MINIMIZE EXPONENCE:
ECONOMY EFFECTS ON A MODEL OF THE MORPHOSYNTACTIC COMPONENT OF THE GRAMMAR

by

Daniel Siddiqi

A Dissertation to the Faculty of the

DEPARTMENT OF LINGUISTICS

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2006

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Daniel Siddiqi
entitled Minimize Exponence: Economy Effects on a Model of the Morphosyntactic Component of the Grammar
and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

____ Date: 5/10/06
Andrew Carnie, Committee Chair

____ Date: 5/10/06
Michael Hammond

____ Date: 5/10/06
Heidi Harley

____ Date: 5/10/06
Simin Karimi

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.
I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

____ Date: 5/10/06
Dissertation Director: Andrew Carnie

STATEMENT BY AUTHOR

SIGNED: Daniel Siddiqi

ACKNOWLEDGEMENTS

This dissertation could not have been written without the support of a host of different people throughout my graduate years at the University of Arizona. First among this list is my wife, Julianna, who followed me out to the desert and endured four long years of seclusion from her family, constant submersion in a field she couldn't care less about, and hundred plus degree summers just so that I could fulfill my dreams. The last year and a half must have been especially hard for her given the sign I've worn across my forehead reading: WE'RE SORRY, THIS MARRIAGE HAS BEEN INTERUPTED SO WE CAN BRING YOU THIS LONG PERIOD OF DISSERTATING. Julie has been with me through all of the good times and the hard times of my graduate career. Everything herein is made both possible and worthwhile by her constant love and support.

What finally ended up on the page here, in particular, has come about through the efforts of several different groups of people (with significant overlap of course). The germination for the ideas presented here came in the form of a term paper I submitted for Heidi Harley's Linguistics 522. I want to thank Heidi for encouraging me to take this particular project on and turn it into my dissertation. Were it not for her input, I likely would have produced a much more boring dissertation (Old English inflection was my original idea...shudder). That paper was developed into my preliminaries paper and then was part of my orals examination. I want to thank all the members of those committees including Andy Barss, Andrew Carnie, Dick Demers, Heidi Harley, Simin Karimi, and Andy Wedel whose comments and feedback on earlier versions of these ideas made it possible for me to expand them to the work here. I also want to thank Adam Ussishkin, who, while not on any of these committees, taught the prelim course, Ling 697, and had his hands in the mix as well.

The people most responsible for this work are, of course, my dissertation committee: Andrew Carnie, Heidi Harley, Mike Hammond, and Simin Karimi. I could not have asked for more support, feedback, and inspiration than I received from these great people. Any student would be lucky to have a committee such as mine—I could be quoted as saying, "I have the perfect committee", and I meant it. If it weren't for the constant efforts of these four people, what ended up on the following pages would be largely unreadable and, honestly, embarrassing.

In particular, I want to single out Andrew Carnie, my advisor and chair of all the committees listed above. Andrew has been like a rock for me and has guided me in my evolution from a disenfranchised English major to a real linguist. This, and all work I do for the rest of my career in linguistics, is due in large part to Andrew's efforts.

I also want to thank all the people who have worked in the framework of Distributed Morphology, whose work in and development of DM made this dissertation possible. I want to acknowledge David Basilico, David Embick, Morris Halle, Heidi Harley (again), Alec Marantz, and Rolf Noyer.

It is my firm belief that the single most important part of a graduate education is the interaction and support of your contemporaries. There aren't words to describe my debt to my peers in the graduate program at Arizona, so I won't even try. Some people I want to thank in particular are my graduate student elders Cathy Hicks Kennard and Meghan O'Donnell who served as mentors throughout my career here. Also among the list of elders whose aid was invaluable to me are Jason Haugen, Erin O'Brien, Tania Granadillo, Jianyuan Wang, and Scott Jackson. There are two more groups of my colleagues that I feel I need to acknowledge, as they were crucial to my development here (again, there is overlap). First, I want to acknowledge my incoming class: Shannon Bischoff, Erin Good, Emily Kidder, Naomi Ogasawara, and Lynnika Butler. These are the fellows that were by my side each step of the way and I am indebted to them for their support. Second, I want to thank the Geek Group (the activities of which will be omitted here due to high school fears of derision) whose membership over the years has included Scott Jackson, Dave Medeiros, Peter Richtsmeier, Jaime Parchment, Emily Kidder, Paul Saito, Heather Jackson, Jeff Punske, Michael Anderson, Hannah Soreng, Erin Good, and Erin O'Brien. Paradoxically, the Geek Group was at once crucially important to my finishing and the biggest impediment. Plenty of other friends have come along in the four years I've been in Tucson and among them I want to thank Benjamin V. Tucker, Jordan Brewer, Polly O'Rourke, Bob Kennedy, Lika Chtareva, Stacy Oberly, Yosuke Sato, Azita Taleghani, and Alina Twist.

The department here at the UA has been spectacular. I want to thank them for all their efforts and for their financial support. Faculty I haven't mentioned yet but that I want to thank include Diana Archangeli, D. Terrence Langendoen, and Massimo Piattelli-Palmarini. I want to give a special thanks to the great staff of the department including Marian Wisely, Jennie Bradley, Jennifer Columbus, and Rosemary Emery.

I also want to thank my family for everything they have done for me and for tolerating my (now ten-year) absence from home. Thanks to my mother, Ann Jaline, and my brother Jamal Siddiqi. Also thanks to my new family, Kerry Motts, Cathy Hays, Don Hays, Brad Loomis, Jeremy Motts, and Max Loomis.

There are two people whose contributions to my life were before my coming to Tucson but are nonetheless critical. First, Elmaz Abinader, one of my advisors in a previous life who noticed I was miserable as an English major and encouraged me to pursue linguistics, a field I had never heard of.

Finally, I want to thank Jennifer Farley Scharff. My life and all my accomplishments are a result of her inspiration.

DEDICATION

To my wife, Julianna. I'm sorry I've been an absentee husband while working on this and that you had to join the "Spouses of Dissertators" support group. It's done now. This dissertation is as much yours as it is mine.

TABLE OF CONTENTS

TABLE OF CONTENTS—Continued

TABLE OF CONTENTS—Continued

LIST OF FIGURES

ABSTRACT

Working within the morphosyntactic framework of Distributed Morphology (DM, Halle and Marantz 1993, 1994) within the Minimalist Program (Chomsky 1995), this dissertation proposes a new economy constraint on the grammar, MINIMIZE EXPONENCE, which selects the derivation that realizes all its interpretable features with the fewest morphemes. The purpose of this proposal is to capture the conflicting needs of the grammar to be both maximally contrastive and maximally efficient.

I show that the constraint MINIMIZE EXPONENCE has a number of effects on analyses of morphosyntactic phenomena. I propose that, in order to satisfy MINIMIZE EXPONENCE, the roots in a derivation fuse with the functional heads projected above them, resulting in a simplex head that contains both a root and interpretable features. Following the tenets of DM, this head is now a target for the process of Vocabulary insertion. Since the target node contains both content and functional information, so too can Vocabulary Items (VIs) be specified for both types of information. This allows VIs such as *eat* and *ate* to compete with each other. This competition of forms linked to the same root allows for a new model of root allomorphy within the framework of DM. In this model of root allomorphy, following proposals by Pfau (2000), VIs that realize roots participate in competition in the same was as do VIs that realize abstract morphemes. Since root VIs are participating in competition and are specified for both content and formal features, the need for licensing through secondary exponence as proposed by Harley and Noyer (2000) is removed from the framework. Further, since *eat* and *ate* in this model are different VIs with different specifications that compete with each other for insertion, this model of root allomorphy also eliminates the need for readjustment rules as proposed by Halle and Marantz (1993, 1994) and elaborated on by Marantz (1997). This new model of root allomorphy allows for an account of the blocking of regular inflection in English nominal compounds (e.g. **rats-catcher*), which was problematic for theorists working with DM, given the tenets of the framework.

I also show that the fusion of roots and functional elements driven by MINIMIZE EXPONENCE allows for a new account of subcategorization. The model of subcategorization presented here falls out of the following facts: 1) arguments are introduced by functional heads; 2) those heads fuse with the root they are projected above, resulting in the node containing both the root and the features of the functional heads; 3) since the root now contains both the root and the formal features, the corresponding VI can be specified for both; 4) VIs that realize roots can also be specified for compatibility or incompatibility of the features of the functional heads that license argument structure. The result here is an underspecification model of subcategorization that predicts a number of behaviors of verbs with respect to their argument structure that it is difficult for a full specification model to account for. Those include polysemy (*I ran the ball to Mary*) and structural coercion (*I thought the book to Mary*).

CHAPTER 1: INTRODUCTION

The study of the language faculty must address a central conflict about the grammar: On one hand, we have the need to convey a message and the need for that message to be as clear as possible. On the other hand, we need our message to be as efficient as possible. These opposing forces, contrast and efficiency, are the driving force behind a host of phenomena we see in language. For example, a number of phonological processes such as place or voice assimilation make utterances easier to say while processes such as dissimilation make the contrasts more transparent.

