinction is problematic for a number of reasons. In German, for instance, Case features cannot be [–interpretable] and thus eliminated from the numeration by checking because they must be realized as inflections in PF.

18. A great deal of contextualization might rescue this construction. Hence, we cannot assume that all conjunctions of DP_{ACC} and DP_{DAT} will be ill-formed or ungrammatical, i.e., inter-clausal or pragmatic factors/features can rescue, i.e., render adequately symmetric, some such conjunctions.

the verbs and their objects must be matched, as the mapping algorithms in (14) indicate, and both lexical and syntactic features must match. We note that in the *Survive* model both types of features are matched via a single mapping algorithm.[19] Thus, the division of labor seen earlier in the phase-based model does not exist in this model; this simplification of the grammar model has the advantage for coordinate feature matching of allowing a single matching operation for both feature types. Furthermore, the mapping algorithms, because they extend from one clausal conjunct to another, make global matching feasible.[20] The lack of phases and multiple spell-out in the *Survive* model means that the problems associated with them are avoided.

In the next section we turn to advantages of the *Survive* model related to computational efficiency.

4.2 Working memory and C_{HL} in the derivation of coordinate structures

As we have just seen, the *Survive* model makes no distinction between lexical and syntactic features when it comes to their matching for coordinate symmetry; both can be mapped via an algorithm during concatenation. In a phase-based model, the two types of features must be handled in their respective areas of the grammar; the lexical features must match at selection (since the narrow syntax cannot read them), and the syntactic features in the narrow syntax. This added complexity of a phase-based model comes in addition to the one already noted, namely that because all LIs are selected prior to computation in the narrow syntax, the selection of LIs for conjuncts must proceed in parallel planes or in 3D mode for matching to occur (cf. Moltmann 1992; Muadz 1991; Wesche 1995 for proposals along these lines). By contrast, in the *Survive* model with on-demand selection, an LI that constitutes a second (or third, etc.) conjunct is matched with the previous conjunct via a mapping algorithm when it enters the numeration. In a phase-based model, matching must occur again in the narrow syntax for the syntactic features of paired conjuncts to achieve the required symmetry.

---

19. An anonymous reviewer pointed out correctly that tense mismatches between verbs of conjoined clauses are generally more severe than ACC-DAT mismatches between conjoined DPs like those above. The same is not true, however, with NOM-ACC mismatches, which result in total ungrammaticality:

(i)	Der	Vater	ging	spazieren,	aber	sein/	*seinen	Sohn	spielte	Billard
	the	father	went	walking	but	his-NOM/	his-ACC	son	played	billiards

20. Whether an object is suitable for a verb selected must also be checked, but this can be accomplished when the verb and its object merge (in a local relation) and thus does not require an algorithm unless it involves a coordinate relation in which the second conjunct, for purposes of cooordinate symmetry, must have an equivalent verb-object relation.

This second matching procedure raises an important question: Where are the SOs of the first conjunct located when the second conjunct is derived? If both conjuncts are in the same phase, such as conjoined DP subjects, then nothing has been transferred to the interfaces when the second conjunct is assembled.[21] If, however, the conjuncts are located in two different phases, as is the case in (14), then matching between the interfaces and the narrow syntax – where the second conjunct, located in the next phase, is being assembled – must proceed. This kind of matching would require the derivation to look back from the interfaces into the narrow syntax, and thus by definition from one phase into another. Look-back, however, is prohibited in phase-based models, as formulated in Chomsky's (2000: 108) Phase Impenetrability Condition. This problem obviously does not occur in the *Survive* model since it does not employ phases.

There is nevertheless a problem for the *Survive* model that relates to computational efficiency in the derivation of coordinate structures:

(16) *Memory load problem*: How can multiple matrix-clause conjuncts be merged and matched, if all conjuncts remain active in the numeration (and are thus retained, i.e., not transferred to the interfaces) until derivation (and matching) is completed?[22]

This problem is the counterpart to the look-back problem of a phase-based model addressed earlier in that it is created by not transferring the derivation to the interfaces. We consider below how it is manifested in the German construction in (17):

(17) Multiple conjoined TPs with a shared *wh*-element (copies in bold)

Welches	*Buch*	*hat*	*Frank*	*gern*	*gelesen,*
which	book	has	F	gladly	read,
		Peter	*nicht*	*verstanden,*	
		P	not	understood,	
		Lars	*nicht*	*gekauft …*?	
		L	not	bought	

'Which book did Frank like reading, Peter not understand, Lars not buy…?'

21. Even if the two conjuncts are located in the same phase, the matching of syntactic features is not a trivial matter. Instead of matching, a copy operation might get the job done (cf. Frazier & Clifton 2001; te Velde 2005b).

22. This problem is not restricted to coordinate structures. Any sentence can theoretically be expanded by *n*-number of mergers of embedded clauses. The memory load problem with the derivation of embedded clauses can be dealt with uncontroversially by transferring clauses (or portions of them, phases) to the interfaces as they are merged, since they, unlike conjuncts, do not have to be retained in the numeration for matching purposes (though potentially other syntactic relations may prevent transfer). Related problems arise from the "bottom-up" approach assumed in much minimalist work: Either look-ahead or the management of working memory are required, see below.

[$_{CP}$ Welches.Q Buch.ACC [hat.3SG [$_{TP}$ Frank.NOM **hat**
[**welches Buch** gern gelesen
[$_{TP}$ Peter.NOM **hat**
[**welches Buch** nicht verstanden
[$_{TP}$ Lars.NOM **hat**
[**welches Buch** nicht gekauft]]]]]]]

Some preliminary observations: The phrase structure in (17) implies a reduced ATB derivation by which *welches Buch* 'which book' is fronted from each VP to the Spec,CP position. It implies a *reduced* ATB derivation in that this element does not make "intermediate stops" in the Spec,CP position of its own clause and then progress upward to the Spec,CP of the next clause, etc. until it reaches the highest one available. This derivation is not tenable in current minimalist syntax for at least two reasons: 1) It violates Phase Theory, which doesn't allow movement from one phasal unit to another without use of the edge (Spec,CP in this case); 2) It requires head movement for raising *hat* to the C^0 position; for principled reasons, head movement is generally assumed to be non-existent (cf. Chomsky 2001).

Parallel Merge in Citko's (2005) proposal would, with some minor modification, be able to derive this construction. However, this proposal, as we have seen, cannot derive some of the other constructions under consideration here; for this reason we will leave it aside here.

Deriving this construction using the *Survive* principle requires a certain assumption about bottom-up derivation, as it applies to coordinate structures of this sort. My line of argument is that a matrix clause (whether a TP or a CP) defines the limit of a derivation. For the conjunction of matrix clauses this means that the "bottom" of the derivation is the bottom of the first matrix clause in a string of conjoined matrix clauses. No condition or rule needs to be formulated for this assumption, as stated in (18), to apply in the *Survive* model:

(18) A matrix clause defines the limits of a single derivation; the conjunction of *n*-number of matrix clauses consists of *n*-number of such derivations that thereafter become the syntactic object of 'conjunction', a "complex" derivational operation that merges the output of a simplex derivation with the output of a preceding symmetric simplex derivation.[23]

23. This assumption addresses the dilemma faced by the bottom-up derivation of coordinate structures: If merge begins at the bottom of constructions like those in (i), then look-ahead is required into the previous clause:

(i) a. Sam bought the beer, but he forgot the milk
b. The car sped down the street, and then it flipped over the railing

For the derivation of (17), this means that first the matrix clause *Welches Buch hat Frank gern gelesen* 'Which book did Frank like reading' will be derived first (cf. (13) for an example of how a *wh*-question is derived using the *Survive* principle). The entire sequence of derivations for (17) looks like this:

(19) Sketch of the derivations required for (17) using the *Survive* principle

a. Derive *Welches Buch hat Frank gern gelesen* a la (13) with the addition of WB merger for the DP *welches Buch*

b. Derive ***Welches Buch hat*** *Peter nicht verstanden* with **copies** of *welches Buch* and *hat* mapped to the CP domain via an algorithm

c. Conjoin the merged matrix clauses[24]

d. Repeat b for the next conjunct ***welches Buch hat*** *Lars nicht gekauft*

e. Conjoin with the previous clauses

f. Elide at the left edge, eliminating the redundant CP domains[25]

The copies of *welches Buch* 'which book' and *hat* 'has' mapped to the CP domain of the second and third conjuncts contain all the syntactic and semantic features of the originals. Lacking are the (morpho-)phonetic features, which have been "deleted" by "coordinate ellipsis," an economy-based operation that makes non-realization of phonetic features possible when the deletion site is c-commanded by [&] and recovery is possible in LF when the coordinate symmetry requirements have been met. A more detailed discussion would take us too far afield, and is already available in Chapter Four of te Velde (2005b). Relevant here is that remerge is not required for this coordinate ellipsis operation itself. Further work is needed to determine whether this operation is fully compatible with derivation by (re)merge as conceived in *Survive*-minimalism.[26]

If derivation proceeds as proposed in (18), then look-ahead is coordinated with the logic and temporal sequence of the construction. See also the discussion of *Link!* in §4.3 and 5.

24. Since this construction is a hypothetical, "limitless" conjunction of CPs, no coordinating conjunction is indicated and need not be merged as long as the appropriate prosodic features are in place.

25. For an analysis of left-edge deletion in the conjunction of CPs and a proposal using a phase-based approach, see te Velde (2005a,b). This proposal would be compatible with the *Survive* model on the assumption that the mapping algorithms established at (re)merge for coordinate matching purposes are also suitable for the "deletion" (marking for non-realization in the PF component) of redundant elements. Further work is needed.

26. It is important to note that the derivation of each conjunct, with all the required mergers for θ-role and Case assignment and subject-verb agreement, are completed in each conjunct, with the relevant features mapped to the next conjunct as in (14) for purposes of coordinate symmetry.

In order for conjunction in steps c and e to proceed, it is necessary to have all of the clauses merged up to those points active in the numeration, if we assume that the conjunction of matrix clauses is a syntactic merge operation. If it were not, it remains a puzzle how a coordinating conjunction could be merged, which is required before the last of a string of conjuncts in German, English and many other languages. But assuming that conjunction is a syntactic merge operation brings us back to the question posed in (16): How can *n*-number of conjuncts (of *any* size) be managed in working memory? The dilemma of managing a potentially infinite number of conjuncts while maintaining conjunction as a syntactic merge operation forces us to look beyond the syntactic mechanisms of the *Survive* model presented so far. This is the objective of the next subsection.

4.3 *Link!*

The primitive function *Link!* in the *Survive* model is described as follows:

(20) *Link!* (Putnam 2006: 6)
If XP is a syntactic object (SO) bearing the interpretable feature α, and a syntactic head Y^0 enters the derivation bearing the interpretable feature α, XP will remerge into the structural position immediately dominating Y^0 [SpecYP] at which point the matching features will amalgamate and be interpreted at the interfaces.

The central idea with *Link!* is to have features and concatenated SOs be checked for interpretability as soon as appropriate configurations arise in the narrow syntax. One of these "appropriate configurations" in (19) is the first CP derived. Upon completion of its derivation, *Link!* determines on the basis of the Q-feature and the finite verb that it can be interpreted as a matrix *wh*-question. *Link!* thus brings the derivation closer to conjunction in that its output is identifiable as a SO that can be conjoined with another of its kind that is concatenated immediately thereafter.

Link! does not by itself solve the working memory dilemma, however. For that *Link!* must extend beyond the interpretive component, call it LF for convenience, to a new area of memory, call it short-term memory. I will assume here that short-term memory is the area of human memory where meaningful (interpretable) utterances can be stored, in contrast to working memory where linguistic computations are performed. If C_{HL} indeed has access to short term memory via *Link!*, then working memory is in fact relieved of a burden with this linkage. Further research will need to verify whether this transfer to short-term memory actually occurs. If so, a potential answer to the question in (16) is available.[27]

27. The transfer of a phase to LF (the Conceptual-Intentional Interface in Chomsky's terminology) also relieves working memory of a burden. The difference between the phase-based

Relieving working memory of the burden of holding a matrix clause until another is derived is of no help to the conjunction of the matrix clauses, however, unless the first matrix clause, now in short-term memory, can be recalled by the narrow syntax for conjunction purposes. *Link!* as outlined above does not provide for this kind of recall. For this, a system is needed in which transfer goes both directions between the narrow syntax and the interfaces, i.e., they are symmetric in this regard. Boeckx (2007) discusses a system in which Spell-Out is eliminated and the interfaces have access to the narrow syntax and are not passive recipients of information. He points to a symmetric relation between the narrow syntax and the interfaces that is broken – made asymmetric – or "hidden by time, i.e., disrupted by derivational processes" (Ibid., p. 422). This is precisely the situation with the derivation of conjoined matrix clauses: The derivation extends over the entire coordinate structure, but it is broken by the time it takes to assemble each clause, with each read by the interfaces at the appropriate point. What is needed to make this system ideal for the derivation of coordinate structures is a recall ability; that is, the narrow syntax is able to take advantage of the fundamental symmetry in the narrow syntax-LF system and recall features as needed for coordinate feature matching. This operation is part of one derivation that consists of subparts, matrix clauses that stand in a coordinate relation to each other, that is broken or hidden by time, i.e., by the derivational processes involved with each one, but emerging in the end as one construction by conjunction. Several non-trivial assumptions about the principles of derivation – most notably look-back – must be investigated here.

4.4 Coordinate asymmetry in the *Survive* model

So far we have considered only how the symmetries of coordinate structures can be handled in the *Survive* model. As Johannessen (1998) has made clear, asymmetries occur quite commonly in coordinate structures and cannot be ignored.[28] Some simple examples of asymmetric subject-verb agreement that resemble many

model and the *Survive* model is that the problem with executing the matching required for coordinate symmetry – accomplished via mapping algorithms in the *Survive* model – has not been eliminated in the phase-based model with the alleviation of the working memory burden.

28. I differ with Johannessen on the premise that asymmetry is the defining property of coordinate structures, i.e., that a theory of coordination should revolve around asymmetry. It is precisely the symmetry of coordinate structures that distinguish them from subordinate structures. This distinction does not force us, however, to use 3D or parallel structures that become highly problematic for syntactic theory. My approach here and in earlier work (te Velde 2005b) is to capture coordinate symmetries within an essentially asymmetric phrase structure.

others because of their C/VSO configuration are given in (5), repeated in (21) (a, German, b, Frisian, from van Koppen 2003):

(21) Examples of asymmetric coordination in C/VSO configurations

a. *So verrauschte Scherz und Kuss*[29]
so dissipated.3SG joke and kiss
'In this way joking and kissing dissipated.'

b. *Ik tink dat-st [do en Marie] dit wykein yn Rome west ha*
I think that.2SG [you.2SG and M].PL this weekend in R. been have.PL
'I think that you and Marie were in Rome this weekend.'

Van Koppen (2005, 2006) analyzes asymmetries like the one in (20b) as a result of a probe-goal relation between the complementizer – which has inflection for agreement – and the first subject. Using the *Survive* model, we can assume a very similar relation, but one that does not induce Internal Merge, as does van Koppen's. Rather, the merging of the complementizer *dat* 'that' in Frisian and the remerging of the first subject *do* 'you' results in a checking relation; *do* 'you' survives for remerge and checking with *dat* 'that' because both have ϕ-features. The second, conjoined subject *Marie* does not require a checking relation with *dat* 'that' because it has no such features. In the event there is no Comp-agreement because a verb sits in C^0 as in (22a), a symmetric agreement relation is required, and as (22b) indicates, the SVO configuration also requires symmetric agreement (both examples are Frisian):

(22) a. Ha /*Ha-st [do en Marie] dit wykein yn Rome west?
Have.PL /Have.2SG [you.2SG and M].PL this weekend in Rome been
'Were you and Marie in Rome this weekend?'

b. [Do en Marie] ha /*ha-st dit wykein yn Rome west
[you.2SG and M].PL have.PL /have.2SG this weekend in Rome been
'You and Marie were in Rome this weekend.'

Other West Germanic data indicate that the VSO configuration doesn't guarantee either symmetry or asymmetry; rather, the agreement relation and features chosen for checking determine this, i.e., there is an optionality involved:

(23) a. In den Wald ging /gingen [der Jäger und sein Hund] (German)
into the wood went.SG /went.PL [the hunter.SG and his dog].PL

29. J.W. Goethe, from "An den Mond." The *-te* ending on *verrauschte* also marks [PAST], ignored here for simplicity's sake.

b. In de kamer kom /komen [ik en Jan] vaak (Dutch)
into the room come.SG /come.PL [I and J].PL often

We must keep in mind that the SVO equivalents of (23) require symmetric agreement:

(24) a. [Der Jäger und sein Hund] gingen /*ging in den Wald
[the hunter.SG and his dog].PL went.PL /went.SG into the wood

b. [Ik en Jan] komen /*kom vaak in de kamer
[I and J].PL come.PL /come.SG often into the room

In American dialects SVO does not guarantee symmetric agreement, however, for reasons that can be debated (see van Gelderen 1997 for an account); in general, non-prescriptive, colloquial speech in English favors asymmetric subject-verb agreement in coordinate structures. It appears that because English inflectional agreement morphology is either impoverished or non-existent, asymmetry results, given the lack of inflections to make a syntactic symmetry clearly apparent. In other words, what features are available and what speech registers or prescriptive rules come into play determine the symmetry or asymmetry of the agreement.[30]

Other West Germanic data that need to be considered include verb-object and preposition-object agreement that have inflections sufficiently rich to indicate symmetry or asymmetry. The data are too extensive to be investigated here (see te Velde 2005b for discussion), but one generalization can be made: When a feature-checking head precedes conjoined DPs, asymmetry is sometimes tolerated, at times even required, as in Frisian constructions with Comp-agreement like (21b). The reverse configuration shows less asymmetry; in languages like German with relatively rich inflectional systems, asymmetry in the [DP & DP]–head configuration is ungrammatical.

This limited amount of data is insufficient for coming to final conclusions, but a picture begins to emerge: rich inflectional agreement systems favor symmetry in the configuration in which the checking head surfaces *after* the conjoined DPs. This correlation suggests that for a merge system like the one used in the *Survive* model, the remerging of a checking head in a position preceding the conjoined elements with which it agrees can result in the breakdown of symmetry. This breakdown is predicted by the fact that the remerger of a head in this position results from an *additional* checking requirement. For instance, in (21b) with Comp-agreement in which the asymmetric, Comp–first conjunct agreement relation is required, this additional checking requirement – the remerger of the coordinate structure [*do en Marie*]

30. Emonds (1986) provides an interesting study on the role of sociolinguistic factors in the choice of (a)symmetric agreement in coordinate structures.

'you and Marie' in the CP domain after feature checking with *ha* 'have' in the VP[31] – is induced because both Comp and the first conjunct *do* 'you' have the same feature in Frisian.

What exactly the make-up of the feature for Comp-agreement is must be determined on the basis of data from Germanic dialects that require it. In (17) the remerger of the finite verb in C^0 is required for checking the Q-feature present on the interrogative in Spec,CP; the same checking occurs for the same reasons in (22a), with the difference that in the latter no lexical element, just this feature, occupies Spec,CP. The feature that induces the remerger of the finite verb in C^0 in constructions like (23a,b) is obviously different, for here the remerger is optional. Hence we must assume that in the *Survive* model there are features that induce stylistic changes as well as the required syntactic operations. Given the flexibility of this model – we saw earlier that it makes no distinction between interpretable and uninterpretable features – "stylistic features" can presumably be selected and merged with SOs in the same way as required syntactic features. However, this is another area for further research.

5. Evaluation and further research

Several other aspects of this analysis leave open areas for further research. One of the most complex involves the precise make-up of the mapping algorithms required for the matching of coordinate features. Much has been left up to them, but further work is required to determine whether mapping algorithms can (i) accomplish the kind of matching required for coordinate symmetry, and (ii) handle both lexical and syntactic feature checking, both without resort to mechanisms that render the grammar non-derivational. This work relates to the division of labor in a derivational grammar and what has been proposed here for the *Survive* model: Is the "shared" labor – in the sense discussed earlier, i.e., that both syntactic and semantic features are handled in the narrow syntax by mapping algorithms – both welcome and feasible? Related to the nature of mapping algorithms is the question of whether they are also suitable for ellipsis, as assumed for the derivation outlined in (19).

Another area for further research is the investigation of whether *Link!* enables the transfer of derivational units via LF to short-term memory, thus relieving the narrow syntax of a memory burden (essentially two questions). The extension of

31. Presumably tense feature checking occurs in the TP domain as well, as it does in (22a) before *ha* raises to check the Q-feature in the CP domain.

Link! proposed here – a symmetric relation between the narrow syntax and LF enabling the recall of features as needed for coordinate feature matching – also requires more investigation.

Finally, the statement in (18) on 'conjunction' as a complex derivational operation that merges the output of simplex derivations must be made more concrete and precise. The prediction here is that this derivation concatenates matrix clauses as SOs in much the same manner as other derivations concatenate LIs. Nevertheless, the syntactic structure required and the status of the coordinating conjunction within it – especially what feature it has that must match with the conjunct it merges with – must be investigated further.

References

Boeckx, C. 2007. Eliminating spell-out. *Linguistic Analysis* 33(3–4): 414–425.

Camacho, J. 1997. The syntax of NP coordination. Ph.D. dissertation, University of Southern California.

Camacho, J. 2000. On the structure of conjunction. In *Ellipsis in conjunction*, K. Schwabe & N. Zhang (Eds), 23–50. Tübingen: Niemeyer.

Chomsky, N. 1995. *The minimalist program*. Cambridge MA: The MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In *Step by step*, R. Martin, D. Michaels & J. Uriagereka (Eds), 89-155. Cambridge MA: The MIT Press.

Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A life in language*, M. Kenstowicz (Ed.), 1– 52. Cambridge MA: The MIT Press, 134–166.

Chomsky, N. 2008. On phases. In *Foundational issues in linguistic theory*: *Essays in honor of Jenn-Roger Vergnaud* R. Freidin, C.P. Otero & M.-L. Zubizaretta (Eds), Cambridge MA: The MIT Press.

Chomsky, N. 2007. Approaching UG from below. In *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, U. Sauerland & H.-M. Gärtner (Eds), 1–29. Berlin: Mouton de Gruyter.

Citko, B. 2005. On the nature of Merge. External Merge, Internal Merge, and Parallel Merge. *Linguistic Inquiry* 36(4): 475–496.

Emonds, J. 1986. Grammatically deviant prestige constructions. In *A festschrift for Sol Saporta*, M. Braeme, H. Contreras & F. Newmeyer (Eds), 93–129. Seattle WA: Noit Amrofer.

Frazier, L. & Clifton, C., Jr. 2001. Parsing coordinates and ellipsis: Copy α. *Syntax* 4(1): 1–19.

Frazier, L. 2008. Processing ellipsis: A processing solution to the undergeneration problem? In *Proceedings of WCCFL* 26, C.B. Chang & H.J. Haynie (Eds), 21–33. Somerville MA: Cascadilla Proceedings Project.

van Gelderen, E. 1997. *Verbal agreement and the grammar behind its breakdown*. Tübingen: Niemeyer.

Grootveld, M. 1994. *Parsing coordination generatively*. Den Haag: Holland Institute of Generative Linguistics.

Johannessen, J.B. 1998. *Coordination*. New York NY: OUP.

Johnson, K. 2002. Restoring exotic coordinations to normalcy. *Linguistic Inquiry* 33: 97–156.