This conflict is prevalent in much of the linguistic literature. For example, in the realm of phonology and prosodic morphology, Optimality Theory (Prince and Smolensky 1993) captures this conflict using two classes of constraints on the grammar: faithfulness constraints ensure that contrastive meanings of an utterance aren't destroyed by the need to be efficient while markedness constraints capture that need for efficiency. In such a model, a grammar with all faithfulness constraints ranked above markedness constraints would result in a language that is maximally contrastive but is a mouthful to use, while the opposite ranking would result in a language that is maximally efficient but has eliminated the vast majority of its contrasts—making it unable to actually convey any meaning.

In historical linguistics, these forces are seen in cases where a language changes to make itself more efficient or to make itself more contrastive. For example, the loss of a sound in a language means a more efficient (easier) system at the cost of losing a contrast (e.g. the loss of the ɛ sound in some dialects of English results in the loss of the contrast between the words *pin* and *pen*). Similarly, adding a contrast (e.g., some dialects have added a new form for the second person plural, *you guys, y'all,* or *youse*, to contrast against the second person singular *you*) makes a system more contrastive at a cost of efficiency.

In many ways, language is shaped by this central struggle between contrast and efficiency. However, in the realm of syntax (and by extension morphosyntax), this struggle is largely ignored by theoreticians. A grammar that was maximally contrastive would not have structural ambiguity (e.g. *I saw the man with the telescope*) yet, would result in pronouncing every single functional head and every single formal feature in a maximally contrastive manner (e.g. one feature = one morpheme). On the other hand, maximum efficiency would entail summing all the features of an utterance into one word (e.g. *Yup.*).

Since the Minimalist tradition is composed of models of language competence rather than performance, it's not at all surprising that these competing forces are marginalized. The study of syntax is largely a question of what the system is and isn't capable of, not how that system is put to use. However, this results in a certain loss of predictive power. Limiting the discussion to the maximal pronunciation of formal features, there are immediately two concerns relevant to a model of UG that need to be discussed. The first is the more obvious question: why don't languages pronounce more functional morphemes (thus realizing more formal features)? Interpretable features, especially in languages like English, are largely unpronounced, despite the fact that, according to the Minimalist tradition, they must be present in each derivation. Second, why do languages differ in the amount of the interpretable features that are realized by morphology and how they are realized?

Minimalist tradition has included a number of economy constraints on the grammar whose purpose is to select the most economical derivation measured in energy used to create a derivation (e.g. greed, procrastinate, shortest move, etc, see Chomsky 1989, Rizzi 1990, Adger 1994). The purpose of this dissertation is to propose a new economy constraint: one that selects the most economical derivation measured in energy used to *produce* it.

We can create an economy constraint that captures the competing forces on the grammar—in particular the balance necessary in pronouncing all the interpretable features of a given derivation in the most efficient way possible. I propose the following constraint:

(1.1) MINIMIZE EXPONENCE

> The most economical derivation will be the one that maximally realizes all the formal features of the derivation with the fewest morphemes.

The gist of this constraint is that the best utterance is the one that conveys the most amount of information with the least effort (measured in number of morphemes that have to be pronounced). In terms of the production of an utterance, this constraint captures the struggle between the need to be maximally contrastive and the need to be maximally efficient.

This dissertation focuses on exploring the role that this type of constraint would have on the model of the grammar by specifically looking at the effects of MINIMIZE EXPONENCE on analyses of familiar morphosyntactic phenomena.

This dissertation is couched within the morphosyntactic framework of Distributed Morphology (DM), proposed in the early 1990's by Halle and Marantz. Chapter 2 is intended to provide the reader with the requisite understanding of DM for following the major claims of this dissertation. Section 1 provides a survey of DM, specifically focusing on how it is different from Lexicalist Minimalism. Section 2 outlines those of the morphological processes available to DM that this dissertation employs.

Chapter 3 proposes an analysis of root allomorphy (e.g. *mouse/mice*) within the framework of DM that showcases the economy constraint MINIMIZE EXPONENCE. Chapter 3 also accomplishes two other things: First, it proposes revisions to how the framework of DM explains root allomorphy. Second, it provides an analysis of nominal compounds in English (e.g. *mousetrap*). Nominal compounds in English uniquely disallow certain kinds of inflectional morphology (e.g. **rats-catcher*). While this phenomenon has been treated in the Lexicalist tradition, a proper treatment of it in DM is problematic. Section 4 of Chapter 3 uses MINIMIZE EXPONENCE to solve this theory internal problem.

Chapter 4 extends the analysis of the effects of MINIMIZE EXPONENCE beyond the realm of morphology and into the realm of syntax. Chapter 4 provides an analysis of subcategorization within the tenets of DM and characterizes the effects that MINIMIZE EXPONENCE has on a model of subcategorization within DM.

Finally, Chapter 5 is a remaining issues chapter, discussing several topics that, due to the focus of this dissertation, are not large enough to warrant their own chapter but still deserve discussion. First, MINIMIZE EXPONENCE has an interesting potential effect on the syntax's interface with event semantics. Since the rest if this dissertation focuses mainly on English data, Section 2 of Chapter 5 details some of the typological predictions of the inclusion of MINIMIZE EXPONENCE in UG. Section 3 considers a model of verb classes in a MINIMIZE EXPONENCE model of the grammar. Section 4 provides a discussion on the nature of features specification and the Elsewhere Condition in light of MINIMIZE EXPONENCE and the feature blocking system proposed in chapters 3 and 4.

CHAPTER 2 DISTRIBUTED MORPHOLOGY

This chapter provides a brief overview of the framework of Distributed Morphology. Section 1 describes the structure of the grammar in DM as opposed to that Lexicalist Minimalism. Section 2 describes all the morphological operations at work in DM.

2.1. The Structure of the Distributed Morphology Grammar

Distributed Morphology (henceforth DM, Halle and Marantz 1993, 1994) provides a proposal concerning the structure of the grammar—in particular, the lexicon, the interface with syntax, and constraints on semantics—as an alternative to other current models of Universal Grammar (UG). In recent models of the grammar within the tradition of Government and Binding Theory (GB, Chomsky 1981), UG is structured in such a way that the lexical module precedes the syntax in the derivation and feeds the syntax words for manipulation. In particular, the GB approach is similar to many other models of the grammar (such the Lexical Functional Grammar (Bresnan 2001, Falk 2001) and the Head-driven Phrase Structure Grammar (Pollard and Sag 1994)) in assuming that the lexicon is a generative component of the grammar, independent from any generative syntactic component. The lexicon in GB, following Chomsky (1970), is formulated as a word-building component that is structured independently of the syntax. The Lexical Hypothesis (Chomsky 1970) assumes that the words that are fed to the syntax come out of the lexicon fully formed, regardless of whether or not they are multimorphemic or simplex.

Distributed Morphology is a framework within the Minimalist program which rejects the Lexicalist hypothesis and the notion of a generative lexicon. This position is discussed below in Section 1.1 of this chapter. In DM, there is only one generative component of the grammar, the syntax, whereas in Lexicalist Minimalism, there are two: the syntax and the lexicon. The three key differences between DM and Lexicalist Minimalism to be discussed here are late-insertion, morphosyntactic decomposition, and underspecification.

2.1.1 Late-insertion

In DM, unlike in GB and its Lexicalist derivatives, rather than manipulating fully formed words, the syntax only manipulates abstract formal features to generate syntactic structures. These morphosyntactic features (such as [plural] and [past]) are selected from a fixed list of abstract features (or feature bundles) rather than being selected from the output of a generative lexicon. The late insertion hypothesis (Halle and Marantz 1994) holds that the phonology which represents the morphological features manipulated by the syntax is provided at PF rather than being present throughout the derivation. At spellout, syntactic terminals in DM are entirely comprised of interpretable features (including roots). Only once all syntactic processes are finished with the structure is phonological content added.

This phonology is provided by a component of the grammar called the Vocabulary. The Vocabulary is a static list of items whose function in the grammar is to provide phonology to realize the interpretable features contained in the terminal nodes of a derivation so that that derivation can be pronounced. Individual items within this list are called Vocabulary Items (or VIs for short). These VIs represent the basic sound/meaning correspondences of a language. Thus, the Vocabulary is the inventory of signs available to the language, which can be used to represent the featural nature of the syntax.

For example, imagine that syntactic operations result in a tree containing a terminal node comprised of the features [present], [singular], and [3rd person] (Each terminal node must be spelled out by some VI or other, including this 3psgPres one). In English, those three features are realized by the affix *–s*. This means there is an item in the Vocabulary, *-s*, which is inserted into that node at spellout to realize those features with the overt phonology /z/ (2.1).

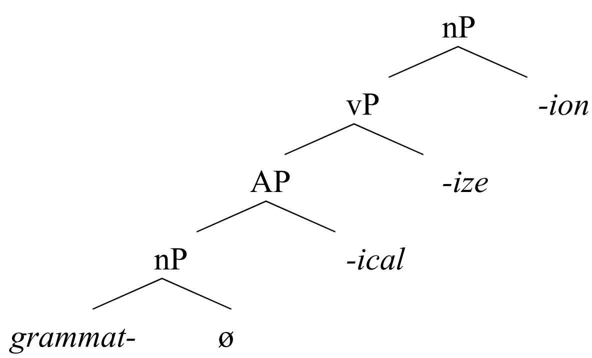
(2.1)	[present] [singular] [3rd person]	↔	*-s* /-z/

Since the phonology of any given derivation is added after spellout, Distributed Morphology is considered a late-insertion model. This crucial difference between DM and Lexicalism reduces to the point at which the phonological material is added to the derivation.

2.1.2 Morphosyntactic Decomposition

One of the strengths of the Distributed Morphology framework is the parallel between syntactic structure and morphological structure. Since the grammar of DM manipulates only syntactic features, the complex structure of a word is created in the same way as is the complex structure of a sentence. Spelling out a complex constituent of the syntax as a "phrase" or a "word" depends on the nature of VIs in the structure. For example, a complex NP such as *dogs of war* has obvious syntactic structure, but, DM claims, so does *wardogs.* To show a more complex example, consider the word *grammaticalization*. According to the tenets of DM, *grammaticalization* is a complex noun phrase (notated in DM as nP, since the dominating constituent is a projection of the functional nominalizer n—read 'little-n') composed of a nominalizing head and the verb phrase *grammaticalize.* That verb phrase is itself composed of a verbalizing head, *-ize*, and an adjective phrase *grammatical,* and so on (2.2).

(2.2) *grammaticalization*



The reason that the complex verb phase *grammaticalization* is spelled out as one complex word rather than a phrase containing isolated functional morphemes (such as *of*) is entirely a result of the available inventory of the English Vocabulary and the application of morphological processes (such as head movement and adjunction, see below). The VIs realizing the functional heads of A, v, and n happen to be affixes, which adjoin to their complements. The same structure in another language (such as an isolating language like Mandarin where the word *grammaticalization1`* is instead three isolated morphemes resulting in *yu fa hua*—literally "language" "law" "ization"[1]) could very well use free morphemes rather than bound morphemes to realize the same syntactic structure.

Since morphological structure derives from syntactic structure, lexical decomposition and phrase structure are identical—i.e. both are the result of the application of Merge in the syntactic component of the grammar. In the literature of DM, this transparency of morphosyntactic structure has been called "syntactic hierarchal structure all the way down" (Harley and Noyer 1999) or the "Pervasive Syntax Perspective" (Haugen 2004).

2.1.3 Underspecification

Distributed Morphology uses underspecification in the insertion of Vocabulary Items into a terminal node of the syntax. The insertion of a VI is governed by the subset principle (see below), which allows for a VI with certain specifications to be inserted into any node that satisfies those specifications, regardless of whether or not it exceeds those specifications. For example, consider the English coplula *are*, an example I will return to throughout this chapter. *Are* can appear in 1st person plural present tense, 2nd singular present, 2nd plural present, and 3rd plural present. The distribution of the VI, *are,* is attributable to the fact that its specification—just the feature [present]—is a subset of all four environments.

[1] Thank you to Jian Wang for this Mandarin example.

Many of the incarnations of the Lexicalist hypothesis, including Lexicalist Minimalism, HPSG, and LFG, are models that require full specification of all lexical entries. For example, in Lexicalist Minimalism, a word in the numeration is fully inflected and always contains the exact set of features that the syntax needs to check—otherwise the derivation crashes. This set is always the same and represents all and only the features that that word realizes.

Since DM uses underspecification, it makes very different predictions from Lexicalist models as to the behavior of syntactic elements (e.g. polysemy and syncretism). It also predicts a vastly smaller number of required lexical items since one item can fill many different roles.

2.1.4 Why reject Lexicalism?

The most cogent defense of Lexicalism is Chomsky's (1970) "Remarks on Nominalization." Chomsky argues that there must be a generative lexicon to account for the two different types of nominalizations that occur in English: gerunds and derived nominals.

Gerunds can be formed productively from a subject-predicate form whereas derived nominals can't always be (e.g. **John's growth of tomatoes* is not allowed though *John's growing tomatoes* is). Gerunds exhibit a regular relationship between the meaning of the verb and the meaning of the gerund whereas the same relationship between the derived nominal and its corresponding verb is not always as regular (e.g. *laughter, marriage, construction, actions, activities, revolution,* etc.). Finally, gerunds do not have the same structure as an NP (for example, the possessive cannot be replaced by a determiner such as *that, the,* or *no*).

Chomsky argues that there are two different possible ways for the grammar to account for this data: either we could suppose that there is a generative lexicon that creates these nominals before the syntactic component or we could build these components in the syntax—for example, by allowing the syntax to build verb phrases and then subjecting them to a transformation such as "Nominalization", thereby making the syntax more complex. Ultimately he decides that the Lexicalist position is correct for a number of reasons: a) derived nominals are not wholly productive and the syntax is, so the syntax must not be responsible for creating nominals; b) since the relationship of the meaning of the verb and the meaning of the derived nominal is often idiosyncratic, it is better to claim that this relationship is stored rather than created by the syntax; and c) since derived nominals behave like nouns not verbs (i.e. they cannot carry aspect) it is better to claim that they are not derived from VPs but are rather just nouns with verb-like selectional restrictions.

Marantz (1997b) approaches the conclusions Chomsky makes in “Remarks” (1970) through the eyes of contemporary grammatical theory, specifically the Minimalist Program (Chomsky 1995). Marantz uses nominals like *growth* and *laughter* to show that the lexical transformations of causitivization and nominalization actually occur in the syntax. For example—if *grow* is causativized in the Lexicon, then there is no way of ruling out Nominalization being applied to the causativized *grow* as well as the inchoative *grow*. If causitivization and nominalization are in the syntax instead of the lexicon, however, the ungrammaticality of *John's growth of tomatoes* arrises because:

a) √*grow* does not inherently take an agent argument
b) it can get an agent argument only through becoming a verb, via the agent-introducing causative head (v)
c) becoming a verb and becoming a (root-derived) noun are incompatible
d) ergo *growth* can't have an agent.

There need not be a relationship between the structure of a derived nominal and that of the corresponding verb. Marantz claims that the relationship between words like *receive* and *reception* is not that of a transformation from one to the other, but actually of a linking to a common root, √RECEIVE, which is unspecified for a category. Rather, it becomes a noun or a verb based on the environment it appears in. In a verbal environment, the root is realized as *receive*, but in a nominal environment the root is realized as *recept-* (see section 2.1.4 below for details). Marantz shows that the apparently idiosyncratic and somewhat contrary behavior of some nominalizations relative to the corresponding verbs is an effect of the combination of the functional projections above the root (for example, the variety of little-v projected above the root) and the class of that root (see Marantz 1997 for more details). In short, Marantz shows that Chomsky did not need to create a generative lexicon to capture the data. The syntactic mechanisms necessary to capture the differences between gerunds and derived nominals have already independently developed. With the syntactic mechanisms already in place, there is no need to propose a second grammatical engine for the composition of words.

Part of the Lexicalist Hypothesis is that the word has a special status in the grammar. In particular, the word is the locus for special sound processes, special structural process, and idiosyncratic non-compositional meaning. Since the word has all these special attributes, it follows that the language faculty contains a component dedicated to the construction and storage of just words. Marantz (1997b) refutes each of these "special" qualities attributed to the word.

The special status of the word derives from the fact that in phonology the prosodic word is one of the crucial units or levels of representation (Marantz 1997b). Lexicalism entails that the prosodic word correlates to the basic units needed for the syntax (i.e. lexical items). However, this claim is only unilateral. While syntacticians have assumed that the prosodic word and the "lexical item" overlap, phonologists make no such claim. Within the realm of prosodic phonology and morphology (cf. Kiparsky 1982, Prince & Smolensky 1993), in general there is no empirical evidence to show that syntactic structure is identical to prosodic structure. Marantz claims that zero-level units are often too small a unit for lexical phonology and at other times those same zero-level syntactic units contain complex prosodic structure. Marantz further argues that, even if the prosodic word and the zero-level unit did happen to be identical, there is no a priori reason for that to show that the phonological formation of the prosodic word must happen before the syntax unless LF was in some way sensitive to that level of the phonology.

Marantz also disputes that the lexicon is the location of idiosyncratic sound/meaning correspondences. Chomsky's assumption is that the words are the largest unit of structure that contain idiosyncratic meaning, whereas all structures larger than the word are constructed by the syntax and have wholly predictable meanings as a consequence. Marantz points to Jackendoff (1996) who reiterates the evidence that there is crucially no empirical difference between the idiosyncratic meaning of words and the idiosyncratic meanings of idiom phrases (such as *kick the bucket*). DM recognizes that there can be idiosyncratic meaning at any level[2], using the term idiom to refer not only to complex phrases and to complex words with non-compositional meanings, but also simple words such as *cat* (see below).

2.1.5 Construction of an Utterance

The structure of the Distributed Morphology grammar expands upon the familiar Y-model of Principles and Parameters, but differs in interesting and important ways (see Figure 1 in 2.3). The numeration in DM includes only morphosyntactic features. As discussed above, no phonological words are produced by any generative lexicon. Syntactic operations (i.e Merge and Move) operate on the abstract features to create semantically interpretable syntactic structures just as in Lexicalist Minimalism. These syntactic operations discharge uninterpretable features from the derivation.