Kayne, R. 1994. *The antisymmetry of syntax* [Linguistic Inquiry Monograph 25]. Cambridge MA: The MIT Press.

van Koppen, M. 2003. First conjunct agreement. Frisian, Tegelen Dutch and Hellendoorn Dutch. Paper presented at the Comparative Germanic Syntax Workshop 18, University of Durham, Sept. 18–20, 2003.

van Koppen, M. 2005. *One Probe – Two goals: Aspects of agreement in Dutch dialects*. Utrecht: LOT.

van Koppen, M. 2006. One probe – Multiple goals. Paper presented at GLOW 2006.

Moltmann, F. 1992. Coordination and comparatives. Ph.D. dissertation, MIT.

Muadz, H. 1991. Coordinate structures: A planar representation. Ph.D. dissertation, University of Arizona.

Munn, A. 1987. Coordinate structure and X-bar theory. *McGill Working Papers in Linguistics* 4: 121–140.

Putnam, M. 2006. Scrambling in West Germanic as XP-adjunction: A critical analysis of prolific domains. Ph.D. dissertation, University of Kansas-Lawrence.

Putnam, M. 2007. *Scrambling and the Survive principle.* Amsterdam: John Benjamins.

Putnam, M. & T. Stroik. 2008. Syntactic relations in Survive-minimalism. Ms.

Stroik, T. 2009. *Locality in minimalist syntax*. Cambridge MA: The MIT Press.

Stroik, T. 2008. The numeration in *Survive*-minimalism. Ms.

te Velde, J.R. 2005a. Unifying prosody and syntax for right- and left-edge coordinate ellipsis. *Lingua* 115: 483–502.

te Velde, J.R. 2005b. *Deriving coordinate symmetries. A phase-based approach integrating Select, Merge, Copy and Match*. Amsterdam: John Benjamins.

Uriagereka, J. 1999. Multiple spell-out. In *Working minimalism*, S.D. Epstein & N. Hornstein (Eds), 251–282. Cambridge MA: The MIT Press.

Wesche, B. 1995. *Symmetric coordination. An alternative theory of phrase structure* [Linguistische Arbeiten 332]. Tübingen: Niemeyer.

PART III

Covert and non-movement operations in Survive-minimalism

Syntactic identity in Survive-minimalism

Ellipsis and the derivational identity hypothesis*

Gregory M. Kobele
Humboldt-Universität zu Berlin

Over the years, a number of counter-examples to the hypothesis that ellipsis resolution is mediated via syntactic identity have been identified. However, in the same time evidence which seems to require comparison of syntactic structures in ellipsis resolution has also been unearthed. On top of this empirical puzzle, survive minimalism places an additional theoretical constraint: syntactic structures, once assembled, are opaque to further search or manipulation. In this paper, I show that a simple perspective shift allows us both to view the purported counter-examples as providing glimpses into the nature of the operations which build syntactic structure, and to satisfy the theoretical constraints imposed by survive minimalism's derivational take on syntactic structure.

1. Syntactic identity in ellipsis

One of the basic tasks of a theory of ellipsis is to explain what elliptical sentences can mean.[1] An observation so obvious it hardly bears mentioning is that the interpretation of an ellipsis site is not free, but rather is constrained in various ways by aspects of its surrounding context. As an example, in an empty discourse context, sentence 1 has only one meaning, which is synonymous with 2, and sentence 3 means only what 4 does, although only the matrix verb differs between the two.

(1) Jesús praised every girl that Adam did.
(2) Jesús praised every girl that Adam praised.
(3) Jesús met every girl that Adam did.
(4) Jesús met every girl that Adam met.

---

*My thanks to Hans-Martin Gärtner, Mike Putnam, and an anonymous reviewer for their helpful suggestions. The usual disclaimers apply.

1. A complete theory of ellipsis must account not only for how utterances of elliptical sentences are interpreted, but also for the distribution of ellipsis sites in discourse. It seems fruitful to approach the study of ellipsis by pursuing these two tasks independently. I will here largely ignore the task of specifying *when* ellipsis is licensed (though see Hardt (1993); Lobeck (1995); Murguía (2004)).

1.1 Two perspectives on ellipsis

There are roughly two approaches to determining the meaning of an elliptical sentence. The first view is that the meaning of 1 is computed differently from the meaning of 2. In particular, 1 is taken to be significantly syntactically different from 2 in that the embedded verb phrase in 1 has no internal structure (it can be thought of as a pronoun as long as one does not adopt the position that pronouns have internal structure (Parsons, 1978; Elbourne, 2002)). This first perspective I will refer to as the *proform theory of ellipsis*. The opposing view holds that the meaning of elliptical sentences is computed in exactly the same way as that of non-elliptical sentences; in other words, that ellipsis sites are syntactically structured objects.[2] According to this view, sentences 1 and 2 have the same meaning in virtue of the fact that they have the same or similar syntactic structure; the major difference between them is that some of this syntactic structure was not pronounced in 1. I will call this view the *non-pronunciation (or deletion) theory of ellipsis*. Note that the non-pronunciation theory of ellipsis has reframed the task of explaining what elliptical sentences mean as explaining under what conditions parts of sentences can be left unpronounced. Two options suggest themselves. Namely, that a sentence part can be left unpronounced just in case it is similar to an antecedent either syntactically, or semantically. I will refer to these different realizations of the non-pronunciation theory as the *deletion under syntactic/semantic identity theory of ellipsis*.

1.2 Deletion vs the pro-form theory

As the major structural difference between the deletion and pro-form theories lies in the syntactic representation they assign to the ellipsis site (with the deletion theory assigning to it the standard articulated structure, and the pro-form theory taking it to be syntactically simplex), the obvious way to argue for the deletion theory is to show that the ellipsis site exhibits characteristically syntactic behaviour, and to argue for the pro-form theory, that it doesn't.

1.2.1 *Island sensitivity*

It is well known that expressions can be dependent upon, descriptively speaking, positions internal to an ellipsis site, as in 5.

(5) Jesús loves Susan, but I don't know who Adam does.

---

2. This, of course, is predicated on the widely held assumption that meaning is read compositionally off of syntactic structure, or, in other words, that a syntactic description of an expression is simultaneously a description of its meaning.

Intuitively, *who* in 5 saturates a semantic argument of the recovered verb phrase meaning. While this has a natural description in terms of syntactic structure, we might think of equally natural semantic characterizations, and so the mere existence of sentences like 5 can't be taken as an argument for the deletion theory. However, characteristic of syntactic dependencies is that they are subject to various geometrical ('island') constraints, which do not seem to have a natural characterization in semantic terms. Merchant (2001) has shown that expressions (like *who* in 5) which seem to be dependent on ellipsis-site internal elements are subject to island constraints, as in 6, where *which subject* would be related to a gap within a relative clause (*girls that like t*), as is explicit in 7.

(6) *Jesús goes for girls that like geometry, but I don't know which subject Adam does.

(7) *Jesús goes for girls that like geometry, but I don't know which subject Adam goes for girls that like.

It is well-known that ellipsis sometimes ameliorates island violations. Although this might seem to favor the pro-form theory, the fact that there are *ever* island effects in elliptical contexts poses a severe problem for the pro-form theory, as it doesn't have access to the syntactic distinctions which seem necessary to correctly discriminate between those sentences which violate island constraints, and those which do not.

1.2.2 *Voice mismatches*

Passive and active sentences (as in 8 and 9) are semantically equivalent, although clearly they differ syntactically.[3]

(8) Gorbachev could have released this information.

(9) This information could have been released by Gorbachev.

As we have here syntactic distinctness, but semantic identity, if ellipsis is sensitive to the distinction between active and passive sentences, the proform theory would be at a loss to explain this. Although voice mismatches in verb phrase ellipsis (as in 10, from Hardt (1993)) are not exactly the clear cases upon which linguistic theories should be based (Chomsky, 1956), there is psycholinguistic (Kobele et al., 2008)

3. Following standard terminology, semantic equivalence boils down to identity of truth conditions. Thus, to claim that passives and actives are semantically equivalent is to claim that a passive sentence cannot be true while its active counterpart is false, or vice versa. In the early days of generative syntax, 'passivization' was a canonical example of a meaning preserving syntactic transformation. Note that it is perfectly reasonable to hold that actives and passives are semantically equivalent, while maintaining that they are information structurally distinct.

and corpus (Hardt, 1993) evidence which provide independent confirmation of their grammaticality.[4,5]

(10) This information could have been released by Gorbachev, although he decided not to.

And so at first glance, it would seem that the facts favor the pro-form theory. However, closer inspection reveals this not to be the case. As noted by Merchant (2007, 2008), neither sluicing (11, 12) nor (psuedo-)gapping (13, 14, 15, 16) allows for voice mismatches between antecedent and elided VP.

(11) *Someone released the information, but I don't know by whom.

(12) The information was released (by someone), but I don't know by whom.

Sentence 11 has an active antecedent (*someone released the information*), and a passive ellided VP (*by whom the information was released*). In sentence 12, we see both that passive ellided VPs are acceptable when their antecedents are passive, and that prepositional phrases (*by whom*) are acceptable even without an explicit antecedent, thus suggesting that the reason for the unacceptability of 11 stems indeed from the voice mismatch.

(13) *Mary was praised by Adam, and Jesús, Susan.

(14) Adam praised Mary, and Jesús, Susan.

(15) *Adam praised Mary, and Susan, by Jesús.

(16) Mary was praised by Adam, and Susan, by Jesús.

In the unacceptable 13, the antecedent clause is passive, and the ellipsis clause active. Sentence 14 shows that the unacceptability of 13 is not due to a ban against the ellipsis clause being in the active voice. Sentences 15 and 16 show the same for passive ellipsis clauses.

---

4. Recall that a sentence is grammatical if and only if it is generated by our theory of grammar, and that it is acceptable if and only if people like it. These properties needn't be coextensional: center-embedded sentences (such as *The boy the girl the cat the dog chased licked likes died*) are frequently claimed to be, although unacceptable, grammatical. Arregui et al. (2006) claim that sentences like 10 are, although acceptable, ungrammatical.

5. To allay any confusion, this is just a convenient way of saying that if we decide to treat voice mismatches in verb phrase ellipsis as ungrammatical (and therefore write our grammars in such a way as to rule them out), their corpus-attestedness and relative acceptability would be tedious to account for. Arregui et al. (2006) start with the assumption that voice mismatches are ungrammatical (i.e., not generated by the grammar), and are led to introduce various complicated psycholinguistic 'repair rules' to explain why people find voice mismatches relatively acceptable. Note that their account is not able to explain why people ever produce such sentences in the first place.

The fact that purely syntactic (i.e., non-semantic) distinctions can be relevant in determining the well-formedness of elliptical sentences argues against not only the pro-form theory, but also the deletion under semantic identity theory.

1.2.3 *Conclusions*

We have seen two examples (island effects and voice mismatches) where purely syntactic properties internal to the ellipsis site are crucial in determining the grammaticality of their containing sentences. Others are presented in Merchant (2001) and Chung (2006). Only the deletion under syntactic identity theory is currently capable of correctly accounting for the patterns of grammaticality assumed above. It may well not be the whole story, but, as we have seen, it must be part of it. And so, in the remainder of this paper, I will see exactly how far we can get with just the deletion under syntactic identity theory, in its most restrictive form.

2. On syntactic identity

I propose to take seriously the idea that 'deletion under syntactic identity' is licensed by *exact* syntactic identity. As I intend it here, syntactic structure is what is constructed during the derivational process, and, in line with the tenets of survive minimalism, 'destructive' operations are not allowed (this is also known as the non-tampering condition). This has the consequence that operations like *trace conversion* (Sauerland, 1998; Fox, 1999), which changes syntactic copies to (something like) traces, are not permitted. Indeed, given the copy theory of movement, and the non-tampering condition, we are led immediately to the following conclusion:

> If a syntactic object SO_1 is elided under identity with SO_2, then
> SO_1 and SO_2 have been derived in *exactly the same way*

It is important to note that, without the ability to alter already built structure, syntactic identity boils down to being derived in the same way, or *derivational identity*.[6] Shifting our perspective from identity of structure, to identity of the processes which construct this structure allows us to conceptualize deletion under syntactic identity in such a way as to render it compatible with another, even more fundamental, tenet of survive minimalism, which is that operations which search through already created structure should be banned. Derivational identity can be checked while still respecting the principles of survive minimalism simply by building identical expressions simultaneously—by synchronizing derivations (see Kobele (2006)).

6. As far as I am aware, the view that 'syntactic identity' should be understood as derivational identity was first articulated by Lees (1960).

Most importantly, viewing syntactic identity in terms of derivational identity helps us determine the right question to ask when we are confronted with a case of ellipsis which seems to fly in the face of our current analysis. Given that we can't change the structure we have built up (by applying a destructive operation), we must instead revise our analysis so as to allow us to directly construct the right structure.[7] This has a number of interesting methodological consequences, which I will talk more about later, among them that, as derivational identity can be formulated independently of any particular syntactic theory, ellipsis can be used to decide between different analyses of a particular phenomenon written in different frameworks.

Although simple, and for that reason appealing, the hypothesis that ellipsis is deletion under derivational identity is simply a programmatic assumption, and needs to be shown worthwhile to be interesting. Accordingly, in the remainder of this section I show how the derivational identity hypothesis (DIH) can be fruitfully applied even to complicated data. Before we begin, a note is in order. The DIH is a methodological assumption, and is not the type of thing which can be true or false. It can be only fruitful, in which case we should continue to use it, or not, in which case we should abandon it. It is my goal to show that it is indeed a fruitful assumption.

2.1 Derivational identity vs data

Although the DIH cannot bear a truth value, it *can* be *incompatible* with other assumptions, such as those of survive minimalism, Tree-Adjoining Grammar (TAG), or Combinatory Categorial Grammar (CCG). It is, in this sense, like a constituency test, which, as a way to interpret data, cannot possibly be true or false, although it certainly *can* be incompatible with a particular analysis of that data (hence its great value).

2.1.1 *The reality of the verb phrase*

Consider sentence 17, in which the verb and object of the second conjunct are deleted.

7. This is similar to the *deforestation* program transform in the functional programming literature (Wadler, 1990). Deforestation is a general technique that allows two functions g and f that apply in serial to be interleaved. This allows the input i to be directly mapped to the output $g(f(i))$ without first building an intermediate representation $f(i)$ (as the representations in question are usually trees, deforestation amounts to getting rid of intermediate trees, hence the name). In our case, we have a derivation which results in a structure d, and then a destructive operation *trace-conv*, which applies to this structure to yield a structure *trace-conv*(d) which is appropriate for syntactic identity. What we are doing, then, is attempting to find an alternative way of deriving the sentence in question, so that we directly arrive at this desired structure; the mechanism proposed later on can be viewed as performing the trace-conversion derivationally.

(17) Jesús will praise Mary, but Adam won't ~~praise Mary~~.

We have two analytical options. First, we can assume that there were two independent applications of a deletion operation, one targeting the verb, and one targeting the object. Second, we can assume that both elements were deleted in one fell swoop, by a single application of the deletion mechanism. These options are schematized in Figure 1. As a general rule, we will prefer analyses that involve fewer applications of the deletion mechanism. This is not forced upon us by anything, but seems reasonable to assume as a default for the following two reasons. First, although we are currently working on building a theory of *how* ellipsis works, we ultimately hope to extend it to a theory of *when* ellipsis is licensed. If we analyze the sentences above as involving independent applications of deletion to the verb and to the object, it seems like it will be difficult to rule out a sentence like 18 below, which also involves two applications of deletion.

S ~~V~~ ~~O~~ S ~~V O~~

Figure 1. Analytical options

(18) *Jesús will praise Mary, but Adam will ~~praise~~ the man who loves ~~Mary~~.

Second, if we don't make this assumption (at least as a default), we can't do anything interesting.

Preferring, then, to assume as few applications of deletion as possible, we are led to the conclusion that the sentences above (and more generally, transitive sentences) can be derived so that the verb and object form a constituent to the exclusion of tense, and the subject. Note that we are *not* entitled to conclude anything about what derivations are *not* available, only which ones are. Thus, we cannot conclude, for instance, that transitive sentences *always* are constructed with the verb and object coming together before the tense and subject.

As a consequence, any analysis of English according to which transitive sentences are built up exclusively from left to right (such as would seem to arise in a system like that proposed by Philips (2003)) would be incompatible with the interpretation of the above data forced upon us by the DIH. More interestingly, analyses such as Steedman's [2000] CCG analysis of English which, although it allows the verb and object to be derivational constituents some of the time, disallows it in cases like 19, are ruled out by the DIH.

(19) Jalapeños, Jesús hates. Adam does ~~hate jalapeños~~, too.

Passive is Phrasal, not (sentential or lexical) Turning now to the voice mismatches in VP-ellipsis which had once been interpreted as an argument against a syntactic account of ellipsis (examples are given as 20 and 21), we see that the DIH leads us to the conclusion that passivization is a VP level process. In other words, whatever structural differences there may be between active and passive sentences, the derivation of a sentence begins in the same way, regardless of its voice, up to at least the point where the object is merged.[8]

(20) This information could have been released by Gorbachev, although he decided not to ~~release this information~~.

(21) The children asked to be squirted with the hose, so we did ~~squirt with the children with the house~~

Unlike the previous conclusion (that verb and object may be derivational constituents), the present one places us on one side of an important theoretical disagreement, ruling out nearly all current analyses of English passivization. In particular, all analyses that have been written in the TAG framework are incompatible with this conclusion, as are all previous CCG analyses, as in both of these grammar formalisms analyses of passivization have been uniformly lexical. The kind of analysis of sentences like 20 and 21 forced upon us by the DIH is given in Figure 2. Note that the very same conclusion (that passivization applies to verb phrases) has been argued for independently on very different grounds (Bach, 1980; Keenan, 1980).

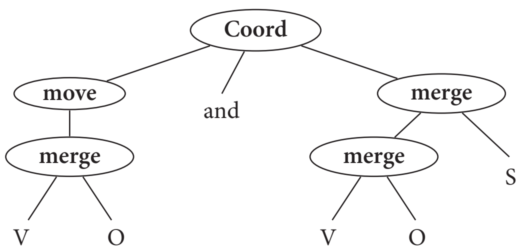


Figure 2. The gross structure of voice mismatches

2.1.2 *A paradox?*

We have seen above that, far from being incompatible with a syntactic approach to the identity condition in ellipsis, voice mismatches (together with the DIH) provide us with a powerful argument for a transformational approach to voice, leading us to the position that actives and passives have a common initial

8. Note as well that, given the morphological differences between the passive and active form of the verbs above, the DIH is incompatible with the position that words come into syntax fully inflected (Chomsky, 1995), and, more generally, with any grammatical organization in which the output of morphology is the input to syntax.

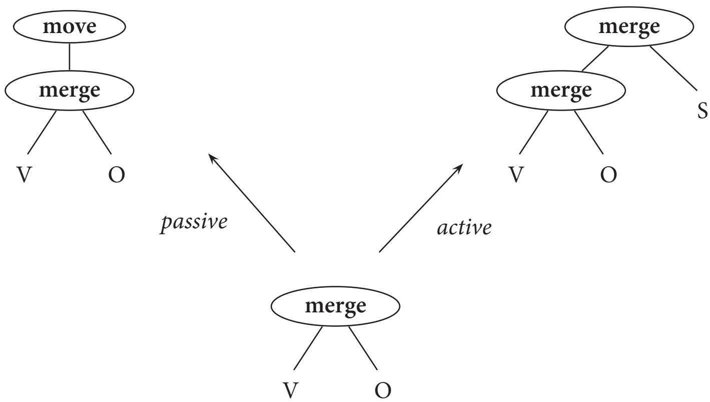


Figure 3. The relation between actives and passives

subderivation (as in Figure 3). That this cannot be the entire story is confirmed by the fact that passives can antecede passives, as in the sentences below.[9]

(22) Mary was praised, and Susan was ~~praised~~, too.

(23) Mary doesn't seem to have been praised, but Susan does ~~seem to have been praised~~.

Although sentence 22 could be treated as simple deletion of the verb (although non-trivial assumptions about morphology would have to be made), in sentence 23 the entire verb plus non-finite complement is deleted. According to the DIH, because the surface subjects *Mary* and *Susan* are different, they cannot be present in the derivation at the point where ellipsis occurs. In other words, we need a second structure for passives, as shown in Figure 4. Although the structure on the left in Figure 4 depicts the derivation that we expect to have given current (default) assumptions about the nature of passivization made in the context of government and binding theory, and its successors, the structure on the right is

9. Matched-voice ellipsis is by far more acceptable than mismatched voice ellipsis (Tanenhaus and Carlson, 1990; Arregui et al., 2006; Kobele et al., 2008). Thus, even if we decide to remove the mismatched examples from the purview of syntactic theory (i.e., call them ungrammatical, and explain their attestedness some other way), we are left with the same problem, as active-active ellipsis forces us to assume that transitive sentences can be constructed with a derivational VP, and passive-passive ellipsis that they can be constructed without a derivational VP. Note that this is precisely the problem that led Sag (1976) to abandon the deep structure identity condition of his contemporaries. The solution proposed in this paper can be thought of as a syntactic reimplementation of his.

more mysterious. Still, it is not completely without precedent in the minimalist community, something like it having been proposed by Manzini and Roussou (2000) in the context of A-movement. Slightly farther afield, the structure on the right is precisely the kind of derivation one would expect, if we had slash-feature percolation instead of movement. In the remainder of this paper, I will present a system which allows for both kinds of passive derivation demanded by the DIH.

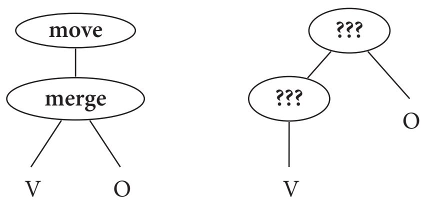


Figure 4. Two derivations of passive sentences

3. (Survive) minimalism

As presented in Stroik (2008), survive minimalism departs from more traditional variants of minimalism primarily in that movement is viewed not as being driven by the attractor, but rather by the moving object itself. This means that it is known whether an expression will stay in its current position as soon as the expression is merged there. This allows Stroik to adopt a strong 'no-tampering' condition on syntactic objects, treating them as black boxes which are opaque to search and manipulation. Intuitively, whereas the 'standard' minimalist would derive sentence 24 along the lines of 25, necessitating a search for a goal with features matching the probe in C (25.4), the survive minimalist would already have prepared the 'goal' for movement as soon as it was merged (26.1), and have it on-hand for remerge into SPEC-CP as soon as the C head is introduced (26.4).

(24) What did Jesús eat?

(25)
1. [eat what$_{wh}$] (merge eat and what)
2. [Jesús [eat what$_{wh}$]] (merge Jesús)
3. [C$_{wh*}$ [Jesús [eat what$_{wh}$]]] (merge C)
4. [what$_{wh}$ C$_{wh}$ [Jesús [eat ~~what~~$_{wh}$]]] (probe for and remerge what)

(26)
1. [eat ~~what~~]; what$_{wh*}$ (merge eat and what; what 'survives')
2. [Jesús [eat ~~what~~]]; what$_{wh*}$ (merge Jesús)
3. [C$_{wh}$ [Jesús [eat ~~what~~]]]; what$_{wh*}$ (merge C)
4. [what$_{wh}$ C$_{wh}$ [Jesús [eat ~~what~~]]] (remerge what)

Because we will be here concerned very much with the structure of derivations, and little differences will matter a great deal, we will have to flesh out a theory of grammar well beyond what is usually done in Stroik's work and elsewhere. For this purpose, we will use minimalist grammars (MGs), introduced in Stabler (1997) as a working out of some of the core ideas in Chomsky (1995).[10] There is an interesting connection between minimalist grammars and the intuitions in Stroik (2008). Although Stabler's original presentation of minimalist grammars was in terms of 'destructive' operations on trees (as in the standard picture on derivations outlined in 25), it was proven in Michaelis (2001) that we could adopt a perspective on derivations similar to the one in 26 without changing any of the properties of the formalism. In other words, MGs can be thought of as a precise variant of survive minimalism.