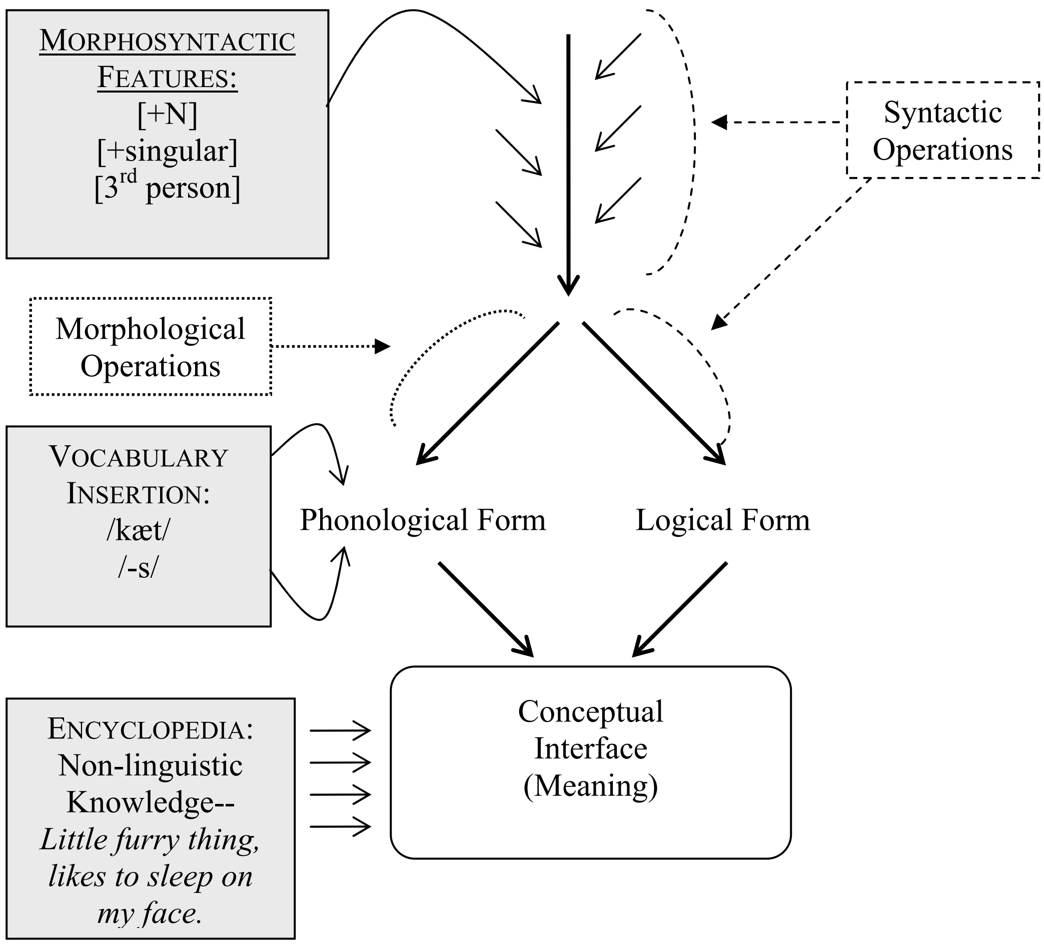
Spellout works a little differently in DM than in Lexicalist Minimalism. The two models are similar in that there is a division where the derivation proceeds along two different paths. Down one path, the derivation continues with some syntactic operations that eventually create the Logical Form of the utterance. The other path leads to

[2] The upward boundary of idiomatic meaning is often assumed by the literature on idioms to be somewhere between vP and CP. Ultimately, this distinction is not relevant to the presentation of DM given in this dissertation.

Phonological Form. However, along the path to PF, DM differs in proposing that there are some purely morphological operations (such as fusion, fission, morphological alignment, morphological merger, etc) which can alter the derivation in certain restricted ways such as (but not limited to): a) changing the featural content of a given terminal node; b) adjoining adjacent terminal nodes into a complex terminal node; or c) realigning or even adding morphemes (terminal nodes) to the structure, in accordance with the language's morphological rules. These operations do not affect the featural content of the interpreted path (LF). Thus, they are LF neutral.

At PF, Vocabulary Items are drawn from the Vocabulary and inserted into the terminal nodes of the structure produced by the derivation. The insertion of these VIs discharges all the interpretable features—i.e. it removes unpronounceable formal features from the derivation and replaces them with pronounceable "words". Following insertion, any applicable readjustment rules (see below) are implemented as well as are any phonological rules that may change the phonological form. Different from the traditional Lexicalist model, however, is the fact that the two arms of the Y-model join again at the conceptual interface. Here is where the features in LF and the semantics of the root VIs provided at spellout are interpreted by the Encyclopedia.

(2.3) Figure 1



The Structure of the DM Grammar

2.1.6 Spellout, Competition, and the Subset Principle

The primary operation through which "words" are constructed in DM is the syntax. Each terminal node in a given derivation is composed of one or more interpretable features. The terminal nodes now need phonology in order to be pronounceable. In DM, after spellout is when the "words" (i.e sound-meaning correspondences) enter the derivation—through Vocabulary Insertion.

The Vocabulary contains entries linking a formal feature (or features) to a series of sounds that realize that feature. These entries are called Vocabulary Items (VIs). For example, the familiar verbal inflectional morpheme, -s, in English realizes three features. Its VI might look like this (2.1 repeated here as 2.4):

(2.4)	[present] [singular] [3rd person]	↔	-*s* /-z/

Vocabulary Items are specified for the features they realize. For example, the Spanish determiners *los* and *el* are both specified for the formal feature [definite] and the phi feature [masculine]. However, *los* will also be specified for [plural] whereas *el* is not specified for number at all. The spelling out of functional morphemic material with these phonological items is done through a process of competition where different VIs compete with each other to be inserted into a syntactic node. In the Spanish example above, when *el* and *los* compete for insertion into a D° terminal node specified for plural, *los* will win because it is better specified for the features present in the node. Both are eligible for insertion, since neither contains any features which conflict with the fully specified D°, but in a competition for insertion, the VI which is specified for the largest number of features without being specified for any features that are not in the target node will win the competition. This is called the Subset Principle.

(2.5) **Subset Principle**:

> The phonological exponent of a Vocabulary item is inserted into a morpheme... if the item matches all or a subset of the grammatical features specified in the terminal [node]. Insertion does not take place if the Vocabulary item contains features not present in the morpheme. Where several Vocabulary items meet the conditions for insertion, the item

matching the greatest number of features specified in the terminal morpheme must be chosen (Halle 1997).

Competition for insertion in DM is governed by the following constraints:

a) The VI that has the most features matching the node features is inserted.

b) The node may contain more features than the VI is specified for, but the VI may not be more highly specified than the node. (The VI will still discharge all the features in the node, even if it is underspecified)

c) Since DM is an underspecification model, the "elsewhere condition" (or default) is the VI in the appropriate competition that is specified for the fewest features.

d) The Vocabulary Item may be sensitive to features in surrounding nodes, as well as to features on their loci of insertion.

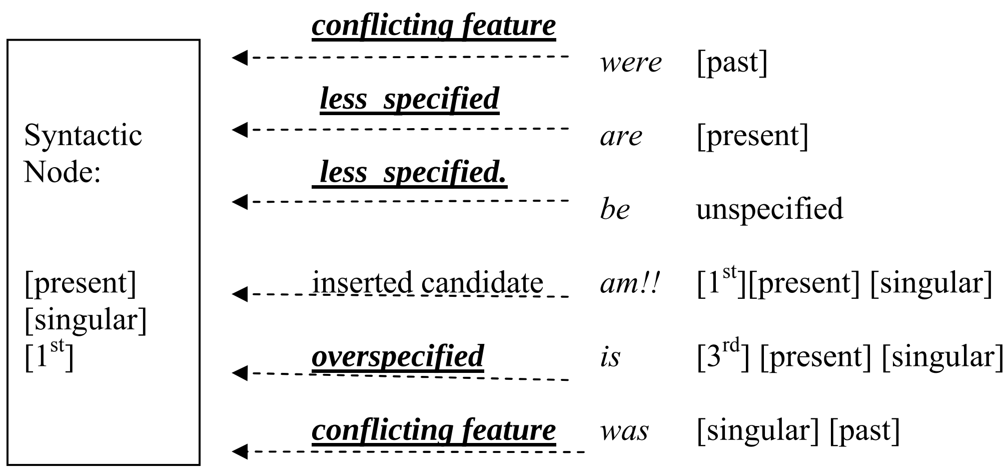
e) The operation of insertion is usually taken to be cyclic (Bobaljik 2002), allowing stems to be inserted before affixes. To use an English example, this allows affixes to 'see' the stems they will be attaching to, for instance, Latinate affixes may be sensitive to Latinate stems.

As a quick exercise in how insertion works, let's take the case of the English copula.

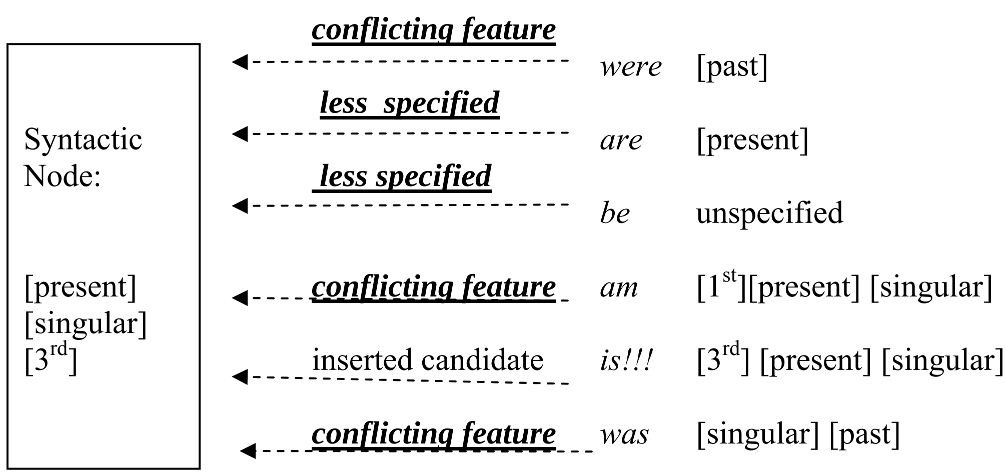
Let's assume that the different forms of the copula are specified as follows (simplifying the data for ease of presentation):

(2.6)	*were*	is specified for the feature	[past]
	are	is specified for the feature	[present]
	be	is the elsewhere condition	
	am	is specified for	[1st] [present] [singular]
	is	is specified for	[3rd] [present] [singular]
	was	is specified for	[singular] [past]

(2.7.a) Competition into *I am happy.*



(2.7.b) Competition into *He is happy.*



(2.7.b) Competition into *They are happy.*

Syntactic Node: [present] [plural] [3^{rd}]			
	conflicting feature	*were*	[past]
	inserted candidate	*are*	[present]
	less specified.	*be*	unspecified
	conflicting feature	*am!!*	[1^{st}][present] [singular]
	conflicting feature	*is*	[3^{rd}] [present] [singular]
	conflicting feature	*was*	[singular] [past]

In the schema of competition as described above, it is possible for two morphemes to arrive at a tie. The method for breaking ties proposed by Halle and Marantz (1993, 1994) is that competition was simply (but crucially) extrinsically ordered. The winning candidate was that the VI to come first in this crucial ordering. However, Noyer (1997) alternatively proposes that such ties are resolved by the Universal Hierarchy of Features: VIs that realize features higher on the hierarchy are preferred for insertion. For example, 1^{st} person and 2^{nd} person outrank 3^{rd} person in the hierarchy. Should a tie occur between a VI that was specified [3^{rd}] and one that was specified for [1^{st}], the VI specified for [1^{st}] would win because the contrastive feature in it is higher ranking. Similarly, since [1^{st}] outranks [2^{nd}], a VI specified for [1^{st}] would win in a tie with a VI specified for [2^{nd}].

Another possible result from competition is that a VI will be inserted into a node that contains more features than does the VI. Insertion of a VI can be understood as the replacement of the formal semantic features in a terminal node with something that is pronounceable (i.e. phonology). It follows that the formal features are removed from the derivation by insertion (so as not to be realized repeatedly). This removal of features is called feature discharge, or 'exponence'. An important aspect of insertion in DM is that the insertion of a VI into a node discharges *all* of the features contained in that node, not just the ones that the VI realizes. For example, in the competition given above in (2.7), the VI *are,* which is specified for the feature [present], is inserted into a node containing [present], [3rd], and [plural]. Even though *are* only realizes [present], it discharges all three of the features from node. In this way, every terminal node need only be spelled out by one VI.

A simple way of thinking of it is this: the phonology is presented with this complex syntactic object and instructed, 'pronounce this!' There are two simple conditions on this pronunciation: a) only terminal nodes must be pronounced, since they are the only linearizable parts and b) every terminal node must be pronounced. Providing a terminal node with phonological content—*any* phonological content—satisfies these two conditions. Therefore, it doesn't matter if a VI doesn't realize all the features of a node, as long as it provides it with phonological content. Competition will ensure that whatever phonology a node gets will be the best available representation of its content.

Noyer (1997) showed that there are times where the insertion of a VI is also dependent on features present in other nodes. To account for this type of data, Noyer (1997) shows that VIs may be specified both for things that they are the *primary exponents* of (i.e the features in the node into which they are inserted), as well as for things that they are *secondary exponents* of (i.e. features in nodes other than the node into which they are being inserted). This is called *secondary exponence.* For example, the allomorph of the A head /-t/ (which is not susceptible voice assimilation) is selected via secondary exponence in forms like *burnt* (cf. the tensed verb form *burned*)

Another good example of the usage of secondary exponence is the example of Spanish determiners seen above. The competition between *el* and *la* is not settled by the featural content of the node they are being inserted into (by hypothesis, the nodes are identical), rather it is settled because *el* is a secondary exponent of the feature [masc], present on the stem, while *la* is a secondary exponent of the feature [fem]. The gender feature is not in the determiner node and is not discharged by the insertion of a determiner VI, but rather is likely in the noun and is discharged by the noun. Similarly, the difference between *el* and *los* is that one is a secondary exponent of [sing] and the other of [plural], again not a feature that is in D°, but rather is located in Num°

One of the usages of secondary exponence is using it to license VIs that realize roots (proposed by Harley and Noyer 2000). For example, the VI *thrash* realizes a root but also can only be a verb (a verb is a root dominated by v, a noun is a root dominated by n). Harley and Noyer propose using secondary exponence to require that *thrash* only be inserted into a node dominated by the little-v head, even though the little-v head is realized by another VI. This use of secondary exponence for licensing root VIs is called licensing and is elaborated below and addressed in detail in Chapter 3.

2.1.7 The Distinction between Functional and Contentful

In DM, the definitional distinction between functional morphemes and content morphemes is the presence of a root. Roots are “abstract morphemes” linked to a basic concept (the root for *cat* is √CAT). They are formal elements of the grammar that, unlike other syntactic featutres, are linked to extragrammatical information (such as reference or encyclopedic knowledge). The major effect that this has on the grammar in DM is that the VIs that realize functional features participate in competition but those that realize roots don’t.

The original proposal for DM did not include different “varieties” of roots, such as √CAT or √DOG. Rather, there was only one syntactic element “√”, which was just called a root. This √ meant that the terminal node containing it needed to be realized by some sort of content morpheme. The syntax was indifferent to which content morpheme realized that √. Thus, the syntax component of the grammar was not sensitive to the difference between *cat* and *dog.* All that was relevant to the syntax was the formal syntactic features. This division of form from referential or encyclopedic meaning was originally considered one of the strengths of DM. However, Pfau (2000), in an account of speech error phenomena, proposed that the roots manipulated by the syntax were specific to the particular concept they were linked to (such as √DOG). This change has largely been adopted by the DM community, and I assume it in this dissertation[3]

[3] I have not attempted the analyses presented here in a model of DM with “general” roots (√) other than in the most rudimentary form and do not know how rejecting that assumption would affect the analyses presented here other than to say that certain parts, particularly chapter 3 would need to be reviewed

Since *cat* and *dog* were both specified for √, they could not compete with each other (they have identical specification). Thus, root VIs could not participate in competition. Since the addition of Pfau's specific roots, the groundwork has been laid for the competition of root VIs – each VI is linked to a different root.

For example, Harley and Noyer's *l-morpheme hypothesis*[4] (1999, 2000) recapitulates the difference between roots and functional morphemes. They claim that VIs which realize roots (since they do not participate in competition—they all have identical primary exponence) are licensed for insertion through secondary exponence (see above). Given that *dog* and *cat* could compete with each other in Pfau's model, the licensing proposed by Harley and Noyer reduces to specific secondary exponence requirements on particular VIs. However, as of yet, no one has proposed a model of the grammar where roots participate in competition.

Within the older (pre-Pfau) model of DM, Marantz (1997) points out that, since root VIs aren't in competition with each other, irregular forms of the same root, such as *rise* and *raise* or *eat* and *ate*, which look to be different forms of the same verb, also cannot be competing with each other even though they appear to be (being conditioned by different structural configurations). That means that the change from *rise* to *raise* (or from *eat* to *ate*) needs to be some type of syntactically-conditioned allomorphy. In pre-Pfau DM, there had to be two types of allomorphy that are conditioned by morphosyntax (excluding phonologically conditioned allomorphy, such as voicing assimilation in

[4] Harley and Noyer name roots "l-morphemes" and call fuctional morphemes "f-morphemes". Given the preponderance of different names for this crucial division, in this dissertation, I will try to maintain the names "functional morphemes" vs. "content morphemes/roots".

inflectional suffixes in English). The allomorphy that results from competition of different VIs (thus limited to functional morphemes) is suppletion. However, content morphemes have to be subject to some other kind of allomorphy. In this case, (pre-Pfau) DM proposes that there is one VI, which undergoes readjustment (discussed below in Section 2 of this chapter).

This division between the processes that condition the insertion of functional VIs (competition) and contentful VIs (licensing and readjustment) largely means that the pre-Pfau model of Distributed Morphology had two separate grammars, one dedicated to the insertion of functional morphemes and the other to content morphemes. As discussed above, Pfau's proposal of contentful roots allows a model of DM where root VIs compete with each other (in this model *cat* and *dog* would be specified for different roots, therefore their specifications aren't identical—allowing them to compete). Again as above, licensing then is just a matter of the secondary exponence requirements of the VI. As of yet, nobody has proposed how to incorporate Pfau's hypothesis with the conditioning of the insertion of root VIs. Such a proposal is one of the purposes of Chapter 3 of this dissertation.