3.1 Derivations

There are two basic operations in minimalist grammars; one that combines two structures, which we will call **merge**, and one that operates over a single structure, which will be called **move**. Entire derivations can be represented as single trees. Given three lexical items, *A*, *B*, and *C*, the following derivation can be represented by means of the single derivation tree in Figure 5.

1. **merge**$(A, B) = [_A\ A\ B]$
2. **move**$([_A\ A\ B]) = [_A B\ [_A\ A\ t_B]]$
3. **merge**$(C, [_A\ B\ [_A\ A\ t_B]]) = [_C\ C\ [_A\ B\ [_A\ A\ t_B]]]$

In the case of the **merge** operation, we adopt the convention that the first of its two arguments is the one which projects over the other. Although the concept of a

10. One advantage of working within such a formalism is that everything—every mechanism, derivational step, and data structure—is perfectly explicit (it is a *generative grammar*, in the sense of Chomsky (1965)). Only with this kind of precision is the theoretician able to demonstrate both why a particular sentence is derivable in his or her system, and also (a point too often neglected) that a particular sentence is *not* derivable. Furthermore, it allows for construction of learning algorithms (Stabler et al., 2003), parsing algorithms (Harkema, 2001), and serious psycholinguistic processing models (Hale, 2003), as well as meaningful comparisons with other grammar formalisms (Harkema, 2001; Michaelis, 2001; Kobele, 2002; Stabler, 2003).

The fact that everything is completely worked out does not imply that it is set in stone. Frey and Gärtner (2002) introduce asymmetric feature checking to account for scrambling and adjunction, Gärtner and Michaelis (2007) study the result of adding various locality constraints to the formalism, Kobele (2005) studies a certain kind of feature percolation, Stabler (2006) investigates a version of sideward movement, and Kobele (2002) investigates Brody-style mirroring (Brody, 2000).

numeration will play no role in the theory developed here, it can be straightforwardly defined in terms of the **yield** of derivation trees.[11]

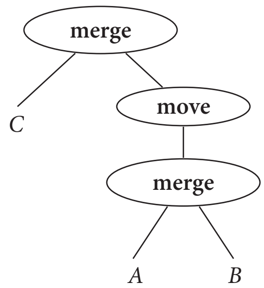


Figure 5. Representing derivations with trees

3.2 Features

As the generating functions **merge** and **move** are taken to be universal and invariant, differences between languages reside in the lexicon. Lexical items have various features, which determine how they behave in the derivation. In addition to movement being feature driven, I will assume that merger is as well, and that, for ease of presentation, the kinds of features which are relevant for the **merge** and the **move** operations are distinct, and will be called **selection** and **licensing** features, respectively.[12] Each feature has an **attractor** and an **attractee** variant (Figure 6), and these must match in order for an operation to apply. Each time an operation is applied, it checks both an attractor and an attractee feature, of the appropriate variety.

	attractor	attractee
merge	=x	x
move	+y	-y

Figure 6. Features

11. By a result of Hale and Stabler (2005), in the context of minimalist grammars a string of lexical items uniquely determines a derivation tree.

12. Treating syntactic features as falling into these two distinct groups is really for simplicity only. (It is simpler in the sense that it cuts down on the bookkeeping needed to ensure that lexical items that shouldn't interact don't.) We could just as well treat there as being only one kind of feature relevant for both **merge** and **move** operations (say, *z and z) without changing the essential formal properties of the system.

Although the features used here will be suggestively named (such as V for VP, v for little-vP, n for NP, and so on), the features themselves have no content apart from that which is given to them by the role they play in the formal system. In other words, in a minimalist lexicon, all that matters is which features match which others, and not what they are called.[13] We will also not attempt here to model inflection or agreement, which is perhaps best seen as a side-effect of formal feature checking. The hope is, of course, that the (too) simple system studied here will synthesize seamlessly with more sophisticated accounts of agreement.

3.3 Lexical items

Syntax relates form and meaning. Lexical items are the building blocks of this relation, atomic pairings of form and meaning, along with the syntactic information necessary to specify the distribution of these elements in more complex expressions. Here, simplifying somewhat, we take lexical items to be pairings of abstract lexemes such as dog, cat, $bank_1$,. . . with feature bundles. Feature bundles are not unstructured sets, but are rather ordered, so that some features can be available for checking only after others have been checked. We will represent feature bundles as lists, and the currently accessible feature is at the beginning (leftmost) position of the list. An example lexical item is shown in Figure 7. Its feature bundle '=d V' indicates that it first selects a DP argument, and then can be selected for as a VP.[14] There are a number of advantages of treating feature bundles as lists, instead of as sets. First, certain lexical items, such as the Saxon genitive's, are naturally thought of as selecting two syntactic arguments, an NP complement and a DP specifier. As the notions of 'complement' and of 'specifier' reduce to 'first-merged'

⟨praise, = d V⟩

Figure 7. A lexical entry for *praise*

13. It is interesting, in this context, to think of how to argue that a feature in one language should bear the same name as a feature in another language. Intra-sentential code-switching, as implemented in minimalist grammars by MacSwan (2000), is one way of imbuing this question with determinate empirical content.

14. More specifically, in order to be checked, its first feature =d requires that it be merged with an expression which has a matching attractee feature, d. After this, the resulting feature bundle will have first feature V, in order to check which, the expression must be merged with another which has the matching attractor feature, =V, as its first feature.

and 'not-first-merged' in the context of the minimalist program, we need a way of ensuring that the first merged argument of 's is the NP, and not the DP. This not being obviously derivable from anything else, here we can simply structure the feature bundle for 's so as to have the noun phrase selected before the determiner phrase.[15] A second advantage is that we can state directly that (for example) the case feature of a DP must be checked before its *wh*-feature.

3.4 Syntactic objects

I am writing lexical items using the notation $\langle \alpha, \delta \rangle$, where α is a lexeme (such as praise), and δ is a feature bundle (such as '=d V'). Complex expressions I will write using a labeled bracket notation, as per the following:

$[_{\delta}\ \alpha\ \beta]$

The above expression has feature bundle δ. As an example, if we assign the lexical entries ⟨'s, =n =D D⟩ and ⟨brother, n⟩ to the Saxon genitive and to brother respectively, then the complex expression 's brother, which is the result of merging brother as the complement of 's, is represented as the below.

$[_{=\text{D D}}$'s brother]

As can be seen, the above expression has feature bundle '=D D', which means that after it merges with an expression with first feature D, it will itself be such an expression.

$[_{\text{D}}$ Jesús ['s brother]]

By representing the relevant syntactic information about an expression in its feature bundle, we technically no longer need to make constituency distinctions, and can (and sometimes for convenience will) write the above expression as per the following.

$[_{\text{D}}$ Jesús 's brother]

Moving expressions, that is, expressions which have 'survived' from earlier merger, will be written after a semi-colon following the 'main' expression. For example, the expression below is a sentence (IP) which contains a *wh*-phrase which has not yet checked its -wh feature.

$[_{\text{i}}$ Jesús will praise]; $[_{\text{-wh}}$ who]

15. Another option is to have the semantics take over responsibility for this distinction, perhaps by assigning to 's the meaning s of type $(eet)((et)t)(et)t$, such that $s(R)(Q)(B) = Q(R^{-1}[B])$, where $R^{-1}[B] := \{a : \exists b \in B.\ aRb\}$, and attributing the unacceptable **doctor's every boy* to a semantic type mismatch. The problem with this is that we no longer have a direct description of the well-formed

3.5 Merge

The **merge** operation applies to two arguments, *A* and *B*, resulting in the new object *A* + *B*, just in case the head of *A* has some selector feature =x as the first unchecked feature in its feature bundle, and the head of *B* has the corresponding x as the first unchecked feature in its feature bundle. In the resulting *A* + *B*, both first features used in this derivational step are checked, making available the next features of both feature bundles. There are three cases of the **merge** operation, depending first on whether the *B* argument will continue moving, and then, if not, whether *B* will surface as a specifier or as a complement of *A* (Figure 8). The first case corresponds to the operation **Survive!** in survive minimalism. Let us assume that all DPs move for case, which we will here represent with an abstract feature 'K', the attractor version of which is +k, and the attractee version is -k. Then, provisionally, we assign the feature bundle 'd -k' to DPs like Jesús, Adam, Mary,. . . Now, whenever a DP is selected by an expression (with first feature =d), we know that it will not surface in that position, as it still has a case feature that must be checked (via movement to some later specifier position). Consider the derivation step below, in which praise selects Jesús as its complement.

$$\textbf{merge}(\langle \text{praise}, =\text{d V}\rangle, \langle \text{Jesús}, \text{d -k}\rangle) = [_{\text{V}} \text{ praise } t_{\textit{Jesús}}]; [_{\text{-k}} \text{ Jesús}]$$

Because Jesús still needs to move for case, it 'survives' by not being trapped within the main expression. As Kobele (2006) shows how to directly model-theoretically interpret each derivational step, obviating thereby any semantic need for traces, I will usually leave out explicit mention of the trace in the expression above, writing instead:

$$\textbf{merge}(\langle \text{praise}, =\text{d V}\rangle, \langle \text{Jesús}, \text{d -k}\rangle) = [_{\text{V}} \text{ praise}]; [_{\text{-k}} \text{ Jesús}]$$

$$\textbf{merge}(A, B) = \begin{cases} A; B & B \text{ survives} \\ [_A\ A\ B] & B \text{ is a complement} \\ [_A\ B\ A] & B \text{ is a specifier} \end{cases}$$

Figure 8. Cases of merge

sound-meaning pairs; they are now defined indirectly as the result of applying a partial function to a set. (The desire to have a direct description of the well-formed sound-meaning pairs is expressed in the guise of *crash-proof syntax* by Frampton and Gutmann (2002).) Kracht (2007) has argued in this vein that constituency (order of application) is a fundamentally syntactic notion, and should be excised from the semantics.

Important to note is the fact that in the resulting expression, both matching features of the arguments (=d and d) are checked (which for present purposes means gotten rid of).

The next case of **merge** is the merger of a complement, or 'first-merge'. Let us represent the head assigning case to the object in English (AgrO) with the following lexical item: ⟨-ϵ, =V +k agrO⟩.[16] Because the first feature in the feature bundle of this lexical item is =V, it can be merged with the VP *praise Jesús* we derived above. There are two things of note to watch out for here. First, one of the arguments to merge has 'surviving' expressions attached to it. These will simply continue to survive in the result. The same would be true if both arguments had surviving expressions attached to them. Second, the selecting lexical item is a suffix (marked by the hyphen preceding the lexeme). This triggers head movement from its complement.[17]

merge(⟨-ϵ, =V +k agrO⟩, $[_{\text{V}}$ praise]; $[_{\text{-k}}$ Jesús])

= $[_{\text{+k agrO}}$ praise-ϵ t_{praise}]] ; $[_{\text{-k}}$ Jesús]

head movement

Leaving out the traces, and unnecessary internal structure, this expression can be abbreviated as the below.

$[_{\text{+k agrO}}$ praise-ϵ]; $[_{\text{-k}}$ Jesús]

Note again that both of the matching first features of the arguments to **merge** (=V and V) have been checked in (i.e., deleted from) the result.

The last case of **merge** is merger of a specifier into its surface position. This happens whenever the first argument to **merge** is not a lexical item, and the second argument has no unchecked features (other than the matching selectee feature which drives the **merge** operation). This will not occur in any of the examples we will encounter in this paper.

3.6 Move

The **move** operation applies to a single syntactic object *A* just in case it contains a surviving expression *B*, and the first unchecked feature of *A* is +y and the first

16. The symbol ϵ is the empty string. The hyphen in front of it indicates that it is a suffix, and thus triggers head movement; its complement's head raises to it. Its feature bundle =V +k agrO indicates that it first merges with a VP (=V), then provides a case position for the object to move to (+k), and finally becomes itself selectable as an AgrOP (agrO).

17. See Stabler (2001) for more details. The basic idea is that head movement is treated as a quasi post-syntactic operation, along the lines of Matushansky (2006). In other words, head movement is not feature-driven movement.

unchecked feature of B is -y. In order to rule out indeterminacy, **move** will only be defined if *exactly one* surviving expression begins with a matching –y feature.[18] Just as with the **merge** operation, **move** checks the matching features of the expression which it applies to. There are two cases, according to whether the moving element survives or not. As an example of the first case, consider the expression below (which is similar to *praise Jesús* derived above, but with who having been merged instead of Jesús).

$[_{\text{+k agrO}}$ praise-ϵ]; $[_{\text{-k -wh}}$ who]

The **move** operation applies to this expression, as the main expression has as its first feature a licensor +k, and there is exactly one surviving expression with first feature the matching licensee -k. In the result, because the moving expression still has features left to check (-wh), it survives.

move($[_{\text{+k agrO}}$ praise-ϵ]; $[_{\text{-k -wh}}$ who]) = $[_{\text{agrO}}$ t_{who} [praise-ϵ]]; $[_{\text{-wh}}$ who]

Again, leaving out traces and internal structure, we write the expression above as per the below. Note again that both features involved in the move operation (+k and -k) are deleted/checked in the result.

$[_{\text{agrO}}$ praise-ϵ]; $[_{\text{-wh}}$ who]

The second case of the move operation is when the moving element is moving to what will be its surface position; in other words, its last feature being checked, it no longer needs to survive. The expression *praise Jesús* we derived earlier illustrates this.

move($[_{\text{+k agrO}}$ praise-ϵ]; $[_{\text{-k}}$ Jesús]) = $[_{\text{agrO}}$ Jesús [praise-ϵ]]

Ignoring internal structure, the expression above becomes the below. This case of the **move** operation permits surviving expressions to become reintegrated with the main expression.

$[_{\text{agrO}}$ Jesús praise-ϵ]

3.6.1 *Eliminating movement*

In the minimalist syntactic literature, there has been much interest in unifying the operations of **merge** and **move**. In a purely set theoretical sense (i.e., viewing

18. This is a radical version of the shortest move constraint (Chomsky, 1995), and will be called the SMC – it requires that an expression move to the first possible landing site. If there is competition for that landing site, the derivation crashes (because the losing expression will have to make a longer movement than absolutely necessary). Using this as a constraint on movement has desirable computational effects (such as guaranteeing efficient recognizability–see (Harkema, 2001; Michaelis, 2001)), although other constraints have been explored in Gärtner and Michaelis (2007), who show that adding constraints on movement can sometimes actually increase the expressivity of the formalism.

functions as sets of pairs), this is easy to do: **merge** is a binary function and **move** is unary, their domains are therefore disjoint, and so their union **merge** ∪ **move** is also a function. Equally clear, however, is the fact that this technical trick 'should not count.' It seems that what is going wrong is that the complexity of description of the grammatical operations is not changed by this particular 'unification'—the easiest way to describe the function **merge** ∪ **move** remains as the union of the functions **merge** and **move**, not directly. Thus, the real question at issue is whether there is a way of describing the action of the generating functions according to which it is simpler to describe **merge** and **move** as special cases of a single function, rather than as operations in their own right. Another question which should then be asked is whether such a unified perspective on **merge** and **move** is simpler than the current, non-unified perspective.

Stroik (2008) has attempted to unify **merge** and **move** by making extensive use of a numeration (or 'work space'); in other words, by treating the inputs to and outputs of the generating functions not as trees, but as (multi-)sets of trees (Chomsky, 1995). According to this perspective, 'surviving' expressions, which we have here written as part of a single expression, are reinserted into the work space. The **merge** operation is given an entire work space as its argument: it selects two objects therefrom to be merged, these are removed from the work space, and in their place are inserted the result of merging these objects, as well as any surviving expressions. '**Move**' is simply what happens when a surviving expression is chosen to be merged.

Although this interpretation of movement eliminates it as an operation in its own right (it has been reduced to a special case of a much more complex presumably non-functional relation we continue to call **merge**), some mechanism must be in place to ensure that merger of surviving expressions results in well-formed chains.[19]

The expressions used in this paper, with their explicit record of which expressions survive within them, needn't be thought of as concrete proposals of what grammatical objects look like. Instead, we can understand them as abstract representations not only of these objects, but also of whatever mechanism is used to ensure that only well-formed chains are built. In other words, when I write an expression like $[_{agrO}$ praise]; $[_{-wh}$ who], I am indicating that we have an expression $[_{agrO}$ praise], which must be involved in a larger derivation which uses the surviving expression $[_{-wh}$ who] which is currently in the work space. Therefore, the present system is perfectly compatible with the survive minimalist's reduction of **move** to a subcase of **merge**. It is however

19. 'Chains' as used here is a descriptive term: the dependencies into which a single expression enters must be related to each other in a certain way. This, as a fact about language, must be ensured even by grammar formalisms which do not reify chains as grammatical objects.

equally compatible with an alternative perspective on survival (also proposed in Stroik (2008)), according to which an expression survives by moving to adjoin to a projection of the current head of the tree, as well as with a host of other, equally valid yet heretofore undiscovered alternatives. As these concerns are orthogonal to the main point of this paper (which is to argue for adoption of the DIH), I continue to treat **merge** and **move** as different. The reader whom this discomforts should simply substitute each occurrence of the word **move** with the phrase "the subcase of **merge** where one of the two arguments is a surviving expression."

4. Passivization in English

In this section, I will present a grammar for a fragment of English comprising passivization. This will serve simultaneously as an introduction to the minimalist grammar formalism (and the fragment-style of language analysis), as well as a reintegration with the theme of this chapter, of providing a grammatical analysis of English consistent with the demands of the DIH.

Analyses in the transformational tradition tend to assign similar derivational histories to semantically related construction types. Passives and actives are related in virtue of having been constructed in the same way up to a point (as per Figure 3). It is precisely this perspective which we are pushed toward by the DIH.

The basic clause structure I will be assuming is as in Figure 9. I will assume that objects move overtly to an AgrO position located underneath the subject θ-position (Koizumi, 1995), which is introduced by little-v. This assumption is forced by the SMC constraint on the **move** operation: as all argument DPs will be assumed to have the feature bundle 'd -k',[20] as soon as the subject is merged into its θ-position, it survives with feature bundle '-k'. If the object has not checked its case feature by that point, there will be two surviving expressions with the same first feature (-k), which is precisely the configuration ruled out by the SMC. We can derive sentences with this clause structure with the lexical items in Figure 10. Using the lexical items in Figure 10, we can derive sentences like *Adam will have been criticizing Jesús*, as shown below.[21]

---

20. An alternative is to treat nominative and accusative cases as syntactically distinct, and then DPs might have either -ACC or -NOM features. The DIH actually argues against this alternative: as passives may antecede actives (as in 20), the DP inserted in a passive sentence must be identical to the DP inserted in an active one. Therefore, case differences must be left to the morphology.

21. In fact, we can derive exactly eight sentences (four of which violate condition C of the binding theory).

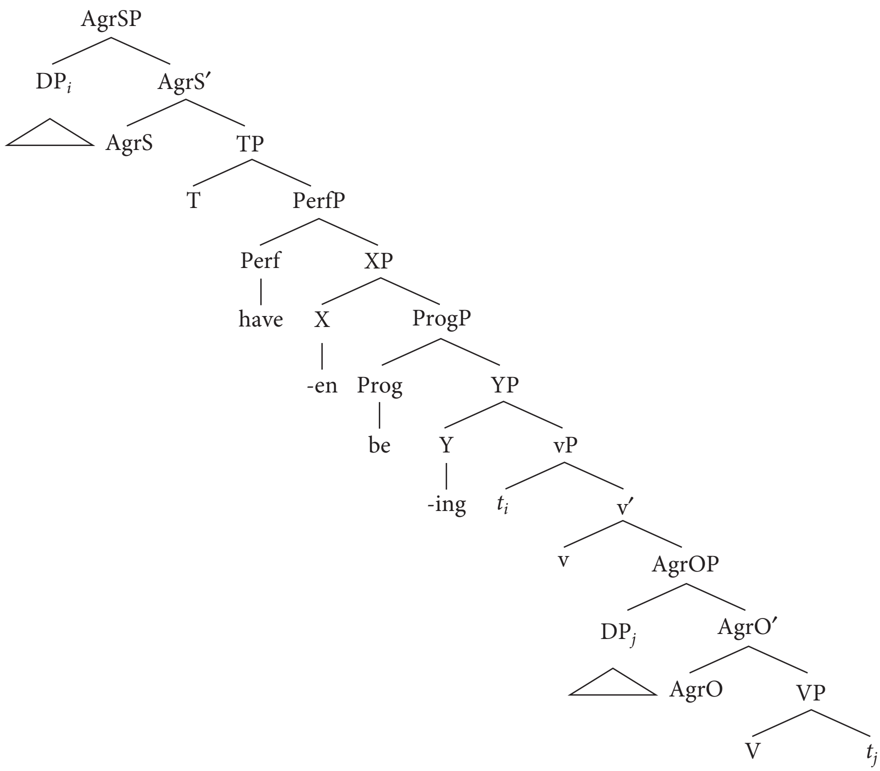


Figure 9. Basic clause-structure

⟨ϵ, =t +k agrS⟩	⟨will, =perf t⟩ ⟨-ed, =perf t⟩	⟨have, =x perf⟩ ⟨-en, =prog x⟩	⟨be, =y prog⟩ ⟨-ing, =v y⟩
⟨-ϵ, =agrO =d v⟩	⟨-ϵ, =V + k agrO⟩	⟨praise, =d V⟩ ⟨criticize, =d V⟩	⟨Jesús, d -k⟩ ⟨Adam, d -k⟩

Figure 10. Lexical items for Figure 9

We begin by merging together the lexical items criticize and Jesús. This is defined because the first features of the respective feature bundles are matching attractor (=d) and attractee (d) features. Because Jesús still has a feature to check (-k), it survives.

1. merge(⟨criticize, =d V⟩, ⟨Jesús, d -k⟩)

$[_V$ criticize $t_d]$; $[_{-k}$ Jesús$]$

The next step is to merge the lexical item ⟨-ϵ, =V +k agrO⟩ with the expression just derived. Because it is a suffix, it triggers the head of its complement (here: *praise*) to raise up to it.

2. **merge**(⟨-ϵ, =V +k agrO⟩,1)

$[_{\text{+k agrO}}$ criticize-ϵ [t_V t_d]]; $[_{\text{-k}}$ Jesús]

The first unchecked feature of the above expression is the attractor feature +k, which is checked by movement of the surviving expression *Jesús*, whose first feature is the matching attractee feature -k.

3. **move**(2)

$[_{\text{agrO}}$ Jesús [criticize-ϵ [t_V t_d]]]

Next little-v is merged, triggering head movement of *criticize*.

4. **merge**(⟨-ϵ, =agrO =d v⟩, 3)

$[=_{\text{d v}}$ criticize-ϵ-ϵ [Jesús[t_{agrO} [t_V t_d]]]]

Next the external argument, Adam, is merged, checking the =d feature of little-v. Because Adam has an unchecked case-feature, it survives.

5. **merge**(4, ⟨Adam, d -k⟩)

$[_{\text{v}}$ t_d [criticize-ϵ-ϵ [Jesús [t_{agrO} [t_V t_d]]]]]; $[_{\text{-k}}$ Adam]

In the next five steps, the tense and aspect heads are merged, triggering head movement when applicable.

6. **merge**(⟨-ing, =v y⟩, 5)

$[_{\text{y}}$ criticize-ϵ-ϵ-ing [t_d [t_v [Jesús [t_{agrO} [t_V t_d]]]]]]; $[_{\text{-k}}$ Adam]

7. **merge**(⟨be, =y prog⟩,6)

$[_{\text{prog}}$ be[criticize-ϵ-ϵ-ing [t_d [t_v [Jesús [t_{agrO} [t_V t_d]]]]]]]; $[_{\text{-k}}$ Adam]

⇝ $[_{\text{prog}}$ be [criticize-ϵ-ϵ-ing Jesús]]; $[_{\text{-k}}$ Adam]

8. **merge**(⟨-en, =prog x⟩,7)

$[_{\text{x}}$ be-en [t_{prog} [criticize-ϵ-ϵ-ing Jesús]]]; $[_{\text{-k}}$ Adam]

9. **merge**(⟨have, =x per⟩,8)

$[_{\text{perf}}$ have [be-en [t_{prog} [criticize-ϵ-ϵ-ing Jesús]]]]; $[_{\text{-k}}$ Adam]

10. **merge**(⟨will, =perf t⟩,9)

$[_{\text{t}}$ will [have [be-en [t_{prog} [criticize-ϵ-ϵ-ing Jesús]]]]]; $[_{\text{-k}}$ Adam]

Finally, we merge agrS, which allows the case feature of *Adam* to be checked.

11. **merge**(⟨ϵ, =t +k agrS⟩,10)

$[_{\text{+k agrS}}$ ϵ [will [have [be-en [t_{prog} [criticize-ϵ-ϵ-ing Jesús]]]]]]; $[_{\text{-k}}$ Adam]

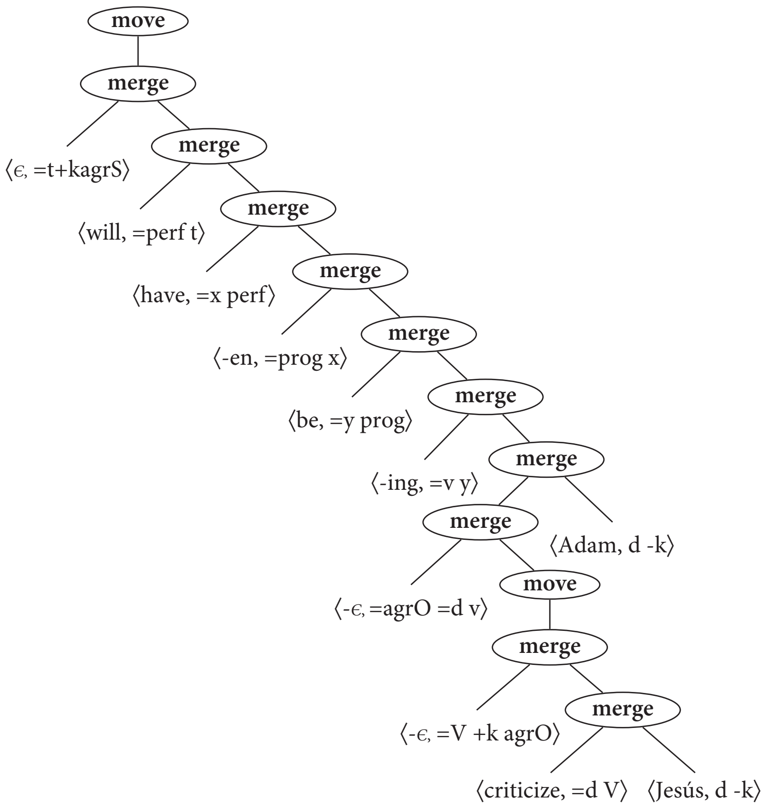


Figure 11. A derivation tree for *Adam will have been criticizing Jesús*

12. **move**(11)
$[_{agrS}$ Adam [ϵ [will [have [be-en [t_{prog} [criticize-ϵ-ϵ-ing Jesús]]]]]]]

The entire derivation of this sentence can be compactly represented as a single tree, shown in Figure 11. In order to allow for the optionality of *be -ing and have -en*, we need to explicitly introduce lexical items which allow for vP to be selected by XP, and ProgP by TP, respectively. This can be done with the lexical items below.

⟨-ϵ, =v prog⟩ ⟨-ϵ, =prog perf⟩

With these additions to our lexicon from Figure 10, we can derive sentences without either *be -ing, have -en*, or both. To derive the sentence *Adam will be criticizing Jesús* (i.e., our previous sentence without *have -en*), instead of steps 8 and 9, we substitute 8′ below.

8′. **merge**(⟨-ϵ, =prog perf⟩,7)
$[_{perf}$ be-ϵ [t_{prog} [criticize-ϵ-ϵ-ing Jesús]]]; $[_{-k}$ Adam]

The derivation can then continue with 10–12. Note that the expression in 8′ has exactly the same featural make-up as the expression in 9 (the main expression has the feature bundle 'perf', and the single surviving expression has the feature ti bundle '-k').

4.1 The passive

One major point of difference between passives (28) and actives (27) in English is that in the passive, the external argument is not required to be present (if it appears at all, it is in an optional by-phrase), and that the internal argument raises to the subject position and assumes all the canonical subject properties, such as nominative case.

(27) Adam was criticizing Jesús.

(28) Jesús was being criticized.

In the context of our present analysis (i.e., the lexicon in Figure 10), objectual case is assigned in AgrO, and the external argument introduced in v. In order to block both case assignment to the object and assignment of an external θ role (selection of the subject), we simply allow lexical items particular to the passive to select the VP, do their thing, and then result in a vP. We introduce the two new lexical items in Figure 12. The first, ⟨en, =V z⟩, selects a VP, triggers head movement of the verb (causing it to appear as a participle), and then returns a ZP.

⟨-en, =V z⟩ ⟨be, =z v⟩

Figure 12. Lexical items for the passive

1′. **merge**(⟨-en, =V z⟩, [$_V$ criticize t_d]; [$_{-k}$ Jesús])

[$_x$ criticize-en [t_V t_d]]; [$_{-k}$ Jesús]

The next, ⟨be, =z v⟩, selects a passive participle phrase (ZP), and returns a vP.

2′. **merge**(⟨be, =z v⟩,1′)

[$_v$ be [criticize-en [t_V t_d]]]; [$_{-k}$ Jesús]

The resulting expression (2′) has the same featural make-up as the expression in 5 (the main expression has the feature bundle 'v', and the only surviving expression the feature bundle '-k'), and thus is syntactically interchangeable with it. The derivation of the sentence *Jesús was criticized* is as in Figure 13. Our current fragment can define passives and actives of sentences with various tenses and aspects. While not particularly exciting it will suffice for the purposes of this paper.[22] We now return to these.

22. Kobele (2006), from which these lexical items are drawn more or less verbatim, shows how this fragment can be extended naturally to account for raising to subject and object, control, expletive-*it*, and quantifier scope ambiguities.

5. Satisfying the DIH

The derivational identity hypothesis, in response to sentences 20 and 21,forced upon us the conclusion that actives and passives must have a derivational constituent in common. Comparing the derivations of active andpassive sentences shown in Figures 11 and 13 respectively, we see that they share the subderivation shown in Figure 14. Ignoring for the moment that our fragment isn't expressive enough to generate sentences like (the unelided version of) 20 (we do not have modal auxiliaries *could*, nor control verbs *decide*), if we render phonologically null just the lexical items shown in Figure 14 in 29, we obtain a sentence parallel to 20.

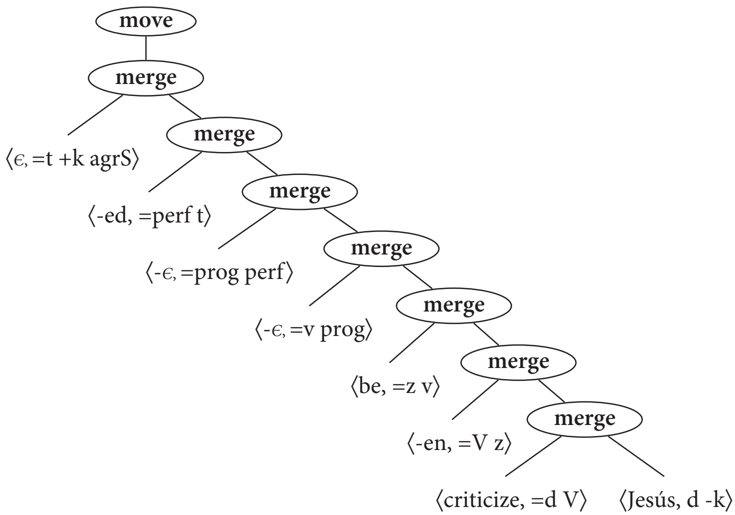


Figure 13. A derivation tree for *Jesús was criticized*

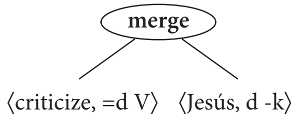


Figure 14. The common subderivation of Figures 11 and 13

(29) Jesús could have been criticized, but Adam decided not to ~~criticize~~-ϵ-ϵ ~~Jesús~~.

Allowing for deletion under derivational identity at the VP level, we can generate the dialogues given in 30 and 31 (extending our fragment in the obvious way for proper names). Both B responses (they are the same) should be read as *Adam did*.

(30) A: Mary criticize-ϵ-ϵ-ed Jesús.
B: (No,) Adam ~~criticize~~-ϵ-ϵ-ed ~~Jesús~~.

(31) A: Jesús be-ϵ-ϵ-ed criticize-en.
B: (Yes,) Adam ~~criticize~~-ϵ-ϵ-ed ~~Jesús~~

Thus, it seems that minimalist-style analyses support the kind of derivational constituents demanded by the DIH. However, they are not able to deliver *all* of the requisite derivational structure, even when we limit our-selves to the relatively simple case of voice. As seen in examples 22 and 23 (repeated below as 32 and 33), passive verb phrases can be elided, despite having different internal arguments.

(32) Mary was praised, and Susan was ~~praised~~, too.

(33) Mary doesn't seem to have been praised, but Susan does ~~seem to have been praised~~.

The problem is that according to the present analysis, the largest derivational constituent shared by both antecedent and elliptical clause is the lexical item ⟨praise, =d V⟩. Looking at Figure 15 (which, substituting Mary and Susan, respectively, for the blank, yields the respective structures of the clauses in 32), we see that the derivation trees for the passive clauses in 32 would have large enough a subtree in common if we were able to abstract away from the difference in internal argument. Although we could simply weaken the derivational identity hypothesis to allow for identical *contexts* instead of just identical *subtrees*

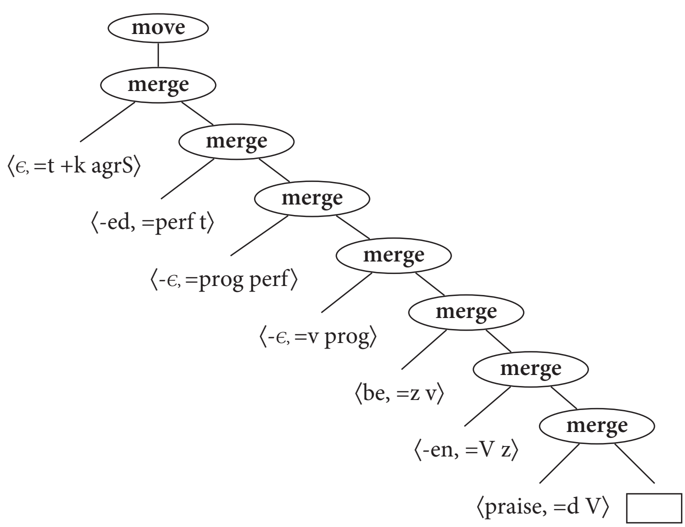


Figure 15. A derivation tree for ▭ *was criticized*

(as shown in Figure 15), this would lessen the analytical force of the DIH. Instead, let us take the DIH seriously, and ask the question of what kind of system would give us the desired derivational flexibility to account for voice-mismatches as in 20, as well as passive-passive ellipsis as in 32. This is the topic of the next section.

5.1 Hypothetical reasoning

The intuition behind the basic extension to the minimalist grammar system that will give us the needed derivational flexibility to satisfy the DIH is that we will be allowed to conduct derivations *as though* other expressions were there. This will allow us to eliminate an asymmetry in the structure of derivations; namely, that all expressions are introduced into derivations in their lowest chain positions. From the perspective of derivation trees, alongside the familiar **merge**-**move** pair shown on the left in Figure 16, we will allow the symmetric **assume**-**discharge** pair on the right (see also Figure 4). How should we interpret these two new operations? The **assume** operation should allow us to 'pretend' we had whatever expression we needed to continue the derivation. In the context of our lexicon in 10, we might have the following derivation.

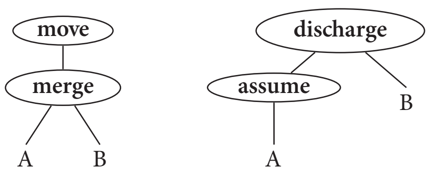


Figure 16. A new symmetry in derivations

1″. **assume**(⟨praise, =d V⟩)

$$[_{V} \text{ praise } t_{d}]; [_{\text{-k}}]$$

In 1″ above, the grammar hypothesized the existence of an expression with feature bundle 'd -k', which survives, as it still has a feature to check (-k). We continue the derivation, just as we would with a 'real' surviving expression, as follows.

2″. **merge**(⟨-ϵ, =V +k agrO⟩,1″)

$$[_{\text{+k agrO}} \text{ praise-}\epsilon\, [t_V\, t_d]]; [_{\text{-k}}]$$

At this point, we would like to apply the **move** rule to the above expression; the first feature of the head is the attractor +k, and there is a (hypothesized) surviving expression with first feature the matching attractee feature -k. As the moving element will have no more unchecked features after this derivational step, it is being put herewith into its surface position. However, it is only a hypothesized element, not a ‘real’ one, and must be replaced by a ‘real’ expression before it can become inactive.[23] To do this, we **discharge** the assumption by introducing an expression which ’would have served’ at the point the assumption was made.

3″. **discharge**(2″,⟨Jesús, d -k⟩)

$[_{\text{agrO}}$ Jesús [praise-ϵ $[t_V\, t_d]]]$

Note that, in order to discharge the assumption, the grammar must have access to information about what the assumption originally was. To make this maximally transparent, we represent hypotheses using the following notation, where δ is the current feature bundle, and γ the original hypothesized feature bundle.

[δ γ]

The cases of the **assume** and **discharge** operations are given in Figure 17. With this addition to the minimalist grammar framework (for more details, see Kobele (2007)), we can account for passive-passive verb phrase ellipsis, by delaying the introduction of **both** internal arguments until they are discharged in their case positions. For comparison, the sentence *Jesús was criticized*, derived already in Figure 13, has now another derivation, shown in Figure 18.

$$\textbf{assume}(A) \rightarrow \quad A;\ [_{\delta}\ \mathrm{x}\delta] \quad A \text{ has first feature } =\mathrm{x}$$

$$\textbf{discharge}(A;\ [_{\text{-}y\delta}\ \gamma]\ ,B) = \begin{cases} A;B & B \text{ survives} \\ [_A\ B\ A] & B \text{ has no more features} \\ & \\ \multicolumn{2}{l}{\text{In both cases, } A \text{ has first feature } +\mathrm{y},} \\ \multicolumn{2}{l}{\text{and } B\text{'s feature bundle begins with } \gamma} \end{cases}$$

Figure 17. Cases of **assume** and **discharge**

5.2 Antecedent-contained deletion

Although introduced to deal with passive-passive ellipsis, hypothetical reasoning allows for solutions to other, well-known elliptical puzzles. Cases of antecedent-contained deletion, as in 34, pose problems for surface-oriented perspectives on ellipsis resolution, as the ellipsis site seems embedded in its antecedent.

(34) Jesús praised every boy that Adam did ~~*praise*~~.

23. Otherwise, we are in a position where the hypothesis is erroneously sent to spell-out, which would cause the derivation to crash.

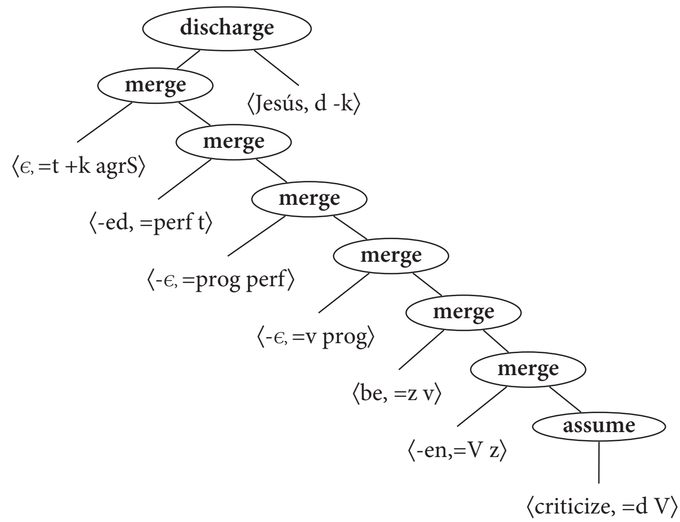


Figure 18. A second derivation tree for *Jesús was criticized*

In particular, for the derivational approach, antecedent-containment is impossible; a (finite) subtree cannot be identical with one of its proper parts. Thus the problem that arises more generally is quite acute in the present setting. The standard remedy to antecedent-containment (which we are forced to by the DIH) is to deny that it exists. In semantic identity theories, this is done in virtue of the fact that the offending DPs are interpreted outside of where they appear on the surface, and in (derived) syntactic identity theories, this is done by moving the offending DP to a point where it is no longer contained in the antecedent. Hypothetical reasoning allows us to do just that, but derivationally. To see this, let us extend our fragment with the lexical items in Figure 19, which will allow us to derive relative clauses. These lexical items implement a raising analysis of relative clauses.

⟨every, =n d -k⟩	⟨boy, n⟩
⟨that, =agrS +rel n⟩	⟨ϵ, =n d -k -rel⟩

Figure 19. Lexical items for relativization

(Kayne, 1994), where the head of the relative clause is formed by merging a noun phrase with a null relative determiner.[24]

24. In Kobele (2006), the null relative determiner is given a simple conjunctive semantics, representable as the lambda term $\lambda P_{et}.\lambda Q_{et}.\lambda x_e.\ P\ (x) \wedge Q(x)$. After combining with the head

1‴. **merge**(⟨ϵ, =n d -k -rel⟩,⟨boy, n⟩)

$[_{\text{d-k -rel}}\ \epsilon$ boy]

The resulting DP has an additional feature (-rel) which causes it to survive past its case checking position. The -rel feature is checked in the specifier of relative *that*, which selects AgrSPs as complements.

2‴. **merge**(⟨that, =agrS +rel n⟩, $[_{\text{agrS}}$ Adam praised];$[_{\text{-rel}}$ boy])

$[_{\text{+rel n}}$ that [Adam praised]]; $[_{\text{-rel}}$ boy]

3‴. **move**(2‴)

$[_{\text{n}}$ boy [that [Adam praised]]]

Note that the expression above has the same feature bundle (n), and therefore distribution, as a common noun (and thus can itself serve as the head of another relative clause).

In the grammar fragment developed in this paper (along the general lines of the 'standard' minimalist analysis), DPs have two positions: their base, or θ, position, and their case position. With hypothetical reasoning, we can thus either insert DPs into the derivation in their θ position (using **merge** and **move**), or in their case position (using **assume** and **discharge**). In order to derive sentence 34, both the embedded DP (the relative clause head $[_{\text{d-k -rel}}\ \epsilon$ boy]), as well as the entire relative clause DP ($[_{\text{d-k}}$ every boy that Adam praised]), must be introduced in their case positions. This allows for the VPs of both clauses to be identical, in this case up to the point shown in Figure 20.[25] A derivation of sentence 34 is as

noun *boy* in 1‴, it denotes the restrictive function on predicate meanings $\lambda Q_{et}.\lambda x_e.\ \mathbf{boy}(x) \wedge Q(x)$. The relative clause in 3‴ can be shown to denote (as is standard) the predicate below, which holds of an entity just in case it is both a boy, and was praised by Adam.

$\lambda x_e.\ \mathbf{boy}(x) \wedge \mathbf{praise}(x)(\mathbf{adam})$

25. The natural extension of the fragment here to control and raising predicates does not allow for the generation of ACD sentences such as i, where the ellipsis site extends higher than the case position of the object.

i. Jesús wants to eat every jalapeño that Adam does ~~want to eat~~.

This can be resolved by postulating a higher movement position, in **both** clauses, that is above the ellipsis site. It is standard to link this position to QR, that is, to movement for scope (Fiengo & May, 1994). However, it seems that this may not be the whole story (Tanaka, 2005; Yoshida, 2008). Another possibility is to relax the restriction that the **discharge** operation applies before a hypothesis reaches its ultimate landing position. I leave this question to further study.

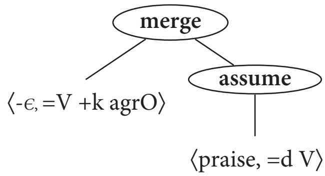


Figure 20. The identical subderivations of the antecedent and ellipsis site in ACD

follows. We begin by simultaneously building up two copies of the verb phrase shown in Figure 20.[26]

1. a. **assume**(⟨praise, =d V⟩)
   b. **assume**(⟨praise, =d V⟩)

$[_V \text{ praise } t_d]; [_{\text{-k}} \text{ d -k}]$

2. a. **merge**(⟨ϵ, =V +k agrO⟩,1a)
   b. **merge**(⟨ϵ, =V +k agrO⟩,1b)

$[_{\text{+k agrO}} \text{ praise-}\epsilon\, [t_V\ t_d]]; [_{\text{-k}} \text{ d -k}]$

Simplifying, we write the twice derived expression above as the below.

$[_{\text{+k agrO}} \text{ praise-}\epsilon]; [_{\text{-k}} \text{ d -k}]$

At this point, we have derived two copies of the above derived phrase, one of which we render phonologically null (deletion under derivational identity with the other copy).

3. **delete**(2a)

$[_{\text{+k agrO}} \text{ \sout{praise-}}\sout{\epsilon}]; [_{\text{-k}} \text{ d -k}]$

We continue building the relative clause with the deleted VP, first dischargeing the hypothesis $[_{\text{-k}}$ d -k] with the relative clause head *boy* derived in 1‴, and then merging in the subject, the tense and aspect heads, AgrS, and the relative determiner.

4. **discharge**(3, $[_{\text{-k -rel}}$ boy])

$[_{\text{agrO}}\ t_k\ [\text{\sout{praise}}\sout{\epsilon}]]; [_{\text{-rel}} \text{ boy}]$

5. **merge**(⟨ϵ, =agrO =d v⟩,4)

$[_{\text{=d v}} \text{ \sout{praise-}}\sout{\epsilon}\text{-}\epsilon\, [t_k\, [t_{agrO}]]] ; [_{\text{-rel}} \text{ boy}]$

26. Synchronizing derivations (as outlined in Kobele (2006)) allows for copying without necessitating inspection of the internal structure of the expressions which are being copied. The idea is simple: if you do the same thing to identical objects, the result is the same (i.e., the resulting objects are also identical).