2.1.8 Lexical Categories in DM

Traditional parts of speech receive a special treatment in DM, which recognizes that some roots may surface as multiple different lexical categories (for example, *book* as a noun or *book* as a verb in *I booked the appointment*; similarly *I saddled the horse* or *Verbing wierds language*). DM asserts that the familiar classifications for content morphemes—noun, verb, adjective, etc—are epiphenomenal. More precisely, the grammar does not store a content item labeled as a noun or a verb. Rather, the properties of lexical classes are derived from the syntactic complex of functional syntactic heads and a root. A root "becomes" a verb by being immediately c-commanded a verbalizing functional head, such as little-v, Aspect, Tense, and perhaps Trans. So a verb is not a simplex word in the traditional notion, but rather it is a complex of a functional head dominating a root. Similarly, nouns are roots c-commanded by a functional nominal head (n). According to Harley and Noyer's *L-morpheme Hypothesis* (as described above), the immediate c-commanding head often licenses the insertion of a contentful VI into the root position, if that VI is restricted in its categorical behavior, and thus is called the licenser.

2.1.9 Special meanings and the Encyclopedia

In Lexicalist syntax, one of the important functions of the Lexicon is to be the locus of special meanings. In particular, theoretical linguistics has always recognized that some complex structures need to be stored whole because the whole structure has an idiosyncratic meaning. Some familiar idioms are *kick the bucket*, *buy the farm*, *rain cats and dogs*. These meanings are not composed from the meanings of the individual constituents of the structure. Often the syntax itself does not function "as it should". For example, the idiom *a lot (of)* is a DP, but almost always functions as an adjective or an adverb. The idiom *the hell out of* is not a traditionally recognized constituent.

DM as I described here seems to have a problem with such structures since the syntax is built up before the words are inserted. In that case, there is no way for a complex idiomatic structure to be stored unless there was a mechanism in spellout that allowed one VI to discharge complex structure. Rather that take this route, DM recognizes that idioms may be simpler than originally assumed. The following is a list of structures that have meanings that are not predictable from composition.

(2.8) a) *let the cat out of the bag*
b) the whole nine yards
c) terrific
d) scissors
e) duck

The examples in a) and b) are widely accepted as idiomatic, because the meaning of the whole phrase is not compositional. However, from the viewpoint on which syntax and morphology are the same mechanism, c) and d) must also be idioms as their meanings are not a sum of their morphemes. *Terrific* and *scissors* just have to be memorized as not being compositional—*terrific* does not have the compositional meaning that *terror-ify-ic* would have (roughly equivalent to *horrific*) but rather means something analogous to *great* or *swell*; similarly, *scissors* refers to just one thing despite apparent plural morphology. It seems then that words by themselves can be idioms. The logical extension of this is to recognize that we already memorize the meaning of arbitrary collections of phonemes when we memorize monomorphemic words such as *duck*. There really is no difference between memorizing the meaning of a large chunk like a) and a small chunk like e). Through this view, we can define an idiom as any grammatical expression whose meaning is not predictable. The grammar then needs not to have a place to store large chunks that are idiosyncratic, because *everything* is idiosyncratic. Rather, the grammar has to have some way of interpreting idiosyncratic meaning. Crucially, this part of the grammar must come *after* insertion, as in the original DM model, the words *cat* and *bag* aren't present in a derivation until after spellout. It must also be sensitive to the final LF form as well since some idioms include structures sensitive to processes that occur between spellout and LF (such as QR. C.f. *Every dog has its day*).

In DM, the mechanism through which these interpretations are made is called the conceptual interface or, more often, the Encyclopedia. The Encyclopedia is the locus for idiosyncratic knowledge of roots (such as the fact that *kick* and *buy* have special meanings when their objects are *the bucket* and *the farm* respectively) and our real world knowledge of the referents of words (for example, the fact that a *cat* is a furry quadruped and *thinking* is something that occurs only in your head). As I will discuss below, this knowledge often clashes with otherwise well-formed grammatical structures (*Mary thought the book to John;* Colorless *green ideas sleep furiously*) Because of this, much of the information that is stored in the Encyclopedia is considered extra-linguistic. This raises the question as to whether or not the Encyclopedia is itself a grammatical component. I have assumed it to be in my sketch of the grammar, but this assumption is debatable.

2.2. Morphological Operations in DM: From Spellout to PF

In Section 1 of this chapter, I detailed the basics of the framework of DM. Since DM is both a theory of syntax and a theory of morphology, there are a number of operations in DM to account for morphological behavior. The majority of these operations occur at spellout or between spellout and PF. For the sake of simplicity, I will use the terminology "at spellout" as I did in the previous section, though many of these operations are crucially ordered.

As shown above, much of the morphological structure is a result of syntactic processes, but some things have to occur in a morphological component that occurs after the syntax proper. In this section, I will discuss some of those processes that are available to DM that are relevant to the work at hand: morphological merger[5], fusion / fission, and readjustment rules.

2.2.1 Morphological Merger

Marantz (1984) first proposed morphological merger as a way of realizing (or replacing) syntactic structure with morphological structure. Marantz (1984) proposes that specific relations between two syntactic nodes can be replaced by affixing those two nodes to each other. Later, Marantz (1988) defined merger to be:

(2.9)

> At any level of syntactic analysis…, a relation between X and Y may be replaced by (expressed by) the affixation of the lexical head of X to the lexical head of Y (Marantz 1988 261).

---

[5] The name for this process is frustratingly similar to the syntactic operation MERGE. They are, however, unrelated.

Fundamentally, merger is the process whereby an abstract morpheme realized by an affix is attached to a stem, creating the difference at linearization between two elements that form one compositional unit (a complex word) and elements that just happen to be next to each other (a difference that is often realized orthographically by a space). Also, merger has been argued by Bobalijk (1994 cal1ed "merger under adjacency") to be one process by which (zero-level) elements adjacent in the syntactic structure come to occupy the same syntactic node (for example, the attachment of the tense node to the verb head in English).

Another recent use of merger has been what has been historically called head movement. Chomsky (1994) has argued that head movement is not a function of syntax but rather something that occurs after spellout. Using morphological merger, head movement can be argued to the merger of two zero-level elements (Baker 1988).

Essentially, merger is the name of the process whereby two independent zero-level nodes are reorganized into one complex zero-level node be it by affixing to each other, re-linearizing, or conjoining.

2.2.2 Fission and Fusion

The process of *fusion* exists within DM to account for situations where merger is not quite adequate to account for the distribution of vocabulary items in a string. This is for cases where the formal features from several different terminal nodes are all realized by one Vocabulary Item. For example, it is common for phi-features and tense to be realized by the same morpheme (as is the case in English and Spanish. What the process of fusion does is reduce all the features of a several distinct terminal nodes (perhaps dominated by a single X°-level element, such as that created by the process of merger) to one simple head containing all the features of the complex head. For example, in English, the features [present], [3rd person], and [singular] are all realized by the affix -*s*. By hypothesis, the phi-features and the tense features are in separate syntactic nodes, AgrS° and T°. The process of fusion applies to the two nodes to combine all the features of the two former nodes into a new, single node that can then be targeted for insertion by –*s*.

Complementary to the process of fusion is ‘fission’, as proposed by Noyer (1997) and Marantz (1997). Whereas fusion takes two positions of exponence (i.e. terminal nodes) and reduces them to one node, fission takes one position of exponence and splits it into many. The most common usage of fission in the literature has been to separate phi-features from one another, since by hypothesis they all are base-generated in a single feature bundle, which appears as one terminal node at the end of the syntactic derivation. Both Marantz (1997) and Noyer (1997) employed this mechanism to spilt AGR into separate heads, each containing a phi feature that is realized independently. Thus, a language, such as Noyer (1997)’s example of Tamazight Berber, can have separate morphemes for person, gender, and number (see 2.10, taken from Harley and Noyer 1999)

(2.10.a) Tamazight Berber Prefix Conjugation. dawa 'cure'

	singular	**plural**
1	*dawa-γ*	*n-dawa*
2m	*t-dawa-d*	*t-dawa-m*
2f	*t-dawa-d*	*t-dawa-n-t*
3m	*i-dawa*	*dawa-n*
3f	*t-dawa*	*dawa-n-t*

(2.10.b) Vocabulary Items proposed by Noyer (1997)

/*n-*/	↔	[1^{st}] [plural]
/*-γ*/	↔	[1^{st}]
/*t-*/	↔	[2^{nd}]
/*t-*/	↔	[3^{rd}] [sing] [fem]
/*-m*/	↔	[plural] [masc] ([2^{nd}])
/*i-*/	↔	[sing] [masc]
/*-d*/	↔	[sing] ([2^{nd}])
/*-n*/	↔	[plur]
/*-t*/	↔	[fem]

2.2.3 Readjustment Rules

Finally, there are types of allomorphy that cannot be captured using the processes described above: Allomorphy whose conditioning environment is not phonological[6], but rather is morphological. In the case of functional morphemes, this morphological conditioning is captured through some combination of competition and secondary exponence. However, since DM assumes root morphemes are not in competition with each other, there is a process whereby the phonology of the VI is changed as a result of morphologically conditioning. For example, *-ceive* and *-cept* are both realizations of the same VI where the change from *-ceive* to *-cept* is conditioned by the nominalizer *–ion*.

[6] such as nasal place assimilation, vowel harmony, voicing assimilation, etc.

DM proposes that there is one basic allomorph (above *–cieve*) and all the other allomorphs are the result of the application of a phonological Readjustment Rule.