6. **merge**(5,⟨Adam, d -k⟩)

   [$_v$ t_d [~~praise~~-ϵ [t_k[t_{agrO}]]]]; [$_{\text{-rel}}$ boy], [$_{\text{-k}}$ Adam]

7. **merge**(⟨ϵ, =v prog⟩,6)

   [$_{\text{prog}}$ ~~praise~~-ϵ-ϵ [t_d [t_v [t_k [t_{agrO}]]]]]; [$_{\text{-rel}}$ boy], [$_{\text{-k}}$ Adam]

8. **merge**(⟨ϵ, =prog perf⟩,7)

   [$_{\text{perf}}$ ~~praise~~-ϵ-ϵ-ϵ [t_{prog}]]; [$_{\text{-rel}}$ boy], [$_{\text{-k}}$ Adam]

9. **merge**(⟨-ed, =perf t⟩,8)

   [$_{\text{t}}$ ~~praise~~-ϵ-ϵ-ϵ-ed [t_{perf} [t_{prog}]]]; [$_{\text{-rel}}$ boy], [$_{\text{-k}}$ Adam]

   ⇝ [$_{\text{t}}$ [~~praise~~-ϵ-ϵ-ϵ-ed]; [$_{\text{-rel}}$ boy], [$_{\text{-k}}$ Adam]

10. **merge**(⟨ϵ, =t +k agrS⟩,9)

    [$_{\text{+k agrS}}$ ϵ [~~praise~~-ϵ-ϵ-ϵ-ed]]; [$_{\text{-rel}}$ boy], [$_{\text{-k}}$ Adam]

11. **move**(10)

    [$_{\text{agrS}}$ Adam [ϵ [~~praise~~-ϵ-ϵ-ϵ-ed]]]; [$_{\text{-rel}}$ boy]

12. **merge**(⟨that, =agrS +rel n⟩,11)

    [$_{\text{+rel n}}$ that [Adam [ϵ [~~praise~~-ϵ-ϵ-ϵ-ed]]]]; [$_{\text{-rel}}$boy]

13. **move**(12)

    [$_{\text{n}}$ boy [that [Adam [ϵ [~~praise~~-ϵ-ϵ-ϵ-ed]]]]]

The expression derived above is an NP, and we will combine it with every in the next step to form a DP. Because the expression takes up so much space on paper, I will abbreviate it even more severely than usual, relying on an unspecified mechanism to convert the stranded affix *-ed* to *did*.

14. **merge**(⟨every, =n d -k⟩,13)

    [$_{\text{d -k}}$ every [boy [that [Adam [ϵ ~~praise~~-ϵ-ϵ-ϵ-ed]]]]]]

    ⇝ [$_{\text{d -k}}$ every boy that Adam did]

At this point, we use the second of the two synchronously derived expressions as the matrix verb.

15. **discharge**(2b,14)

    [$_{\text{agrO}}$ every boy that Adam did [praise-ϵ]]

We continue the derivation as before, merging Jesus as the subject.

16. **merge**(⟨-ϵ, =agrO =d v⟩,15)

  $[_{=\mathrm{d\ v}}$ praise-ϵ-ϵ [every boy that Adam did [t_{agrO}]]]

17. **merge**(16,⟨Jesús, d -k⟩)

  $[_{v}\ t_{d}$ [praise-ϵ-ϵ every boy that Adam did [t_{agrO}]]]; $[_{\text{-k}}$ Jesús]

18. **merge**(⟨ϵ, =v prog⟩,17)

  $[_{\mathrm{prog}}$ praise-ϵ-ϵ-ϵ [t_{d} [t_{v} [every boy that Adam did [t_{agrO}]]]]]; $[_{\text{-k}}$ Jesús]

  ⇝ $[_{\mathrm{prog}}$ praise-ϵ-ϵ-ϵ every boy that Adam did]; $[_{\text{-k}}$ Jesús]

19. **merge**(⟨ϵ, =prog perf⟩,18)

  $[_{\mathrm{perf}}$ praise-ϵ-ϵ-ϵ-ϵ [t_{prog} every boy that Adam did]]; $[_{\text{-k}}$ Jesús]

20. **merge**(⟨-ed, =perf t⟩,19)

  $[_{\mathrm{t}}$ praise-ϵ-ϵ-ϵ-ϵ-ed [t_{perf} [t_{prog} every boy that Adam did]]]; $[_{\text{-k}}$ Jesús]

  ⇝ $[_{\mathrm{t}}$ praise-ϵ-ϵ-ϵ-ϵ-ed every boy that Adam did]; $[_{\text{-k}}$ Jesús]

21. **merge**(⟨ϵ, =t +k agrS⟩,20)

  $[_{\mathrm{+k\ agrS}}$ ϵ [praise-ϵ-ϵ-ϵ-ϵ-ed every boy that Adam did]]; $[_{\text{-k}}$ Jesús]

22. **move**(21)

  $[_{\mathrm{agrS}}$ Jesús [ϵ [praise-ϵ-ϵ-ϵ-ϵ-ed every boy that Adam did]]]

6. Conclusion

In derivational grammar formalisms, the derivation itself can be seen as structuring expressions. By formulating the derivational identity hypothesis, the hypothesis that the structure of the ellipsis site is an exact derivational analogue of its antecedent, I have shown that ellipsis can be used as a powerful probe into the underlying mechanisms of grammar. Recent work on ellipsis by Chung (2006) has suggested that we need to make reference to syntactic structure in order to adequately describe the ellipsis data, and has proposed an account based on (partial) identity of numerations. As remarked in footnote 11, in the context of minimalist grammars, numerations and derivation trees are interchangeable. Although we can view Chung's proposal as being related to the one made here, the statement in terms of identity of derivations is general enough to apply independently of grammar formalism, and thus clear enough to interpret its import on data in a meaningful way.

In tandem with the DIH, ellipsis data puts pressure on a grammar formalism to provide rich derivational structures. In order to accommodate the attested antecedence possibilities, we were led to introduce hypothetical reasoning as a dual of movement. Hypothetical reasoning can be thought of as a derivational version of trace deletion (Sauerland, 1998; Fox, 1999), which acts as a syntactic version of the lambda abstraction used by Sag (1976) to resolve the same issues we confronted here, and introduces a symmetry of sorts into the derivational process, allowing dependencies to be satisfied before an expression is introduced into a derivation. The DIH requires that identity be computed over derivational subtrees. Hypothetical reasoning enables us to 'trick' the formal system into allowing us to, in a very restricted way, compute identity over derivational *contexts*. It is of course possible to weaken the DIH from computing identity over subtrees to (sub-)contexts, and obviate thereby the need for hypothetical reasoning.

Many systematic deviations from surface identity between the (hypothesized) ellipsis site and antecedent have been unearthed over the years (see among many others, Ross (1969); Hardt (1993); Fiengo and May (1994); Chung et al. (1995); Potsdam (1997); Merchant (2001)). Adopting the DIH gives us a principled way to view these apparent deviations from identity between ellipsis site and an antecedent; instead of being problematic exceptions to a rule, they provide information not only about which construction types must be derivationally related to which others, but also about what kinds of structure building operations must be part of an adequate theory of grammar. Here I have provided a case study of sorts in one well-known such deviation, presenting by example the general strategy for a response consistent with the DIH. There remain many more. At the very least, the DIH promises a principled way of thinking about ellipsis, and of approaching its analysis.

References

Arregui, A., Clifton, C. Jr., Frazier, L. & Moulton, K. 2006. Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language* 55: 232–246.

Bach, E.W. 1980. In defense of passive. *Linguistics and Philosophy* 2: 297–341.

Brody, M. 2000. Mirror theory: Syntactic representation in perfect syntax. *Linguistic Inquiry* 31(1): 29–56.

Chomsky, N. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2(3): 113–124.

Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge MA: The MIT Press.

Chomsky, N. 1995. *The minimalist program*. Cambridge MA: The MIT Press.

Chung, S. 2006. Sluicing and the lexicon: The point of no return. In *General Session on Parasession on Prosodic Variation and Change, Berkeley, CA*, [BLS 31], R.T. Corver & Y. Kim (Eds), 73–91. Berkeley CA: Linguistics Society.

Chung, S., Ladusaw, W.A. & McCloskey, J. 1995. Sluicing and logical form. *Natural Language Semantics* 3(3): 239–282.

Elbourne, P. 2002. Situations and individuals. Ph.D. dissertation, MIT.

Fiengo, R. & May, R. 1994. *Indices and identity*. Cambridge MA: The MIT Press.

Fox, D. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30(2), 157–196.

Frampton, J. & Gutmann, S. 2002. Crash-proof syntax. In *Derivation and explanation in the minimalist program*, S.D. Epstein & T.D. Seely (Eds), 90–105. Oxford: Blackwell.

Frey, W. & Gärtner, H.-M. 2002. Scrambling and adjunction in minimalist grammars. In *Proceedings of formal grammar 2002*, G. Jäger, P. Monachesi, G. Penn, & S. Wintner (Eds), 41–52.

Gärtner, H.-M. & Michaelis, J. 2007. Some remarks on locality conditions and minimalist grammars. In *Interfaces + Recursion = Language?* [Studies in Generative Grammar 89], U. Sauerland & H.-M. Gärtner (Eds), 161–195. Berlin: Mouton de Gruyter.

Hale, J.T. 2003. Grammar, uncertainty and sentence processing. Ph.D. dissertation, The Johns Hopkins University.

Hale, J.T. & Stabler, E.P. 2005. Strict deterministic aspects of minimalist grammars. In *Logical Aspects of Computational Linguistics* [Lecture Notes in Computer Science 3492, Ch. 11], P. Blache, E.P. Stabler, J. Busquets & R. Moot (Eds), 162–176. Berlin: Springer.

Hardt, D. 1993. Verb phrase ellipsis: Form, meaning, and processing. Ph.D. dissertation, University of Pennsylvania.

Harkema, H. 2001. Parsing minimalist languages. Ph.D. dissertation, University of California, Los Angeles.

Kayne, R. 1994. *The antisymmetry of syntax*. Cambridge MA: The MIT Press.

Keenan, E. 1980. Passive is phrasal not (sentential or lexical). In *Lexical Grammar* [Publications in Language Sciences 3, Ch.7], T. Hoekstra, H. van der Hulst & M. Moortgat (Eds), 181–214. Dordrecht: Foris.

Kobele, G.M. 2002. Formalizing mirror theory. *Grammars* 5(3): 177–221.

Kobele, G.M. 2005. Features moving madly: A formal perspective on feature percolation in the minimalist program. *Research on Language and Computation* 3(4): 391–410.

Kobele, G.M. 2006. Generating copies: An investigation into structural identity in language and grammar. Ph.D. dissertation, University of California, Los Angeles.

Kobele, G.M. 2007. A formal foundation for A and A-bar movement in the minimalist program. In *Mathematics of Language* 10, M. Kracht, G. Penn & E.P. Stabler (Eds). Los Angeles CA: UCLA.

Kobele, G.M., Kim, C.S. , Hale, J.T. & Runner, J.T. 2008. A processing model for ungrammatical VP ellipsis. Paper presented at the 2008 CUNY Conference on Human Sentence Processing.

Koizumi, M. 1995. Phrase structure in minimalist syntax. Ph.D. dissertation, MIT.

Kracht, M. 2007. The emergence of syntactic structure. *Linguistics and Philosophy* 30(1): 47–95.

Lees, R.B. 1960. *The grammar of English nominalizations*. The Hague: Mouton.

Lobeck, A. 1995. *Ellipsis: Functional heads, licensing, and identification*. New York NY: OUP.

MacSwan, J. 2000. The architecture of the bilingual language faculty: Evidence from intrasentential code switching. *Bilingualism: Language and Cognition* 3(1): 37–54.

Manzini, M.R. & Roussou, A. 2000. A minimalist theory of A-movement and control. *Lingua* 110(6), 409–447.

Matushansky, O. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37(1), 69–109.

Merchant, J. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis* [Oxford Studies in Theoretical Linguistics 1]. New York NY: OUP.

Merchant, J. 2007. Voice and ellipsis. Ms, University of Chicago.
Merchant, J. 2008. An asymmetry in voice mismatches in VP-ellipsis and pseudogapping. *Linguistic Inquiry* 39(1): 169–179.
Michaelis, J. 2001. On formal properties of minimalist grammars. Ph.D. dissertation, Universität Potsdam.
Murguía, E. 2004. Syntactic identity and locality restrictions on verbal ellipsis. Ph.D. dissertation, University of Maryland.
Parsons, T. 1978. Pronouns as paraphrases. Ms.
Philips, C. 2003. Linear order and constituency. *Linguistic Inquiry* 34(1): 37–90.
Potsdam, E. 1997. English verbal morphology and VP ellipsis. In *Proceedings of the 27th Meeting of the North East Linguistic Society, Amherst MA*, 353–368.
Ross, J.R. 1969. Guess who? In *Papers from the Fifth Regional Meeting of the Chicago Linguistics Society (CLS)*, Vol. 5, R. Binnick, A. Davison, G. Green & J. Morgan (Eds), 252–286. Chicago IL: CLS.
Sag, I.A. 1976. Deletion and logical form. Ph.D. dissertation, MIT.
Sauerland, U. 1998. The meaning of chains. Ph.D. dissertation, MIT.
Stabler, E.P. 1997. Derivational minimalism. In *Logical Aspects of Computational Linguistics* [Lecture Notes in Computer Science 1328], C. Retoré (Ed.), 68–95. Berlin: Springer.
Stabler, E.P. 2001. Recognizing head movement. In *Logical Aspects of Computational Linguistics* [Lecture Notes in Artificial Intelligence 2099], P. de Groote, G.F Morrill & C. Retoré (Eds), 254–260. Berlin: Springer.
Stabler, E.P. 2003. Comparing 3 perspectives on head movement. In *Syntax at sunset* 3: *Head movement and syntactic theory* [UCLA/Potsdam Working Papers in Linguistics 10], A. Mahajan (Ed.), 178–198.
Stabler, E.P. 2006. Sidewards without copying. In *Proceedings of the 11th Conference on Formal Grammar*, P. Monachesi, G. Penn, G. Satta & S. Wintner (Eds), 133–146. Stanford CA: CSLI.
Stabler, E.P., Collier, T. C., Kobele, G.M., Lee, Y. Lin, Y., Riggle,J., Yao, Y. & Taylor, C.E. 2003. The learning and emergence of mildly context sensitive languages. In *Advances in Artificial Life* [Lecture Notes in Computer Science 2801], W. Banzhaf, T. Christaller, P. Dittrich, J.T. Kim & J. Ziegler (Eds), 525–534. Berlin: Springer.
Steedman, M. 2000. *The Syntactic Process*. Cambridge MA: The MIT Press.
Stroik, T. 2008. *Locality in minimalist syntax.* Cambridge MA: The MIT Press.
Tanaka, H. 2005. Antecedent-contained deletion without quantifier raising. In *York Papers in Linguistics,* Vol. 3, G. Hicks & G. Walker (Eds), 173–184.
Tanenhaus, M.K. & Carlson, G.N. 1990. Comprehension of deep and surface verb phrase anaphors. *Language and Cognitive Processes* 5(4): 257–280.
Wadler, P. 1990. Deforestation: Transforming programs to eliminate trees. *Theoretical Computer Science* 73: 231–248.
Yoshida, M. 2008. Antecedent contiained sluicing. Ms, Northwestern University.

Evidence for Survive from covert movement

Winfried Lechner*
University of Athens

This paper pursues two goals. First, it motivates a particular view of the Survive principle. Concretely, it is suggested to interpret the Survive principle as the syntactic instance of a more general push-up mechanism that is responsible for triggering movement induced by type incompatibility on the semantic side. Second, I identify a particular set of properties that the Survive analysis predicts for configurations involving multiple covert movements. These diagnostics, which help to discriminate between survive and Attract based models of dislocation, are argued to be manifest in scope restrictions on double object constructions and inverse linking. The critical factor setting apart the two models consists in the observation that only the Survive principle is able to express ordering restrictions between different types of movements (Case driven movement vs. QR) in a natural way. The resulting analysis also supports the phonological theory of QR.

1. Introduction

Traditional historical phonology distinguishes between *drag chains* and *push chains*. In drag chains, diachronic change is triggered by a gap in the system which is filled by a newly emerging element (phoneme). If the gap is thought of as an external *attractor*, the fundamental properties of this model closely resemble an Attract based theory of syntactic dislocation, in which a higher attracting head induces movement of a lower category (Chomsky 1995, 2005, a.o.). By contrast, changes which are associated with push chains are initiated by the emergence of new elements that expel parts of the system from their original position. Conceived this way, push chains create effects similar to those which *The Survive Principle* (Stroik 1999, 2009) postulates for syntactic dependencies.[1] Or, put the other way round, the Survive Principle imports a concept similar to push chains into syntax.

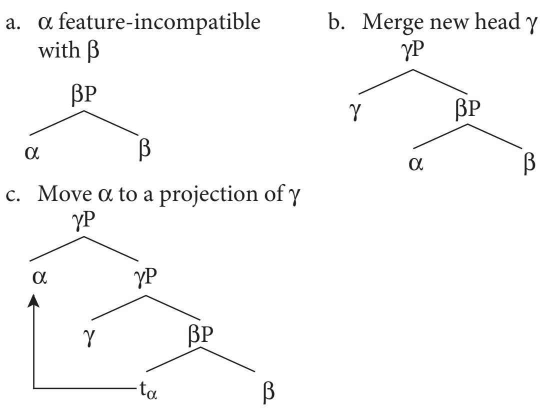
*I am indebted to Elena Anagnostopoulou and Mike Putnam for helpful discussion, and would like to thank Klaus Abels for pointing out a flaw in an earlier version of the analysis of scope freezing. A revised, extended version of this contribution is available at LingBuzz (http://ling.auf.net/lingbuzz/).

1. On push chains in syntax see also Preminger (2008); Putnam (2007); van Riemsdijk (1997) and Stroik (1999, 2009), among others.

In essence, the Survive Principle (henceforth TSP) formulates an alternative for motivating syntactic dislocation. Unlike Attract based models, which make movement contingent upon properties of higher heads, TSP locates the trigger for dislocation exclusively in the relation between the moved category and its local environment. Concretely, TSP states that if a node α is not feature compatible[2] with its sister node, α is pushed into a higher position in order to 'survive' the consequences of feature mismatch. Moreover, the target of movement is determined by the next new head which is introduced into the derivation. One way to make the TSP explicit is as in (1):

(1) *The Survive Principle (first version)*
For any nodes α, β and γ:
If α is not feature compatible with its sister node β,
(i) merge the projection of β with a new head γ and
(ii) remerge α with a projection of γ.

(2) schematically illustrates how TSP triggers dislocation of an α which is not feature-compatible with its sister node β (2a)[3]. In an initial step of the derivation, a new head γ is externally merged with a projection of β (2b). As a consequence of TSP, the incompatible category α is expelled from its original position, moving to a projection of the newly merged head γ (2c):

(2)


2. Roughly, two nodes are feature incompatible if their original feature matrices prior to any checking operation have an empty intersection. I will not attempt to make the notion of feature incompatibility more precise here. Stroik's original version of TSP is (i):

(i) The SURVIVE Principle (Stroik, 2009)
If Y is a SO [Syntactic Object] in an XP headed by X and Y has an unchecked feature incompatible with (i.e., cannot potentially be checked by) the features of X, Y must Remerge from the WorkBench with [a projection of; WL] the next head Z that c-commands XP.

3. For typographic convenience, complements are drawn on the left.

The present paper explores a possible venue for locating evidence in support of TSP based on interpretive properties that are also syntactically encoded. In particular, I will propose that certain restrictions on relative quantifier scope in English can be given a simple explanation if TSP is adopted, but require additional assumptions, which essentially mimic the effects of TSP, in the standard Attract based model.

The analyses to be presented cover two different phenomena: scope freezing in the double object construction (Barss & Lasnik 1986; Richards 2001a; Bruening 2001; Sauerland 2000), and scope restrictions in contexts involving Inverse Linking (Larson 1987; Heim & Kratzer 1998; Sauerland 2000). As a common thread, both analyses will be seen to rest on the assumption of intermediate traces. Such traces are automatically generated by the TSP model, but not by theories that use Attract as a trigger for movement. This finding provides support for integrating the concept of TSP into the syntactic component.

The paper is structured as follows. In Section 2, I discuss relevant background assumptions pertaining to the specific implementation of TSP that will be used (2.1) and the particular model of the grammar (2.2). Section 3 introduces two scope restrictions and develops a common TSP analysis of these phenomena. Section 4 summarizes and comments on the results.

2. Background assumptions

2.1 The survive principle and type driven interpretation

The principle of *type driven interpretation* (TDI; see (3)), to be adopted here, expresses the widely endorsed view that certain covert movements are induced by the need to repair type incompatibilities:[4]

(3) *Type driven interpretation*
If a category α is not type compatible with its sister node β, move α to the next higher position type compatible with α.

Interestingly, the formulation of TDI, which triggers covert dislocation, bears an uncanny resemblance to TSP in (1). To bring out the similarities more transparently,

4. For discussion see e.g., Heim and Kratzer (1998). α is type compatible with β if either (i) or (ii) applies:

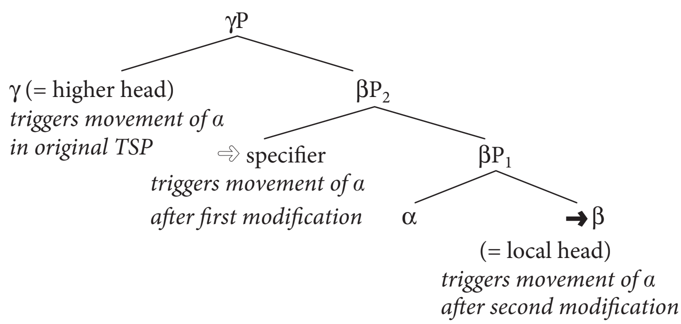
(i) $[\![\alpha]\!] \in D_{\langle\tau,\sigma\rangle}$ and $[\![\beta]\!] \in D_{\langle\tau\rangle}$ (Function Application, with β as argument)
(ii) $[\![\alpha]\!] \in D_{\langle e,t\rangle}$ and $[\![\beta]\!] \in D_{\langle e,t\rangle}$ (Predicate Modification)

(1) can be rephrased as in (4) below:

(4) *The Survive Principle (first version, paraphrased)*
If a category α is not feature compatible with its sister node β, move α to the next higher position projected by a new head γ.

In particular, both principles include an antecedent clause defining incompatibility (type or feature wise) and specify a movement strategy to resolve this conflict in the consequent of the conditional. But while the antecedent clauses of (3) and (4) are identical *modulo* the distinction 'type' vs. 'feature', the consequents are given a different wording in each principle. For TSP, the escape strategy consists in movement to a position above the next head, while for TDI, movement must target a node of suitable type.

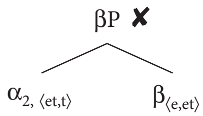
In what follows, I argue for three changes in the definition of TSP that will render this relation more symmetric and at the same time contributes to an overall simplification of the Survive model. The first change will make TSP look more like TDI, thereby removing an imbalance in these two otherwise quite similar principles. The second modification eliminates a stipulation from TSP, rendering it more general than the original version. These two changes affect TSP *incrementally* in that they expand (4), which creates intermediate landing sites only upon insertion of new heads (represented by γ in (5)) to systems that add specifiers (see ⇨ in (5)) and local heads (see ➜ in (5)), respectively, to the list of triggers for movement:

(5)


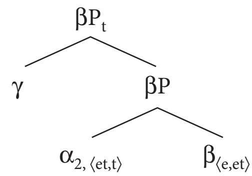
Finally, a last revision extends the range of categories affected by TSP by generalizing the definition of 'compatibility' to include type mismatches, in addition to feature incompatibility.

Turning to the first modification, the contrast between TSP and TDI observed above is a function of another, deeper difference between the two principles. While TDI assumes that heads combine with their complements in the same way in which specifiers combine with heads (or, more precisely, with nodes containing heads and their complements), TSP ignores specifiers all together. To illustrate the relevance of specifiers for TDI, consider the text book case of type-driven movement of an object quantifier in (6).

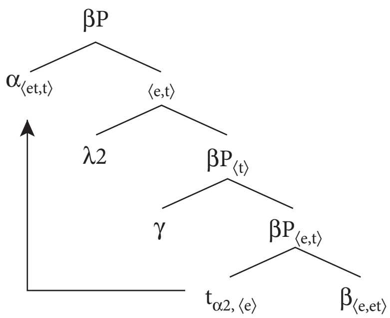
(6) a. α type incompatible with β



b. Merge specifier of β (=Y)



c. Move α to projection of β



In (6a), the generalized quantifier type expression α_2 (type ⟨et,t⟩), originates as a sister node to a transitive verb (= β). On standard assumptions, α needs to attach next to a node of type t in order to be able to combine with the rest of the clause in semantics. Such a node is provided by addition of the *specifier* of β (= γ), i.e., the subject, in (6b).[5] Thus, specifiers are instrumental for TDI in that they create suitable landing site for QR (see (6c)).

I would like to suggest to generalize TSP in the same direction, such that TSP driven movement is not only triggered by new heads, but also by the addition of specifiers. One way to achieve this is made explicit by the revised version of TSP in (7):

(7) *The Survive Principle (second version)*
For any nodes α, β and γ:
If α is not feature compatible with its sister node β,
(i) merge the projection of β with a new category γ, resulting in δ, and
(ii) remerge α with δ.

(7) derives the desired effect of pushing α across newly merged heads as well as across newly added specifiers by not restricting the value of γ to heads. To see how this works, assume again that α is in complement position of a verb. If the projection of β (= VP) merges with a head γ (e.g., v^0), then α will attach to a projection of v^0. Subsequently

5. I assume, as is standard (implicit) practice, that phases and quantifiers make their semantic type visible to syntax.

merging VP with a specifier (e.g., a subject) results in movement of α to a position adjoined to the minimal node containing the specifier and vP, as illustrated by (6c).

The second change of the TSP pertains to the role of heads and eliminates an extrinsic statement from the Survive model. In its current incarnation (7), TSP demands that α moves to the next higher position projected by a new head γ. However, confining movement to contexts in which a *new* head – i.e., a category that is not yet part of the derivation – is inserted only adds a stipulation to TSP and is, as far as I can see, not motivated by empirical considerations. I will therefore assume that movement is triggered after any application of merge, irrespective of whether the externally merged category is a new head or the local head that α has been first merged with, or a specifier. This can be informally summarized as in (8), which provides the pre-final version of TSP:[6]

(8) *Survive Principle (pre-final version)*
For any nodes α, β and γ, such that γ is the mother of β and β c-commands α:
If α is not feature compatible, then merge α with γ.

The definition (8), which essentially mandates that any feature-incompatible α move to the next branching node in the tree, includes two subcases. If α serves as the complement of a head, say V^0, then α will be pushed to a VP-adjoined position. If, on the other hand, α is adjoined to a maximal projection, say vP, which is still to combine with a specifier, e.g., the subject, then α will land in an outer adjunct to vP.

The last modification of TSP removes another unnatural stipulation from the definition of Survive, resulting in further reduction of redundancy. In particular, I would like to suggest that 'compatibility' is not restricted to 'feature compatibility', but extended to also comprise *type* incompatibility. The final definition of TSP is given in (9):

(9) *Survive Principle (final version)*
For any nodes α, β and γ, such that γ is the mother of β and β c-commands α:
If α is not feature or type compatible, then merge α with γ.

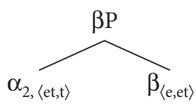
6. An alternative definition is provided by (i):

(i) For any nodes α, β and γ:
If α is not feature compatible with its sister and (i) and (ii) hold, then adjoin α to the result of merging β with γ.
(i) γ reflexively dominates α and
(ii) γ is merged with β and

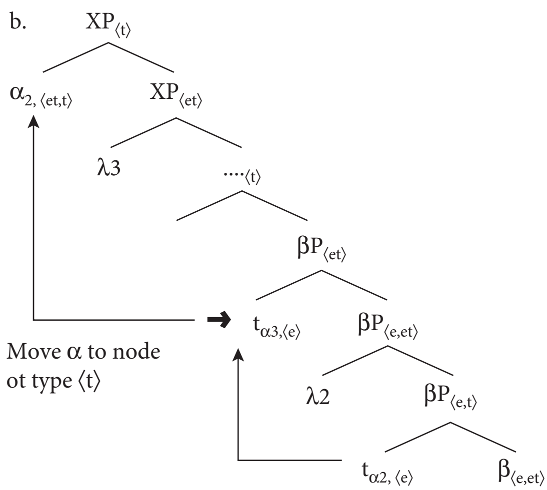
(9) can be understood as an instruction to successively adjoin a category to the root upon insertion of new nodes in the derivation, irrespective whether the category is feature or type incompatible with its sister node. Thus, this final revision enables the Survive principle to also react to type mismatches, essentially subsuming the work of TDI. Moreover, by stating TSP as in (9), it becomes possible to recognize the non-accidental similarities between syntactic displacement and TDI. On the present conception, these two conditions manifest the syntactic and the semantic side of a single, more fundamental principle, which expels semantically or syntactically incompatible categories from their position.

But the current way of treating type mismatches also differs in an important respect from the conjunction of the standard definition of TDI and the original TSP in (4). This disparity manifests itself, among others, in configurations such as (6a) (repeated below as (10a)), in which a generalized quantifier denoting category α is merged with a type incompatible head β. For such environments, (9) generates a prediction which distinguishes it from the combined effects of the original TSP in (4) and TDI. In contrast to the original TSP in (4) or TDI, (9) leads one to expect that the derivation creates intermediate landing sites for α not only *outside* βP, but also in *βP-adjoined* positions (marked by ➜):

(10) a. Type incompatibility



b.



More specifically, TSP licenses movement through βP in (10b) by having α strand an individual type trace, which may directly combine with the verb denotation. (The entire chain including the quantifier finally becomes interpretable once α attaches to a node of type t in a later movement step, also shown in (10b)). Crucially, such an intermediate position could not have been generated by TDI, because the lexical content of α is not interpretable in the intermediate βP-adjoined position. A generalized quantifier denotation cannot combine with a two-place relation without type adjustment. Actual manifestations of the structure (10b) will be encountered in Section 3 (see (20)[7]).

Two concluding remarks on the present implementation of the TSP are in order here. First, although (9) superficially differs substantially from the original formulation of TSP (1)/(4), it strictly adheres to the spirit of the Survive program, which regulates the distribution of push chains in syntax. Moreover, from a methodological perspective, the generalized version facilitates the detection of a wider range of potential Survive phenomena, rendering it thereby better suitable for the search of criteria that help to discriminate between Survive and Attract based models. Second, the current definition of TSP, on which movement always targets the next branching node in the tree, renders storing elements in the WorkBench, and retrieving them again, as in Stroik (2009) unnecessary (see Stroik 2009 for discussion). This significantly simplifies the way in which TSP treats movement.

To summarize, by eliminating stipulations from the original version of the TSP, all movements of feature or type incompatible categories proceed now via local adjunction to the minimally containing node. The final version of TSP in (9), which also subsumes the effects of TDI, accordingly represents a more general, and simpler, manifestation of Survive, which both extends the range of categories affected by TSP and the set of possible landing sites for movement. These two properties will be seen to receive empirical support from the account of scope restrictions to be given in Section 3, where they will prove instrumental for a unified analysis of the phenomena.

2.2 The phonological theory of QR

A second component that is crucial to the analysis to be developed in Section 3 is provided by what has come to be known as the *phonological theory of QR*. On this view, QR does not proceed in a separate, covert level of representation which follows overt syntax, but takes place in the stem of the derivation and interleaves with other, overt movement operations (Bobaljik 1995; Fox & Nissenbaum 1999; Bhatt & Pancheva 2004). The fact that QR does not affect the relative order of constituents is accounted for by the assumption that dislocation by *overt covert*

7. So far, I have only been able to find indirect evidence for such intermediate positions. Potential *prima facie* support for the presence of VP-adjoined traces come from licensing of bound variable readings by quantificational objects into low VP-adjuncts, as in (i):

(i) John read every book before reviewing it.

For (i), it might be argued that the variable is licensed by a VP-adjoined e-type trace of the object.

movement (Pesetsky 1998) leads to pronunciation of the tail, instead of the head, of the chain.

The phonological theory of QR offers both empirical as well as conceptional advantages over the standard LF-based model. As for the latter, it is no longer necessary to postulate a separate post-Spellout portion of the derivation. This is of advantage as the existence of such a level has commonly been inferred on the basis of the observation that LF behaves just like overt movement, with the only exception that QR does not have PF-visible effect. But then, the question arises of why to locate these silent dislocation operations into a separate component in the first place. An integrative model, which employs overt covert movement eliminates this redundancy, rendering the transition from syntax to the semantic component more transparent.

Empirically, the phonological model of QR sheds new light, among others, on the interaction between binding and coreference on the one hand, and dislocation on the other. Fox and Nissenbaum (1999) note, for instance, that overt covert movement offers a ready explanation for the observation that extraposition bleeds disjoint reference effects (Taraldsen 1981):

(11) a. *I showed him$_3$ a book [that Sam$_3$ wanted to read] yesterday
b. I showed him$_3$ a book *t* yesterday [that Sam$_3$ wanted to read]

According to Fox and Nissenbaum, extraposition consists in a two step procedure: overt covert movement of the head of the relative clause *a book*, illustrated by (12a), followed by Late Merge of the relative clause to the head, as detailed in (12b).

(12) I showed him$_3$ a book *t* yesterday [that Sam$_3$ wanted to read]
a. Step 1, overt covert movement of *a book*:
I [$_{VP}$ showed him$_3$ a book] yesterday [~~a book~~]
b. Step 2, Late Merge of relative clause:
I [$_{VP}$ showed him$_3$ a book] yesterday [~~a~~ [~~book~~][that Sam$_3$ wanted to read]]

The combination of overt QR and late merge of the relative results now in a derivation in which the name inside the relative clause has never been inside the c-command domain of the coreferential pronoun. Crucially, this analysis is dependent on the assumption central to the phonological theory of QR that inaudible and overt movement processes behave alike: both apply in the stem of the derivation, and both may serve as attachment sites for late merged relative clauses.[8]

8. Additional evidence in favor of a single output model and discussion of the phonological theory of QR can be found in Bhatt & Pancheva (2004); Bobaljik (1995); Fox (2002); Hulsey and Sauerland (2003); Nissenbaum (2000); Pesetsky (1998) and Takahashi (2006).

3. Survive and scope

In the present section, it will be demonstrated that given the assumptions specified in Section 2, a theory which includes TSP finds support from two different empirical domains related to quantifier scope interpretation.

To be specific, TSP makes possible a simple analysis of two restrictions on relative quantifier scope in English to be discussed in sections 3.1 and 3.2 below, respectively. Although both of these restrictions have already been given detailed analyses in Attract based frameworks (see Bruening 2001 and Sauerland 2000), these accounts were forced to adopt unnatural assumptions, in particular about the interpretation of subjects. Building on these studies, I propose that certain aspects of the Survive model lead to an improvement over these previous accounts in two regards. More specifically, additional machinery will be seen to be superfluous in the TSP model, where all dislocation processes – irrespective whether they affect subjects or objects – proceed in small incremental steps.

3.1 Scope freezing in double object constructions

The conditions regulating possible quantifier scope permutations in English determine that in a number of contexts, the scope order must not alter surface order of the quantificational terms. Prominently among these is the scope freezing generalization for the double object construction (Barss & Lasnik 1986; Richards 2001a; Bruening 2001).

3.1.1 *Previous accounts*

The scope freezing generalization for double object constructions makes explicit the observation that in the dative-accusative serialization, the two internal arguments are unable to change their relative scope order (ibid.):

(13) a. I gave a child each doll. $\exists \succ \forall$/*$\forall \succ \exists$ (Bruening 2001, (2a))
b. The judges awarded a (#different) athlete every medal.
*$\forall \succ \exists$ (ibid., (28c))

In the same context, the direct object may scope over the subject, though (Bruening 2001), as shown by (14). This indicates that the relevant constraint does not limit the scope taking options of direct objects *per se*, but has to be formulated in such a way that it imposes order preservation on the relation between the two internal arguments:

(14) a. A (different) teacher gave me every book. $\forall \succ \exists$ (Bruening 2001, (28))
b. At least two judges awarded me every medal. $\forall \succ$ *at least two*

According to Bruening, scope freezing is a reflex of two properties: First, quantificational DPs need to check a Q(uantificational)-feature on v^0. On this view, QR is not only motivated by TDI, but also feature driven. Second, there is a general syntactic requirement that multiple movements to a single position result in tucking in (Richards 2001a, 2001b) and the formation of crossing paths. The latter condition can be further reduced to general principles of economy (Shortest Move and Shortest Attract). Without going into the details (see Richards 2001b), the essential property of the tucking-in approach consists in an isomorphism from order of movement to asymmetric c-command relations. Whenever two categories α and β are attracted by a single head and α moves prior to β, then α asymmetrically c-commands β in the resulting output representation.

Taken together, these two assumptions derive scope freezing for (13). The indirect object (IO) is base generated in a position above the direct object (DO), and therefore closer to the attracting head v^0, which bears an attracting Q-feature (see (15a)). IO accordingly undergoes QR first, followed by movement of DO. Both QPs land in an outer specifier of vP, as shown in (15b). Moreover, since the order of movement matches the relative order of the landing sites, the two QPs are raised in a crossing dependency, preserving the base order of the internal arguments (see (15c)):

(15) a. $[_{vP}$ SUB $v_{[Q]}$ $[_{VP}$ IO_2 [DO_3]]]
b. $[_{vP}$ IO_2 $[_{vP}$ DO_3 $[_{vP}$ SUB $[_{VP}$ t_2 [t_3]]]]]
c. Mapping from base order to LF: $2 \succ 3 \Rightarrow 2 \succ 3$

Sauerland (2000) notes that Bruening's analysis is challenged by the existence of examples such as (16), which can, among others, be assigned a reading on which the subject scopally interferes inbetween the indirect and the direct object. The relevant parts of the LF are given in (17a):

(16) Two boys gave every girl a flower $\forall \succ 2 \succ \exists$ Sauerland (2000, (49))

(17) a. $[_{vP}$ IO_2 $[_{vP}$ SUB_1 $[_{vP}$ DO_3 $[_{vP}$ t_1 $[_{VP}$ t_2 [t_3]]]]]
b. Mapping from base order to LF: $1 \succ 2 \succ 3 \Rightarrow 2 \succ 1 \succ 3$

(17a) is problematic for the economy/crossing dependencies approach inasmuch as the LF representation does not preserve the order between IO and SUB (see (17b)). While SUB c-commands the IO in the base, the relations are reversed subsequent to QR.

Sauerland proposes to amend this shortcoming by adopting a definition of closeness on which the reconstructed subject and the indirect object occupy two positions at LF which are equidistant to v^0. Together with the assumption that both IO and SUB need to check a Q-feature on v^0, either IO or SUB can be the first category attracted by v^0. Since order of movement translates into asymmetric c-command, both scope orders SUB $\succ$ IO and IO $\succ$ SUB can now be generated.

Attract analyses need to espouse two rather unnatural propositions, though. First, subsequent to reconstruction, subjects need to undergo short QR from their vP-internal base position, despite their being type compatible, and therefore directly interpretable, in SpecvP. Otherwise, it would be impossible to generate the LF-representation (17a), in which the subject is sandwiched inbetween the indirect and the direct object (see Sauerland 2000). Second (and partially related), the assumption of a Q-feature in addition to the principles of TDI partially duplicates the trigger for QR. Concretely, there is only a single case – namely subjects – where the effects of the Q-feature are distinguishable from those of TDI. For object quantifiers, which need to leave VP due to type mismatch anyway, this has the consequence that postulating a Q-feature only obscures the motivation behind QR.

Note finally that Q-features are indispensable for the success of Attract based approaches, and can therefore not simply be dispensed with. More specifically, Q-features are needed for two reasons. First, they are implicated in the creation of crossing paths, thereby ensuring order preservation. And second, a Q-feature is indispensable for driving the subject from its reconstructed base into an intermediate position in (17a). (This follows from the independent assumption that no category can check features in its first merge position). As will be specified below, the Survive based model offers an alternative which does not require QR to be feature driven, and is therefore in a position to avoid these complications.

3.1.2 *Survive and scope freezing*

The greatest strength of TSP lies in its ability to derive the observation that subjects can be evaluated in an intermediate position above SpecvP that is high enough to scope over other operators such as (object) quantifiers:[9]

(18) $[_{vP}$ SUB$_1$ $[_{vP}$ QP/negation/... $[_{vP}$ t_1 ....]]]

This property will be seen to be instrumental in accounting for the scope options of double object constructions.

Earlier, it was assumed that QR proceeds in the overt part of the derivation by overt covert movement. For constructions with quantificational expressions in object position, such as (13), repeated from above, this entails that the VP-internal QPs are expelled in overt syntax, moving successive cyclically to the next higher nodes until they reach a type compatible position (i.e., a position which is sister to a propositional node such as vP).

(13) I gave a child each doll. $\exists \succ \forall / {}^*\forall \succ \exists$ (Bruening 2001, (2a))

The pertinent double object example (13) involves two internal arguments (instead of only a single one) that need to be evacuated from VP by overt covert movement.

9. For additional justification for interpreting subjects in this intermediate site see Bruening (2001, fn. 25); Johnson & Tomioka (1997); Lechner (2007); Sauerland (2000).

This raises a pair of partially related questions, which has been extensively discussed in the literature on Attract, but has received less attention in Survive based models.[10] Which category is expelled by TSP first? and where do the two QPs land? As will be specified below, the Survive model essentially replies in the same way as theories that employ Attract, the only difference being that the economy metric which underlies the answer to the second question is already part of the definition of TSP.

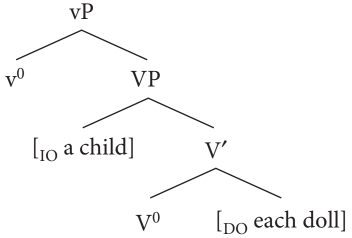
Addressing order of movement first, a theory that lacks attractors evidently cannot resort to calculating distances to an attracting head. I will therefore assume that the derivation privileges recently merged nodes over categories that have already been merged at an earlier point in the derivation. Since in bottom-up derivations, 'recency' correlates with height of attachment in the (stem of) the tree, higher categories move first. Just like in the Attract based model, this forces IO to move prior to the lower DO in structures like (13) (see also 3.2.2 for discussion).

As for the choice of landing sites, the definition of TSP (9), repeated below, already precisely determines the order in which two categories contained inside the same node will land.

(9) *Survive Principle (final version)*
For any nodes α, β and γ, such that γ is the mother of β and β c-commands α:
If α is not feature or type compatible, then merge α with γ.

The effects of (9) on the final order of IO and DO are best explicated by going over the actual derivation of (13).

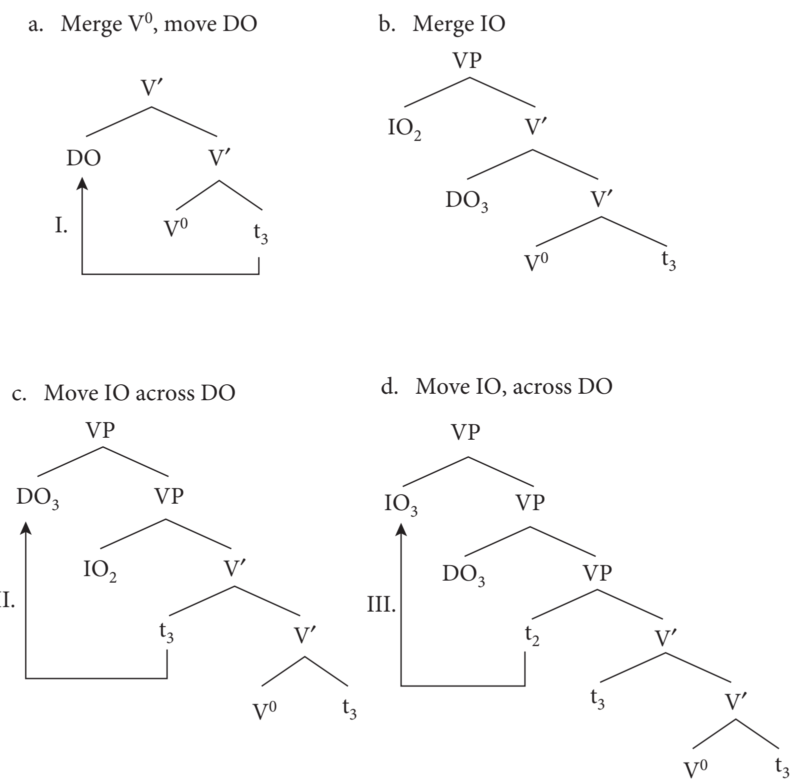
Following Kayne (1984), Beck & Johnson (2006), among many others, the indirect object of double object constructions will be taken to occupy a specifier of the same head that contains the direct object as a complement. Thus, the parse for (13) roughly looks as follows (I use V′ for typographic convenience):

(19)


10. See Stroik (2009, chapter 2) for discussion of multiple overt wh-movement.

Assembling the tree for (13) bottom-up, TSP now creates intermediate landing sites for both type incompatible object quantifiers right above every branching node that they are contained in. In (20a), the initial step triggered by (9), DO adjoins to V′, i.e., the minimal node containing V^0 and DO. Next, externally merging IO in SpecVP (see (20b)) forces movement of DO across IO, as detailed by (20c).[11] In (20d), IO then raises across DO to a higher SpecVP in response to the internal merger of DO:

(20)



At this point, it becomes apparent that some device is needed in order to terminate TSP-driven movement between IO and DO once the derivation has reached

11. Movement of DO creates a derived λ-abstract. As a result, the V^0-denotation will correctly apply to IO, and not to DO, as its external argument.

(20d). Otherwise, IO and DO would continue leap-frogging at the VP-level, resulting in endless regress. This is so as in the further course of the derivation, re-merger of DO in a position above IO (as in (20c)) inevitably triggers movement of IO, which in turn pushes DO up one more node derivation, and so on *ad infinitum*.

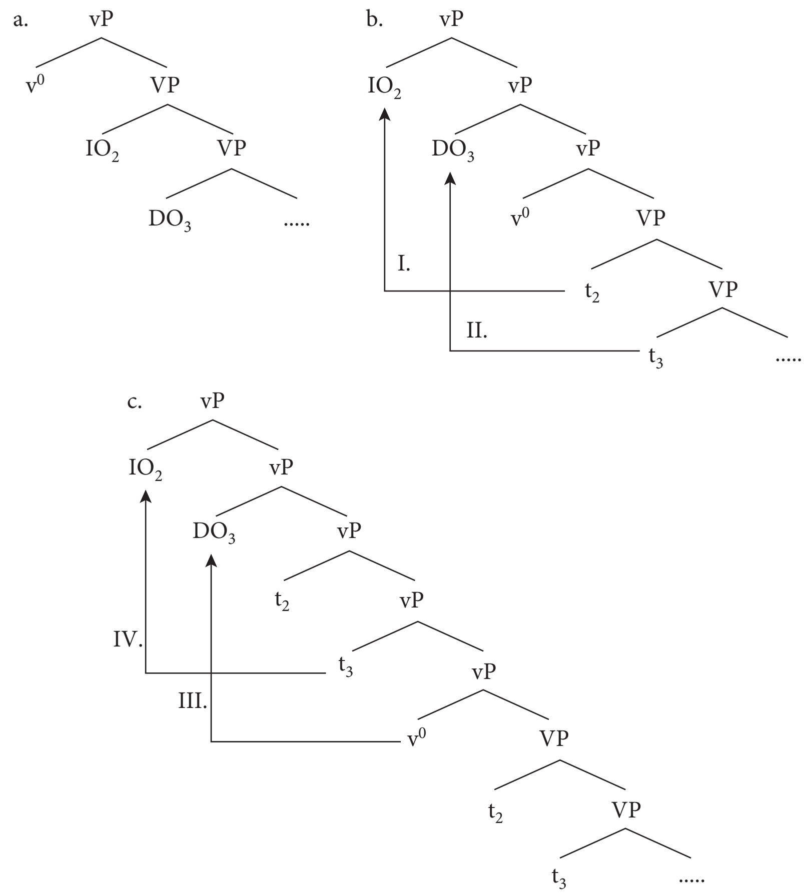
A natural strategy for avoiding these types of loops is to take into consideration that multiple locally interleaving steps of the same category as in (20) do not effect a change either in the feature or the type composition of the subtree. Suppose that in a derivation such as (20), DO has crossed over IO and IO over DO (see (20d)), and that neither of these two operations has resulted in improved feature or type compatibility. In this context, moving DO one more time across IO will equally fail to repair feature and type mismatches. Such changes are only expected if a new category is merged inbetween the two object quantifiers. Since the ultimate motivation for TSP resides in the avoidance of feature/type conflicts, and loops of the type discussed above never lead to resolution of these incompatibilities, it is natural to conclude that configurations which create such loops should be generally disregarded by the TSP. Thus, endless regress in (20) is terminated by the independent minimalist core requirement that all movement processes be motivated (*Last Resort*).[12]

Given the proviso above, the derivation of (13) finally arrives at the vP-level, as illustrated by (21a). Addition of v^0 creates now intermediate landing sites for IO and DO right above vP. In order to find out which of the two internal arguments moves first, recall that on present assumptions, the order of movement between two nodes is inversely related to the order in which these nodes have been merged. Thus, categories that are merged last will move first. As a consequence, movement of IO across v^0 precedes movement of DO across v^0 in (21b), yielding an order preserving configuration in which IO asymmetrically c-commands DO.