In the case of the morpheme, -*ceive* (*receive, deceive, conceive, perceive*), DM hypothesizes a readjustment rule that changes v to p, and shortens i to ε in just this vocabulary item, when the nominalizing suffix -*tion* is attached to it. Similarly, DM hypothesizes a readjusment rule that changes *foot* to *feet* when dominated by the feature [plural] (cf *goose/geese, mouse/mice*). Additionally, there has been some suggestion that readjustment rules can target classes of VIs instead of being idiosyncratic to just one (see Embick and Noyer 2004, Embick and Halle 2004 and others). For example, there could be a readjustment rule in English that changes /aj/ to /ow/ in a VI when it is c-commanded by [past] (*drive/drove, ride/rode,* and *dive/dove*). This readjustment rule would apply to an entire class of VIs instead of just one.

Readjustment rules are different from the syntactic and morphological transformations of older work (see Chomsky and Halle 1968, Kiparsky 1982 and many others). Readjustment rules are taken to be idiosyncratic rules that apply only an individual VI or class of VIs, they are not general phonological rules in the grammar.

CHAPTER 3 ON A THEORY OF ROOT ALLOMORPHY

Root allomorphy is a subset of relationships traditionally called irregular morphology. Root allomorphy comes in two varieties[7]. The first is suppletive allomorphy where the two forms cannot be derived from each other by some sort of phonological process. Some examples of suppletive allomorphy are in (3.1).

(3.1) *go/went*
good/better/best
bad/worse
person/people

The other type of allomorphy is what I call irregular allomorphy, in which there is some common phonology between the two forms. This commonality is usually attributable to some type of historically regular phenomena (such as i/j umlaut) which has since fallen out of the language.

(3.2) *eat/ate*
mouse/mice
receive/reception
sleep/slept

In DM, as discussed in chapter 2, since root allomorphy always involves a root, these relationships are not captured the same way that suppletion of functional morphemes is (i.e. competition). Rather, these relationships are always considered the application of a readjustment rule.

---

[7] I will use the terms suppletive and irregular as contrastive. In actuality, suppletion is usually considered a subset of irregular morphology. I am using the term *irregular* to mean irregular morphology minus suppletive morphology.

The ultimate purpose of this chapter to show an analysis of root allomorphy that uses the same mechanisms used for allomorphy of functional morphemes. This chapter will also aim to do four other things: 1) offer an explanation for data that is otherwise unexplained in DM (i.e. the blocking of inflection in the non-head position of nominal compounds); 2) show that DM does not need two different "grammars" for morphology (one set of operations for roots and one for functional morphemes) as described in Chapter 2, reducing the number of operations proposed by DM and make for a more economic model of the grammar; 3) greatly reduce the number of null morphemes that are predicted by DM—a prediction which is a potential criticism; and finally, 4) show a functional application of MINIMIZE EXPONENCE, the ultimate purpose of this dissertation.

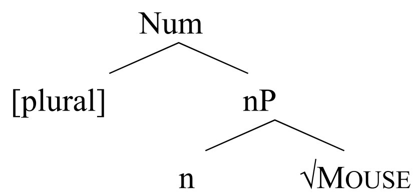
To those ends, I propose that, in order to satisfy MINIMIZE EXPONENCE, the functional heads projected above the root fuse with that root. This results in the root and the formal features being in the same node. Because of this fusion, VIs can be specified for both a root and formal features, allowing *eat* and *ate* to be different VIs that compete with each other for insertion.

3.0.1 Licensing and readjustment in DM

For the purposes of refreshing the reader's memory, in this section I quickly review some of the aspects of DM presented in Chapter 2. The purpose of this section is to show root allomorphy in action in DM and then to showcase some of the concerns about the traditional DM analysis that will be addressed in this chapter.

Consider the typical derivation of root allomorphy in DM. As an example, consider the derivation for *mice*, a form showing what I have called irregular root allomorphy. The syntax results in the complex head found in (3.3):

(3.3)



Were *mice* a regular form, the node containing the feature [plural] would be realized by the VI –*s*. Since the root, √MOUSE, is a noun, its VI would be licensed for insertion by the feature little-n, which is itself realized by a null morpheme. The root itself would be realized as *mouse*. Linearization of the morphemes would result in *mouses*. However, since *mice* is irregular, there are a number of key differences to this derivation. First, instead of the affix –*s*, the null allomorph of [plural] is conditioned by the presence of the root √Mouse. *Mouse* is inserted, again having been licensed by little-n (still realized by a null morpheme). Then, a readjustment rule is conditioned by *mouse* being c-commanded by [plural]. This readjustment rule changes the phonology of *mouse* to *mice.*

In addition to the marked complexity of the derivation of a relatively innocuous word like *mice*, a strange interdependence occurs in the derivation. The null plural morpheme is licensed by the presence of *mouse* and the readjustment of *mouse* to *mice* is licensed by the presence of [plural].[8] What follows will be an alternative analysis of root allomorphy that is ultimately less complex.

---

[8] This interdependence is the effect of both the application of readjustment rule and the secondarily licensed affix and is actually a prediction of the DM grammar. Since both readjustment of a root and a secondarily

3.1.0 Roots in the grammar

Recall from chapter 2, VIs realizing abstract morphemes are specified for formal features such as [past] or [1st]. Thus, an example of a typical specification of a VI is (3.4).

(3.4) Vocabulary Entry for *-ed*

[past] → *-ed*
/-d/

VIs must be specified for the morphosyntactic features that they realize. These features encode the meaning of the syntactic node in the semantic computation. The VIs that realize roots are also so specified. Thus, the VI for *cat* is likely specified for realizing the core meaning of *cat*, which according to DM would be a root (√CAT). Earlier work in DM (Halle and Marantz 1993, 1994, Halle 1997, Marantz 1997, Harley and Noyer 1999, 2000) suggests that roots came in only one variety and the syntax was not sensitive to the different ways that root could be realized. Pfau (2000), on the other hand, suggests that roots are in fact individually distinguished from the beginning of the syntactic computation in order to account for specific types of speech errors (word substitution, etc) within the framework of DM. This change to DM has largely been accepted and in general, in relevant analyses, roots are treated as being specific to the concept they refer to (see Embick and Marantz 2006 for example). That is, the numeration not only

---

licensed affix are possible the grammar predicts any combination of one or both of those to happen, seen in the table below.

regular root and regular affix	*walk*	*-ed*
regular root and irregular affix	*hit*	*-ø*
irregular root and regular affix	*slep*	*-t*
irregular root and irregular affix	*mice*	*-ø*

includes the formal features to be manipulated by the syntax, but the formal concepts that will later be interpreted by the Encyclopedia as well. Following that assumption, *cat* is likely specified to realize a formal instantiation of the concept of cat-ness, which can then manipulated by the syntax. Thus, a VI for *cat* might look like (3.5.a)

(3.5.a.) Vocabulary Entry for *cat*

√CAT → *cat*
/kæt/

The VI seen in (3.2) can only be inserted into a terminal node containing the very specific root √CAT. Just as the *-s* that realizes [present] can't be inserted into a node containing [past], *cat* can't be inserted into a node containing √DOG. However, a VI such as (3.5.b) would be perfectly able to be inserted into such a node:

(3.5.b.) Vocabulary Entry for *dog*

√DOG → *dog*
/dag/

Thus, the VIs in (3.5.a) and (3.5.b) compete against each other for insertion, deciding the winner based on which better matches the contents of the target node. (3.5.a) wins a competition into the node containing √CAT while (3.5.b) wins the competition for insertion into a node containing √DOG. Neither would win a competition to be inserted into a node containing √DUCK.

3.1.1 Licensing Insertion

While in English zero derivation from one "grammatical category" such as verb or noun is not uncommon (meaning the words like *pen, ink,* and *table* move easily between verb and noun without overt morphology), we are all at least intuitively aware that some words must be verbs while others must be nouns. For example, for the most part, *thrash* must be a verb and *technique* must be a noun.

(3.6) a) *Dave techniqued his writing skills.[9]
b) *Dave also awaited the thrash.

Thus, there has to be some way in which VIs restrict their insertion to nodes where the root receives the appropriate category. Currently in DM (Harley and Noyer 2000), the hypothesis is that VIs are licensed by the immediate syntactic environment through secondary exponence—for example, the immediately C-commanding functional head. Thus, *thrash* might have an entry such as (3.7).

(3.7) Vocabulary Entry for *thrash*

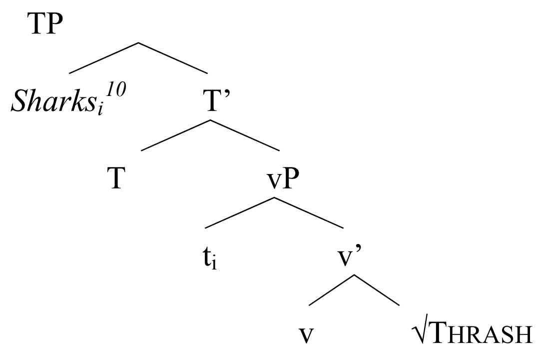
√THRASH* → *thrash*
/θræʃ /

*must be c-commanded by [v]

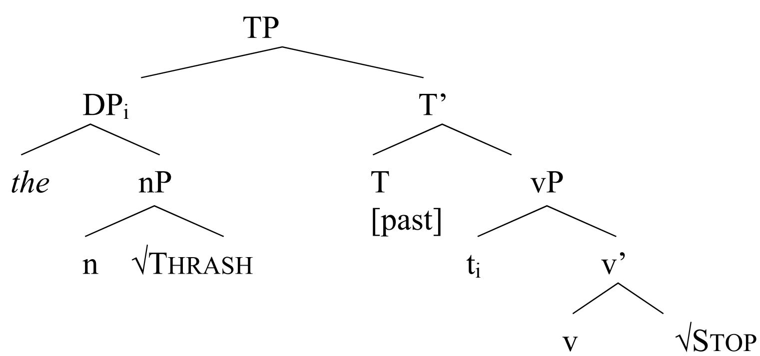
This type of specification allows the insertion of *thrash* into (3.8) where the root is dominated by vP, but blocks it from insertion into (3.9) where the root is dominated by nP.