At this point, IO and DO end up in a new, local relationship, just like in (20b). And just like in (20b), such a configuration triggers two more movement steps, depicted in (21c). As a result, the output of Survive driven multiple dislocation does not only preserve the input order, but also produces derivations in which the higher category (IO) has been manipulated last. (This property will, among others, turn out to be instrumental for the analysis of Inverse Linking below.)

12. This restriction can be built into the definition of movement itself, e.g., by assigning to each category a local memory stack that triggers the instruction to forego movement if two subsequent movement operations are induced by merger of one and the same expression.

(21)



The computation above demonstrated that whenever two nodes are expelled from a single containing category – in this case the highest segment of vP – the higher one will move first (21b). Then, the lower category moves across the newly merged node (v^0 in 21b), 'tucking in' below the copy of the higher category. Since the two raised nodes are now located in a new environment, Survive induces two final, order preserving movement steps, which target DO first and affect the higher category IO last (21c). Borrowing a term (for different) effects from Bobaljik (1995), I will also refer to this way of deriving order preserving relations by multiple interleaving movement dependencies as *leap-frogging*.

Although the present analysis is close in spirit to the traditional Attract account, and empirically replicates the pattern of tucking-in (Richards 2001a, 2001b), it should be noted that there is no reference to 'closeness to an attractor' hidden in any of the definitions or principles underlying the account. Rather, it is the definition of TSP itself, together with the assumption that material that is merged later moves first, which provides the key to the answer for why multiple movements result in crossing, order preserving dependencies. If correct, this conception of the Survive model provides an interesting new perspective on how to derive order preservation effects in syntax.

Finally turning to the derivation of the possible scope orders in the double object construction, recall that while the two internal arguments cannot permute in scope, schematized in (22a), the subject of (16) (repeated from above) may be assigned either intermediate scope inbetween IO and DO (see (22b)) or narrow scope (see (22c)):

(16) Two boys gave every girl a flower $\forall \succ 2 \succ \exists$ Sauerland (2000, (49))

(22) Mapping from base order to scope order:
a. $*2 \succ 3 \Rightarrow 3 \succ 2$
b. $1 \succ 2 \succ 3 \Rightarrow 2 \succ 1 \succ 3$
c. $1 \succ 2 \succ 3 \Rightarrow 2 \succ 3 \succ 1$

Evidently, the leap-frogging strategy at work in (21), which produces order preserving movement, is the source for the scope freezing effect. However, the representation for (13) given under (21b) is not complete, yet, because (21b) is still missing the subject. This is amended by merging the external argument in (23a):

(23) a. $[_{vP}\ SUB_1\ [_{vP}\ IO_2\ \ [_{vP}\ DO_3\ v^0\ \ldots$
b. $[_{vP}\ IO_2\ \ [_{vP}\ SUB_1\ [_{vP}\ t_2\ \ [_{vP}\ DO_3\ v^0\ \ldots$
c. $[_{vP}\ IO_2\ [_{vP}\ DO_3\ \ [_{vP}\ SUB_1\ [_{vP}\ t_2\ \ [_{vP}\ t_3\ \ v^0\ \ldots$

In (23a), TSP continues to induce displacement of the two internal arguments because the object quantifiers need to combine with type-t expressions, the formation of which in turn requires a subject (trace). Thus, leap-frogging of IO and DO is repeated one more time. First, IO moves over the newly merged subject (23b), followed by movement of DO (23c). In (23c), the two internal arguments have now for the first time reached a position in which they are interpretable as generalized quantifiers, and therefore stop to move.

In the last relevant steps of the derivation, the subject is ejected from its base position and moves to TP. Although the subject matches type-wise with its predicate sister, it is *feature* incompatible with the position it originated in. As a result, it needs to leave SpecvP, and lands, in a first movement step, in a position right inbetween IO and DO:

(23) d. $[_{vP}\ IO_2\ \ [_{vP}\ SUB_1\ [_{vP}\ DO_3\ \ [_{vP}\ t_1\ v\ \ \ldots$

From there, it moves on to SpecTP. Unlike IO and DO, the subject quantifier is interpretable in all intermediate positions, as well as in its base, and is therefore free to undergo optional reconstruction into the slot inbetween IO and DO. If the subject reconstructs into this intermediate position, resulting in an

LF-representation akin to (23d), it will be assigned intermediate scope (22b). By contrast, narrow scope of the subject w.r.t. IO and DO ((22c)) is derived by total reconstruction of SUB into SpecvP.

To recapitulate, The Survive Principle not only accounts for scope rigidity among the internal arguments in the double object construction, but also correctly predicts the more liberal behavior of subjects. The key to the success of the TSP analysis is its ability – in conjunction with the phonological theory of QR – to create interaction between movements driven by type incompatibility and movement driven by feature incompatibility. It is this interaction which sponsors the kind of short subject movement observed in (23d). Moreover, in contrast to the Attract based model, the current analysis derives this result without the need to postulate Q-features, which duplicate the effects of TDI, or stipulating that subjects are not interpretable in SpecvP (see Section 3.1.1). If correct, these findings provide support of a Survive based conception of movement, and challenge the traditional Attract model.

Finally, note that the analysis has also consequences for the theory of QR. As discussed in 3.1.1, Fox and Nissenbaum (1999) showed that covert and overt movement both interact with late merge, and should therefore be treated alike. The evidence above supports a second, qualitatively different type of parallelism between covert and overt movement operations, which further vindicates the idea that the two types of processes are not distributed between overt syntax and LF, but take place at one and the same level. More precisely, it was seen that in contexts that involve both overt and covert dislocation, overt movement (of the subject) sometimes seems to follow covert QR (of the objects), instead of preceding it. The order of movement was measured by the possible relative scope orders, which in turn were taken to indicate the availability of intermediate reconstruction sites for the dislocated quantifier phrases. Thus, the findings above not only furnish support for an TSP- based model of dislocation, but also further strengthens the case for a phonological theory of QR.

3.2 Inverse linking

The second freezing phenomenon comes in shape of a scope condition which is operative in context that involve inverse linking, first discussed by Larson (1987) (see also Heim & Kratzer 1998: 234; Sauerland 2000). Restricting the attention to readings in which the embedded QP_3 scopes over its container QP_2, (24) admits a wide and narrow scope interpretation for the subject, but lacks the intermediate construal (24c):[13]

13. Strings of similar structure such as (i) also can be assigned the surface scope order.

(i) Every policemen$_1$ reviewed exactly two reports$_2$ about every candidate$_3$ $(1 \succ 2 \succ 3)$

I assume that this interpretation is derived by interpreting QP_3 *in-situ* (see Heim & Kratzer 1998).

(24) [$_{QP1}$ Two policemen spy on [$_{QP2}$ someone from [$_{QP3}$ every city]]]
a. $2 \succ \forall \succ \exists$
b. $\forall \succ \exists \succ 2$
c. $^*\forall \succ 2 \succ \exists$

For some reason, the subject quantifier may not intervene between the inversely linked quantifier (QP_3) and its container (QP_2). Thus, only two of the three possible mappings from surface to scope order that keep constant the order $QP_3 \succ QP_2$ are actually attested:

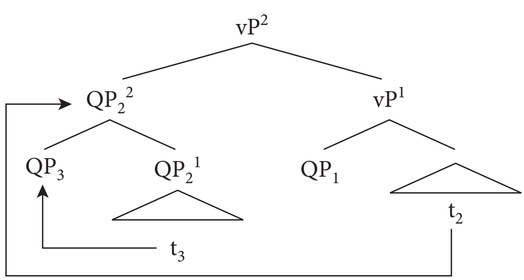
(25) Mapping from base order to scope order:
a. $1 \succ 2 \succ 3 \Rightarrow 1 \succ 3 \succ 2$
b. $1 \succ 2 \succ 3 \Rightarrow 3 \succ 2 \succ 1$
c. $^*1 \succ 2 \succ 3 \Rightarrow 3 \succ 1 \succ 2$

Sauerland (2000) provides a uniform analysis of this restriction and scope freezing in double object constructions. Although descriptively adequate, the account needs to adopt ancillary assumptions that reduce its explanatory force (see previous section for discussion). Given these inherent shortcomings, I will not further expand on Attract based explanations here, but will directly proceed to the TSP analysis (see Sauerland 2000 for details).

3.2.1 *Survive and inverse linking*

On the TSP account, essential parts of which are graphically represented by (26), insertion of the subject quantifier QP_1 (*two policemen* in (24)) induces movement of the direct object QP_2 (*someone from every city*) to an outer vP-adjoined position. Simultaneously, the embedded category QP3 *every city* is pushed out of its base inside the containing QP_2, landing as a DP-adjunct. Both movements are triggered by type mismatch. At this point of the derivation, there are two categories that could potentially undergo TSP-driven movement. The subject quantifier QP_1, which needs to escape SpecvP to avoid feature clash with its head v°, and QP_3 from its intermediate QP_2-adjoined position, which has to move in order to resolve a type mismatch:

(26)

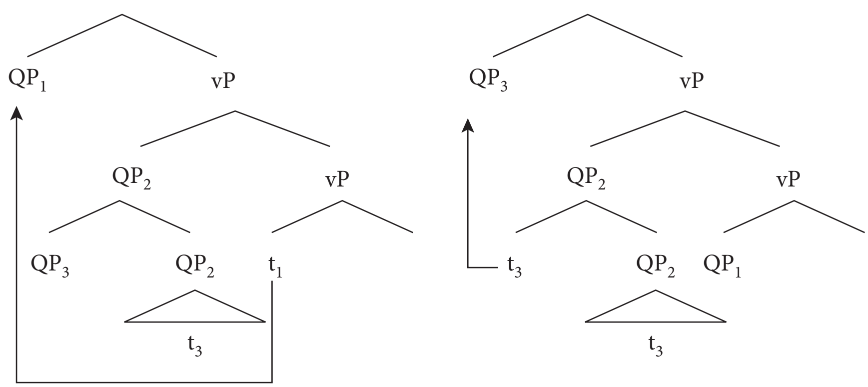


The decision which of these two nodes moves first proves interesting, as it supplies a diagnostic for choosing between the two competing strategies that Attract models and the current inception of TSP employ for ordering multiple dislocations, respectively: closeness vs. leap-frogging. Assume to begin with that precedence of operations were determined on the basis of 'closeness to the target', as is standard practice in Attract based models. Applying this metric to (26), one arrives at an impasse, because QP_3 and QP_1 (the subject) are equidistant to vP^2. Both nodes are separated by exactly one intervening node from the target (QP_3 by a segment of QP_2, labeled $QP_2{}^2$; and QP_1 by the vP-segment vP^1), and in both cases, the intervening node is a segment of a multiple-segment category. Movement of QP_3 and QP_1 should therefore be able to proceed in either order. However, as will be shown below (see also discussion of (29)), a system that admits free movement cannot derive scope freezing by any standardly sanctioned conditions.

If on the other side the current conception is correct, on which the order of movement operations is taken to be regulated by leap-frogging, movement of QP_2 above the subject forces the subject to cross over QP_2, as shown by (27a). Moreover, leap-frogging also excludes the alternative derivation (27b) in which movement of QP_2 is immediately followed by subextraction of QP_3, instead of movement by QP_1:

(27)

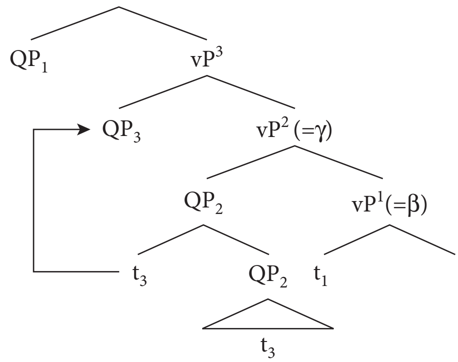
a. Leap frogging b. Non-leap frogging



Adhering to the idea that multiple dislocations are always resolved by leap-frogging, I will assume that (27a) is the correct intermediate representation created by the TSP model.

In the final movement step leading to an interpretable output, the inversely linked QP_3 is expelled from its intermediate position. QP_3 targets the next higher position in the tree, adjoining to vP^2, as detailed by (28). Note on the side that for QP_3, the node that its mother has combined with is vP^1. Hence, QP_3 adjoins to the result of merging QP_2 with vP^1 – i.e., vP^2 – and not to some position above QP_1.

(28)



In (28), all quantifiers have now reached a position in which they are interpretable (the subject will move on to SpecTP, but this is irrelevant for present concerns). What is of particular significance is the fact that the subject has not had the option to strand an intermediate trace inbetween QP_3 and QP_2. As a result, the subject can be read with widest possible scope, yielding reading (25a), or narrow scope, as in (25b), but cannot be construed with intermediate scope, as in (25c).

Thus, TSP provides a simple explanation for scope freezing in inversely linked contexts. The key ingredient of the analysis is again order of movement. Although feature driven subject movement and QR both proceed at the same level, that is in overt syntax, subject displacement precedes inverse linking on the assumption that derivations follow leap-frogging.

3.2.2 *Order of movement*

One open question remains to be addressed. What is the defining property of leap-frogging? Can it be derived from another axiom of the system? One venue to conceptualize the underlying principle governing these effects, which made reference to order of merge, was already explored in the discussion of double object constructions. There, it was concluded that categories which are merged last move first. For instance, in the double object case (20), the higher IO, which was added to the derivation later than the lower DO, underwent movement first. And in (27), the subject is merged at a later point in the derivation than the inversely linked object QP_3 and therefore raises before QP_3. On this view, the metric determining movement orders recycles an independent property of movement, i.e., the order of external merge.

I will end this section with a brief discussion of a related issue that arises in the analysis of inverse linking. As mentioned above, the order of merge informs order of movement. Interestingly, this metric only uses information about external merge operations in (27), while being ignorant about applications of internal merge or remerge. If the order in which categories are internally merged would also count, then derivation (27b) could no longer be excluded. More specifically, suppose that adjunction of the inversely linked QP_3 to its container QP_2 is delayed to a point *after* QP_2 has moved to vP, as detailed by the transition from (29a) to (29b). Then, the node that has been merged last would be QP_3, and QP_3 should therefore move prior to the subject QP_1, resulting in the intermediate representation (29c):

(29)
a. $[_{vP}\ [_{QP2} \ldots QP_3 \ldots]\ [_{vP}\ QP_1$ (result of movement of QP_2)
b. $[_{vP}\ [_{QP2}\ QP_3\ [_{QP2} \ldots t_3 \ldots]]\ [_{vP}\ QP_1$ (adjoin QP_3 to QP_2)
c. $[_{vP}\ QP_3\ [_{vP}\ [_{QP2}\ t_3\ [_{QP2} \ldots t_3 \ldots]]\ [_{vP}\ QP_1$ (move QP_3 to vP)
d. $[_{vP}\ QP_3\ [_{vP}\ QP_1\ [_{vP}\ [_{QP2}\ t_3\ [_{QP2} \ldots t_3 \ldots]]\ [_{vP}\ t_1$
(move QP_1 to vP, resulting in $QP_3 \succ QP_1 \succ QP_2$)

But in such a scenario, the TSP forces the subject to stop inbetween QP_2 and QP_3, as in (29d), feeding the unattested scope order $QP_3 \succ QP_1 \succ QP_2$. Thus, internal merge operations such as movement of QP_3 in (29b) must be somehow rendered invisible for the calculation of movement orders in (27b).

As far as I can see, there are two ways to exclude (27b) and similar derivation. On the one hand, it is possible to explicitly exclude prior movement operations from determining the order of future derivational steps. Although not implausible, this approach only yields suboptimal results. First, it fails for the double object cases, where movement of DO across IO forces the latter to raise. Second, it introduces an imbalance between internal and external merge, in that movement orders would only be dependent on properties of the latter. Finally, this view presupposes the availability of non-local information flow, because the nodes have to 'remember' throughout the derivation the order in which they originated.

Alternatively, one might invoke an independent property for blocking the derivation (27) above, while at the same time keeping the assumption that the order of movement is a function of the history of external as well as internal merge operations. A likely candidate for such a restriction is provided by an earliness requirement of the sort employed by Pesetsky (1989) and Stroik (2009). If operations need to be carried out at the earliest possible point in the derivation, QP_3 has to adjoin to its container QP_2 even before the container has reached the vP-adjoined position in (27). It follows that the category which is merged last is not QP_2, but the subject, and the derivation therefore proceeds, as desired, along the lines of (27a). In contrast to the first option, it is not necessary to have the relation only being defined for instances of external merge. Also, this conception, which will be adopted here, has the added advantage that introducing non-local dependencies can be avoided.

In sum, linking the derivational history of categories to the timing of future derivational steps represents a promising method for ordering movements. Taken together, the system employs two distinct timing metrics. On the one hand, if two (or more) nodes are to undergo TSP induced movement at a specific step in the derivation, then the category that has been merged last will be expelled first. On the other hand, competition between movement processes that cannot be ordered by this metric, e.g., because these processes are part of different subderivations (see inverse linking), is indirectly resolved by the requirement that operations apply as early as possible. At this point, it remains to be seen whether either one of these conditions can be reduced to the other, or whether they can be derived from a common underlying property. In addition, some obvious questions need to be answered. Most importantly, one would like to find a plausible reason why order of merge and order of movement should be linked in such an intimate way, and how this relation can be formally implemented. I will have to delegate these issues to future investigations.

4. Conclusion

In this paper, I tried to identify ways to discriminate between the predictions of an Attract based and a Survive based model of non-visible movement operations (more concretely, QR). The discussion lead to four broader conclusions.

To begin with, rephrasing the TSP in such a way that it generates intermediate position on top of every node a moving category has to cross leads to a definition that is less stipulative, and at the same time naturally subsumes the (methodologically and substantially) closely related principle of Type Driven Interpretation.

Second, the 'push-chain'-TSP based perspective on movement receives support from the fact that two conditions limiting scope taking directly follow from the restrictions that TSP imposes on the way in which feature driven subject raising interacts with QR.

Third, the very existence of such interactions between QR and subject movement can be taken as an indication that both dislocation processes apply at the same level of representation, so as to permit the derivational interactions discussed in Section 3 above. This finding supplies further evidence vindicating the phonological theory of QR.

Finally, only a push chain model, in which moving categories are expelled instead of attracted, leads one to expect that movement dependencies which involve such fundamentally different elements as subjects (bearing feature-incompatible subject features) and quantifiers (which are type-incompatible) systematically interact with each other. The standard Attract theory has to express this conspiracy indirectly, by positing features on quantifiers and a special proviso for the interpretation of subjects. On the Survive model, the interdependencies between QR and subject movement fall out from a basic property of the theory, i.e., the assumption that all feature or type incompatible categories are removed from their local environment, irrespective of the nature of the incompatibility.

If the above conclusions are essentially accurate, contexts involving different types of movements elicit important empirical diagnostics for evaluating the respective strengths of the two competing models, rendering such phenomena interesting candidates for future explorations into the Survive program.

References

Barss, A. & Lasnik, H. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17: 347–354.

Beck, S. & Johnson, K. 2004. Double objects again. *Linguistic Inquiry* 35(1): 97–124.

Bhatt, R. & Pancheva, R. 2004. Late merger of degree clauses. *Linguistic Inquiry* 35(1): 1–45.

Bobaljik, J. 1995. Morphosyntax: The syntax of verbal inflection. Ph.D. dissertation, MIT.

Bruening, B. 2001. QR obeys superiority: Frozen scope and ACD. *Linguistic Inquiry* 32(2): 233–273.

Chomsky, N. 1995. *The minimalist program* [Current Studies in Linguistics]. Cambridge MA: The MIT Press.

Chomsky, N. 2005. On phases. Ms, MIT.

Fox, D. 2002. Antecedent contained deletion and the copy theory of movement. *Linguistic Inquiry* 33(1): 63–96.

Fox, D. & Nissenbaum, J. 1999. Extraposition and scope: A case for overt QR. In *WCCFL 18*.

Heim, I. & Kratzer, A. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Hulsey, S. & Sauerland, U. 2003. Sorting out relative clauses: A reply to Bhatt. Ms, University of Massachusetts at Amherst.

Johnson, K. & Tomioka, S. 1997. Lowering and mid-size clauses. In *Proceedings of the Tübingen Workshop on Reconstruction*, G. Katz, Shin-sook Kim & Winhart, H (Eds), 185–206. Tübingen.

Kayne, R. 1984. Unambiguous paths. In *Connectedness and binary branching*, 129–164. Dordrecht: Foris.

Larson, R. 1987. Quantifying into NP. Ms, MIT.

Lechner, W. 2005/07. Interpretive effects of head movement. Available for download at http://ling.auf.net/lingbuzz.

Nissenbaum, J. 2000. Investigations of covert phrase movement. Ph.D. dissertation, MIT.

Pesetsky, D. 1989. Language-particular processes and the earliness principle. Ms, MIT.

Pesetsky, D. 1998. *Phrasal movement and its kin*. Cambridge MA: The MIT Press.

Preminger, O. 2008. (Im)perfect domains: Yet another theory of syntactic movement. In *Proceedings of 26th West Coast Conference on Formal Linguistics,* C.B. Chang & H.J. Haynie (Eds), 402–410. Somerville MA: Cascadilla Press.

Putnam, M.T. 2007. *Scrambling and the Survive principle.* Amsterdam: John Benjamins.

Richards, Norvin W. III. 2001a. *Movement in language.* Oxford: OUP.

Richards, Norvin. 2001b. Featural cyclicity and the ordering of multiple specifiers. In *Working minimalism*, S. Epstein & N. Hornstein (Eds), 127–159. Cambridge MA: The MIT Press.

van Riemsdijk, H. 1997. Push chains and drag chains: Complex predicate split in Dutch. In *Scrambling*, S. Tonoike (Ed.), 7–33. Tokyo: Kurosio Publishers.

Sauerland, U. 2000. Syntactic economy and quantifier raising. Ms, University of Tübingen.

Stroik, T. 1999. The Survive principle. *Linguistic Analysis* 282–309.

Stroik, T. 2009. *Locality in minimalist syntax*. Cambridge MA: The MIT Press.

Takahashi, S. 2006. Decomposition and identity. Ph.D. dissertation, MIT.

Taraldsen, K.T. 1981. The theoretical interpretation of a class of marked extractions. In *Theory of markedness in generative grammar: Proceedings of the 1979 GLOW conference*, A. Belletti, L. Brandi & L. Rizzi (Eds), 475–516. Pisa: Scuola Normale Superiore.

Language change and survive

Feature economy in the numeration

Elly van Gelderen
Arizona State University

This paper examines what data from language change can tell us about derivations, and in particular the early part of the derivation where lexical items are selected from the lexicon using the Survive mechanism. It is well-known that in changes often referred to as grammaticalization the features of lexical items are not only lost but reanalyzed from semantic to grammatical. I argue this is due to principles economizing derivations. Unlike many using Survive, I argue that uninterpretable features are in fact necessary.

There are many recent attempts to find a crash-proof grammar. Stroik and Putnam are among those and have, in a number of publications, argued for an active role of features in the operation merge. They argue that merge is a local operation driven by local feature checking. Features that are not checked 'survive' until a later application of merge. One of the motivating factors for this approach is to arrive at a crash-proof derivation. The features on their own are interface compatible and are there to ensure merge and remerge. In this paper, I examine what syntactic change can tell us in terms of features and how they are internalized by the language learner.

The paper is organized as follows. In section one, I outline what I see as the basic properties of the Survive Principle within a Minimalist Program. In section two, I review typical instances of language change, and in section three, I speculate on the role of Survive.

1. Survive

The Survive Principle is developed in Stroik (1999, 2009), Stroik & Putnam (2005), and Putnam (2007). It is formulated within the Minimalist Program (Chomsky, e.g., 1995, 2007) and attempts to provide crash-proof derivations. The Survive Principle states that computations should derive interface-interpretable representations. Thus, Survive-Minimalism reduces the computational complexity of a derivation. I will first provide some background to the Minimalist program and then outline the Survive principle and other aspects connected with this view of syntax.

A generative approach starts from the premise that our ability to acquire a language is a biological one, due to a genetic endowment, also referred to as Universal Grammar (UG). Being exposed to linguistic data is essential in building up an internal grammar that then generates language. In this approach, the emphasis is on UG. A Minimalist approach, as in Chomsky (1995; 2005; 2007), emphasizes a third factor, namely general cognitive principles. The three factors are shown in Figure 1, adapted from Chomsky (2005: 6).

1. Genetic endowment
2. Experience
3. Principles not specific to language

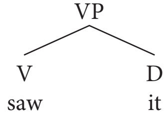
Figure 1. Factors in Language design

The third factor is divided into several types, including principles of efficient computation, which are "of particular significance in determining the nature of attainable languages" (Chomsky 2005: 6). Another change from earlier versions of generative grammar is the emphasis placed on language-external systems. This is formulated as (1):

(1) Strong Minimalist Thesis
Language is a perfect solution to interface conditions (Chomsky 2007: 5)

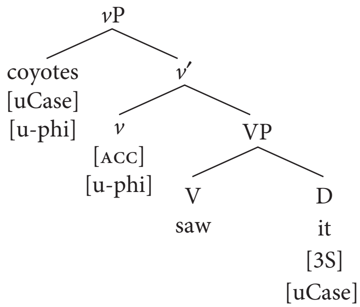
Thus, syntax has to satisfy the requirements of the external systems, the sensory-motor one and the conceptual-intentional one, achieved via the interfaces PHON and SEM respectively.

I'll now turn to an actual instance of a derivation paying special attention to merge and agree. In a late Minimalist derivation, elements are selected from the lexicon into a Numeration and then merged. Chomsky (2007: 6) suggests that the lexicon has "atomic elements, lexical items LI, each a structured array of properties". Abstracting away from features, a lexical array could be {saw, it, T, coyotes}. Secondly, elements are merged (through External Merge), e.g., *saw* and *it* in (2), and one of the two heads projects, in this case V, to a higher VP:

(2)


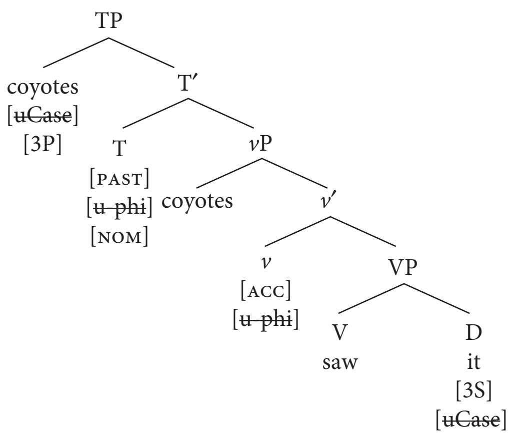
Merge is essential and is the core of the derivational system. Through merge, binary and hierarchical relationships between the merged elements form. We refer to the merged constituents as heads, complements, and specifiers, and to the relations between them via c-command. To the VP in (2), a (small) *v* and subject *coyotes* are merged, as in (3), where all the features are also indicated:

(3)



When *v* is merged, it probes (i.e., searches) into the VP to value its phi-features and at the same time it values the Case of the object. Then, after the functional categories such as T (and C) are merged to *v*P, these heads probe to find a noun with matching features to check agreement and with uninterpretable, unvalued Case (indicated as uCase). So, T has interpretable tense features but uninterpretable phi-features. It probes for a nominal it c-commands to agree with. It finds this nominal, or goal, in *coyotes* and each element values its uninterpretable features which then delete. The final structure will look like (4) where the features that are not 'struck through' are interpretable and not subject to elimination. The subject moves to Spec TP, or in other terms is merged from an internal position (Internal Merge) for language-specific reasons:

(4)



In (4), the CP layer is not indicated. The derivation in (4) uses early lexical insertion, i.e., a lexicalist approach, as in Chomsky (1995; 2004) but nothing hinges on this.

At some point (in steps/phases), the derivation has to be handed over to the Sensorimotor (SM) and Conceptual-Intensional (CI) systems. The former is responsible for linearization, e.g., which copy of *coyotes* is spelled-out in (4) (see Nunes 2004 for an account of how to decide which copies to spell out). For the CI system, two aspects are relevant, the theta-structure (determined in English through position but in other languages through inherent case) and the discourse information (see Chomsky 2002: 113). There are even two mechanisms responsible for the two, external and internal merge respectively.

Stroik (1999; 2009) points out some problems with this checking approach, in particular the possibility that a derivation crashes when, for instance, uninterpretable features survive to the CI system. In addition, he worries about the uneconomical nature of non-local probing. Instead, he wants to restrict "computational operations to a single type of local operation" (Stroik 2009: 2) and to reduce the power of internal merge, i.e., what used to be called 'move', by seeking to eliminate it. The main reason behind move is the EPP or OCC feature but this feature has nothing to do with interface requirements and is therefore problematic according to the SMT in (1). There are also empirical problems with non-local probing, e.g., the lack of a probe to license *a unicorn* in (5), as pointed out by Hazout (2004). Although there are straightforward solutions to this by having the existential *be* probe for partitive case, as in Belletti (1988), the point is well-taken that the EPP is a problem, however:

(5) It is unimaginable for there to be a unicorn in the garden.

Stroik's (2009: 22) alternative is that "displacement … comes from an operation that is local and repel-based. When a constituent X is having its features checked by a head Y, if there are features of X that are not compatible with Y … these features SURVIVE the checking operation and their incompatibility with Y repel them to the next c-commanding head Z". The XP that is repelled ends up in the specifier of a new head so that it has another chance at feature checking. Stroik's take on the selection of lexical items (LIs) from the lexicon is as follows. An LI enters the derivation by being copied into it from the Numeration, but it remains present in the Numeration. The narrow syntax has access to the entire Numeration and features that are checked in the derivation are also checked in the copy present in the Numeration. If all the uninterpretable features are checked, the LI becomes dormant (2009: 24–25). Stroik & Putnam's (2005) view of the lexicon is different from standard Minimalism as well. They assume that LIs and their features, i.e., the lexicon, are contained within the interfaces. Hence, no feature is uninterpretable and "[l]exical features are inherently legible at the interfaces" (Stroik 2009: 39). The narrow syntax need not check these features but is only there to check the legibility of larger constituents.

The data from language change show that having features play an important role in the numeration helps to explain certain phenomena. The data also show that what are known as in Minimalism as interpretable features are reanalyzed as uninterpretable. In what follows, I will therefore entertain a version of the Survive Principle where uninterpretable Edge Features are responsible for movement (see e.g., Richards 2008) so crashes do not occur. I won't assume the model of the lexicon that Stroik and Putnam advocate, i.e., one without uninterpretable features and one with continuous access of the Narrow Syntax to the lexicon. The latter is a burden on the computation as Chomsky has argued in relation to phases.

2. Syntactic change: Grammaticalization

Putnam (2007: 137–142) argues that the loss of scrambling in Modern English and Pennsylvania German can be seen in terms of the Survive principle. In this section, I restrict myself to linguistic change that is often referred to as grammaticalization and analyze it as feature loss. Describing grammaticalization in terms of features that are lost will enable us, in section three, to examine the relevance of Survive.

Grammaticalization is a process whereby lexical items generally lose phonological weight and semantic specificity and gain grammatical functions. The most well-known cases are those of verbs being reanalyzed as auxiliaries and prepositions as complementizers, i.e., of lexical elements changing to grammatical ones. There are also grammatical elements that are reanalyzed into more grammatical ones. These changes result in a need for renewal and this entire process is sometimes referred to as a linguistic cycle (van Gelderen 2007; 2008a). Grammaticalization was identified early on, but coined as a term in 1912 by Meillet. Works such as Lehmann (1985) and Heine & Traugott (1991) have inspired many linguists to pay closer attention to this phenomenon again, especially in a functionalist framework. Recently, however, structural accounts have started to appear (e.g., Roberts & Roussou 2003; van Gelderen 2004) accounting for the cyclicity of the changes involved. Van Gelderen, for instance, uses Economy Principles that help the learner acquire a grammar that is more economical, and as a side-effect more grammaticalized.

Cross-linguistically, motion verbs such as *go,* as in (6), are reanalyzed as future markers, as in (7). This means that, from the set of semantic features of [motion, future, location] in (6), only one is activated. If *go* occurs in the numeration with another verb, e.g., with *see* in (6), it is particularly easy to use only one of the features, in this case [future] rather than all. In that case, a biclausal structure is avoided as well. The difference between (7a) and (7b) lies in the degree of reanalysis. In (6a), the category of *going* is ambiguous between an auxiliary or a full verb, but this is not so in (7b):

(6) I am **going** (to Lahore) to see him.

(7) a. I am **going** to see him
b. I'm **gonna** see him.

Let's consider the numeration for (6). There are two lexical verbs and a main clause followed by an adverbial clause. This sentence could be split into two Numerations and later combined:

(8)	{I,	am,	going}	and	{C,	to, see,	him}
	i-phi	iT	i-ASP		i-FUT		….
	u-Case	NOM					
		u-phi					

The derivation for the main clause is as in (9a) and for the adverbial clause as in (9b). I have added a T as an alternative to (8) where I assume *am* fulfills the same role but nothing hinges on that:

(9)	a.	Merge	(am, going)	{am, going}
		Merge	(T {am, going})	{T {am, going}}
		Merge	(I {T {am, going}})	{I {T {am, going}}}
	b.	Merge	(see, him)	{see, him}
		Merge	(to {see, him})	{to {see, him}}
		Merge	(T {to {see, him}})	{T {to {see, him}}}
		Merge	(C {T {to {see, him}}})	{C {T {to {see, him}}}}

I am ignoring the role of v in licensing the object, the original position of the subject, as well as the theta-roles for now. Adding features and feature checking to (9a), we get the following incomplete list of features included in the first merge:

(10)	(am,	going)	{am,	going}
	i-T	i-ASP	i-T i-ASP	
	u-phi		u-phi	
	NOM		NOM	

Assuming Survive, copies of these LIs are kicked back into the Numeration, or remain alive, since not all the features have been checked (e.g., u-phi). Merging the subject to the derivation will enable the checking of the u-phi:

(11)	(I	{am,	going})	{I	{am,	going}}
	i-phi	i-T	i-ASP	i-phi	i-T	i-ASP
	u-Case	u-phi		~~u-Case~~	~~u-phi~~	
		NOM			NOM	

The Numerations in (8) could also be combined if one of the semantic features of *go* is (re)analyzed as a formal feature. Remember that LIs in the lexicon are nothing but features and can therefore probably easily change, i.e., be analyzed differently. The Numeration for (7a) would be as in (12) with the features that are relevant to the change in bold:

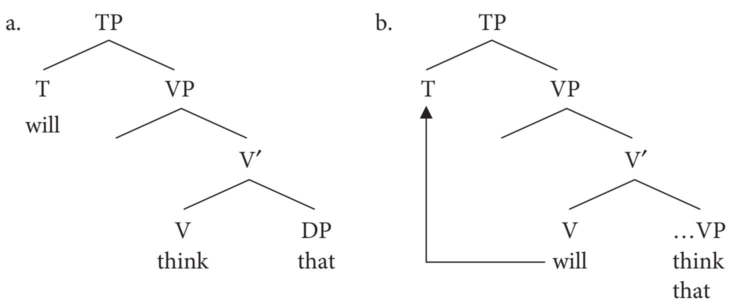
(12)	{I,	am,	going,	to,	see,	him}
	i-phi	i-T	**u-FUT**	**i-FUT**		
	u-Case	u-phi				
		NOM				

Going in (11) has semantic features of [motion away, future] but has lost those in (12). Thus, there is a decrease in semantic features, and the derivation will proceed faster, with fewer features surviving. Language change shows that surviving features may be reanalyzed by the language learner.

A similar example comes from the verb *willan* 'want'. It has semantic features such as [volition, future]. The selection from the lexicon for a sentence as in (13) might include {he, will, think, that}, using Modern English equivalents:

(13) He will think that.

Using tree structures, this could have (14b) as the result. One of the features of *will* could be reanalyzed as a grammatical tense feature with (14a) as the result:

(14)


Similar cases of reanalysis are given in (15) and (16), where the *have* and *to* in the (b)-sentences have fewer semantic features:

(15) a. I **have** a garden.
b. I **have** seen the garden.

(16) a. The highway **to** Tucson.
b. I didn't expect **to** think that.

Other examples of semantic feature loss are given in Table 1.

Table 1. Examples of Semantic Feature Loss

On: P to ASP	VP Adverbials to TP/CP Adverbials
Like: P to C (*like I said*)	Negative objects to negative markers
Modals: v to ASP to T	Negative verbs to auxiliaries
To: P to ASP to M to C	>PP to C (*for him to do that …*)

As in the case of *will* in (13), *have* in (15), and *to* in (16), each of the cases in Table 1 involves features loss. For instance, *on* changes from preposition to aspect marker in the Early Middle English period, as (17) shows, and PP-adverbials change from inside the VP to outside, as in (18). See van Gelderen (2004) for more details:

(17) *þær he was an slæting*
'There he was on hunting'. (Layamon, *Brut* Caligula 6139)

(18) a. *wen ic þæt ge* ***for wlenco*** *… Hroðgar sohton.*
expect I that you for daring … Hrothgar sought
'I expect that you sought Hrothgar because of your daring' (*Beowulf* 338–9)

b. ***forþam*** *Trumbriht wæs adon of þam biscopdome*
'because T had been deprived of his biscopric' (*Chronicle E,* anno 685).

Van Gelderen (2008b) suggests an Economy Principle, as in Figure 2, that expresses the changes that are observed.

Feature Economy
Minimize the semantic and interpretable features in the derivation

Figure 2. Feature Economy

In conclusion, syntactic change results in a system where LIs end up with fewer total features. Assuming a version of Survive, this may be more economical.

3. Conclusion: Survive and feature economy in language change

In the above section, I have shown that a number of syntactic changes involve the loss of features and the reanalysis of semantic features as formal features. The data show that too many surviving features may be a burden on the computation.

Abbreviations

ACC	Accusative
ASP	Aspect
BNC	British National Corpus
CHILDES	Child Language Data Exchange System
CP	Complementizer Phrase
EPP/OCC	Extended projection Principle/Occurrence; these trigger movement
FUT	Future
i-	interpretable
LI	Lexical Item
NOM	Nominative
OED	Oxford English Dictionary
phi	person, number, gender
S	singular
T	Tense
u-	uninterpretable

Notes

Thanks to Terje Lohndal, Tom Stroik, Michael Putnam, and an anonymous referee.

References

Belletti, A. 1988. The case of unaccusatives. *Linguistic Inquiry* 19: 1–34.
Chomsky, N. 1995. *The minimalist program*. Cambridge MA: The MIT Press.
Chomsky, N. 2002. *On nature and language*. Cambridge: CUP.
Chomsky, N. 2004. Beyond explanatory adequacy. In *Structures and Beyond*, A. Belletti (Ed.), 104–131. Oxford: OUP.
Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry* 36(1): 1–22.
Chomsky, N. 2007. Approaching UG from below. In *Interfaces + Recursion = Language*, U. Sauerland et al. (Eds), 1–29. Berlin: Mouton de Gruyter.
van Gelderen, E. 2004. *Grammaticalization as economy*. Amsterdam: John Benjamins.
van Gelderen, E. 2007. The definiteness cycle in Germanic. *Journal of Germanic Linguistics* 19(4): 275–305.
Gelderen, Elly van. 2008a. Linguistic cycles and Economy Principle: The role of Universal Grammar in language change. In Thórhallur Eythórsson (Ed.), *Grammatical Change and Linguistic Theory. The Rosendal Papers*, 245–264. Amsterdam: John Benjamins.
van Gelderen, E. 2008b. Where did Late Merge go? Grammaticalization as feature economy. *Studia Linguistica* 62.3: 287–300.
Hazout, I. 2004. Long-distance agreement and the syntax of for-to infinitives. *Linguistic Inquiry* 35(2): 338–343.

Heine, B. & Traugott, E.C. 1991. *Approaches to grammaticalization*. Amsterdam: John Benjamins.
Klaeber, F. 1922 [1941]. *Beowulf*. Boston MA: Heath & Co.
Lehmann, C. 1985. Grammaticalization: Synchronic variation and diachronic change. *Lingua e stile* 20(3): 303–318.
Nunes, J. 2004. *Linearization of chains and sideward movement.* Cambridge MA: The MIT Press.
Putnam, M. 2007. *Scrambling and the Survive principle*. Amsterdam: John Benjamins.
Richards, M. 2008. Defective agree, case alternations, and the prominecne of person. Ms.
Roberts, I. & Roussou, A. 2003. *Syntactic change*. Cambridge: CUP.
Stroik, T. 1999. The survive principle. *Linguistic Analysis* 29: 278–303.
Stroik, T. 2009. *Locality in minimalist syntax*. Cambridge, MA: The MIT Press.
Stroik, T. & Putnam, M. 2005. The lexicon at the interfaces. LASSO talk.
Thorpe, B. 1861. *Anglo-Saxon chronicle I and II*. London: Longman.

Towards a derivational syntax index

Linguistik Aktuell/Linguistics Today

A complete list of titles in this series can be found on the publishers' website, *www.benjamins.com*

147 **GOMESHI, Jila, Ileana PAUL and Martina WILTSCHKO (eds.):** Determiners. Universals and variation. *Expected October 2009*

146 **GELDEREN, Elly van (ed.):** Cyclical Change. viii, 319 pp. + index. *Expected July 2009*

145 **WESTERGAARD, Marit:** The Acquisition of Word Order. Micro-cues, information structure, and economy. xii, 242 pp. + index. *Expected July 2009*

144 **PUTNAM, Michael T. (ed.):** Towards a Derivational Syntax. Survive-minimalism. 2009. x, 269 pp.

143 **ROTHMAYR, Antonia:** The Structure of Stative Verbs. 2009. xv, 216 pp.

142 **NUNES, Jairo (ed.):** Minimalist Essays on Brazilian Portuguese Syntax. 2009. vi, 243 pp.

141 **ALEXIADOU, Artemis, Jorge HANKAMER, Thomas McFADDEN, Justin NUGER and Florian SCHÄFER (eds.):** Advances in Comparative Germanic Syntax. 2009. xv, 395 pp.

140 **ROEHRS, Dorian:** Demonstratives and Definite Articles as Nominal Auxiliaries. 2009. xii, 196 pp.

139 **HICKS, Glyn:** The Derivation of Anaphoric Relations. 2009. xii, 309 pp.

138 **SIDDIQI, Daniel:** Syntax within the Word. Economy, allomorphy, and argument selection in Distributed Morphology. 2009. xii, 138 pp.

137 **PFAU, Roland:** Grammar as Processor. A Distributed Morphology account of spontaneous speech errors. 2009. xiii, 372 pp.

136 **KANDYBOWICZ, Jason:** The Grammar of Repetition. Nupe grammar at the syntax–phonology interface. 2008. xiii, 168 pp.

135 **LEWIS, William D., Simin KARIMI, Heidi HARLEY and Scott O. FARRAR (eds.):** Time and Again. Theoretical perspectives on formal linguistics. In honor of D. Terence Langendoen. 2009. xiv, 265 pp.

134 **ARMON-LOTEM, Sharon, Gabi DANON and Susan ROTHSTEIN (eds.):** Current Issues in Generative Hebrew Linguistics. 2008. vii, 393 pp.

133 **MACDONALD, Jonathan E.:** The Syntactic Nature of Inner Aspect. A minimalist perspective. 2008. xv, 241 pp.

132 **BIBERAUER, Theresa (ed.):** The Limits of Syntactic Variation. 2008. vii, 521 pp.

131 **DE CAT, Cécile and Katherine DEMUTH (eds.):** The Bantu–Romance Connection. A comparative investigation of verbal agreement, DPs, and information structure. 2008. xix, 355 pp.

130 **KALLULLI, Dalina and Liliane TASMOWSKI (eds.):** Clitic Doubling in the Balkan Languages. 2008. ix, 442 pp.

129 **STURGEON, Anne:** The Left Periphery. The interaction of syntax, pragmatics and prosody in Czech. 2008. xi, 143 pp.

128 **TALEGHANI, Azita H.:** Modality, Aspect and Negation in Persian. 2008. ix, 183 pp.

127 **DURRLEMAN-TAME, Stephanie:** The Syntax of Jamaican Creole. A cartographic perspective. 2008. xii, 190 pp.

126 **SCHÄFER, Florian:** The Syntax of (Anti-)Causatives. External arguments in change-of-state contexts. 2008. xi, 324 pp.

125 **ROTHSTEIN, Björn:** The Perfect Time Span. On the present perfect in German, Swedish and English. 2008. xi, 171 pp.

124 **IHSANE, Tabea:** The Layered DP. Form and meaning of French indefinites. 2008. ix, 260 pp.

123 **STOYANOVA, Marina:** Unique Focus. Languages without multiple wh-questions. 2008. xi, 184 pp.

122 **OOSTERHOF, Albert:** The Semantics of Generics in Dutch and Related Languages. 2008. xviii, 286 pp.

121 **TUNGSETH, Mai Ellin:** Verbal Prepositions and Argument Structure. Path, place and possession in Norwegian. 2008. ix, 187 pp.

120 **ASBURY, Anna, Jakub DOTLAČIL, Berit GEHRKE and Rick NOUWEN (eds.):** Syntax and Semantics of Spatial P. 2008. vi, 416 pp.

119 **FORTUNY, Jordi:** The Emergence of Order in Syntax. 2008. viii, 211 pp.

118 **JÄGER, Agnes:** History of German Negation. 2008. ix, 350 pp.

117 **HAUGEN, Jason D.:** Morphology at the Interfaces. Reduplication and Noun Incorporation in Uto-Aztecan. 2008. xv, 257 pp.

116 **ENDO, Yoshio:** Locality and Information Structure. A cartographic approach to Japanese. 2007. x, 235 pp.