[9] Given that zero derivation of nouns from verbs is so productive in English, this sentence becomes grammatical given enough context. Assume no context for now.

(3.8) Sharks thrash.



(3.9) *The thrash stopped.[11]



This formulation of licensing (through secondary exponence), whereby the VI checks the functional element c-commanding the target node rather than checking the rather than the target node itself is a side-effect of having VIs that realize roots not participate in competition. However, in the Pfau (2000) model of DM where the roots can compete, there is another way to capture the "noun"-ness or "verb"-ness of a given VI.

---

[10] As a simplification tool, I will be using italicized words (rather than triangles) to indicate summarized sections of trees.

[11] The tree shown here assumes *stop* to be unergative, just for ease of presentation. It was the only verb that even began to make sense with *thrash* as a noun.

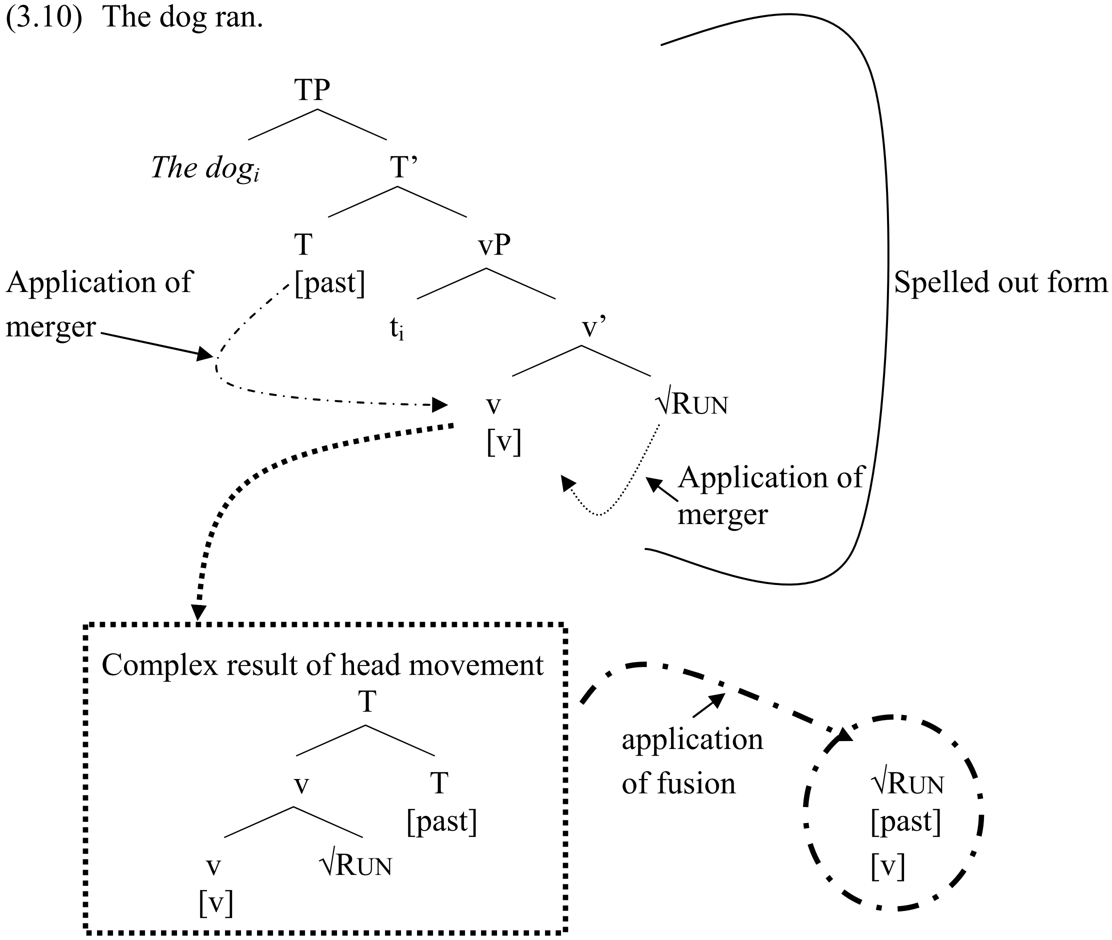
I propose that this is possible because of the applications of morphological merger and fusion to roots and the functional heads c-commanding them. In particular, I propose that the functional verbalizing element—little v (Kratzer 1996)—carries an interpretable feature[12], [v], whose syntactic content is something along the lines of "I am a verb". Similarly, the nominalizing head—little n (Harley and Noyer 2000)—carries the feature [n]. Furthermore, I propose that the root nodes themselves acquire these features through several applications of morphological processes. First, the root undergoes "head movement" (i.e. morphological merger) to adjoin to the functional heads above it. The resulting complex head then undergoes the process of fusion to incorporate all the features of the complex head (including the root) into one simplex head.

If we assume an application of morphological merger (called merger under adjacency) to the tense head and the verb as proposed by Bobaljik (1994), then the resulting head after fusion contains a root, a functional verbal element, and a tense feature.

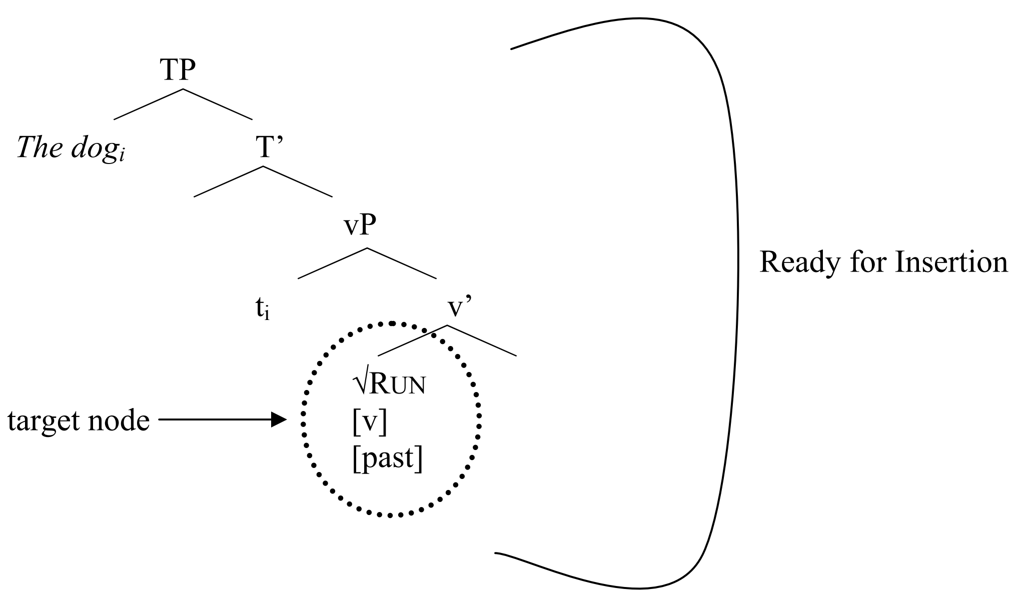
---

[12] or more likely a bundle of features—I will refer to it as one feature until the distinction becomes important in chapter 4.

(3.10) The dog ran.



(3.11) Resulting form



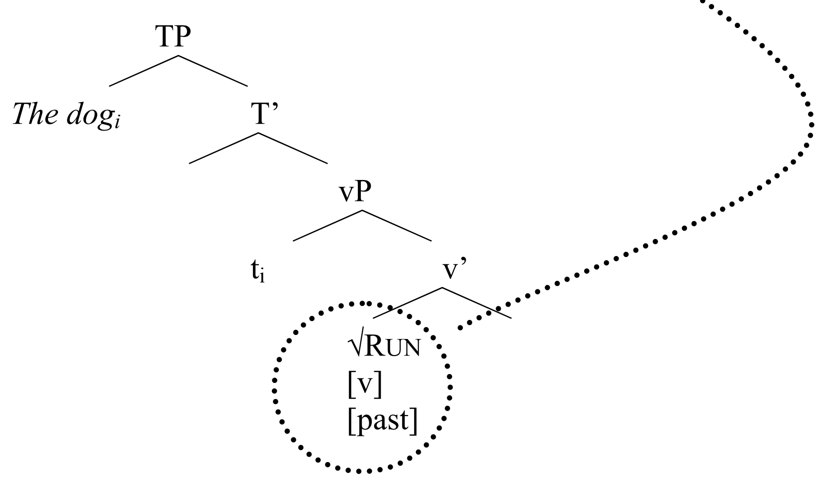
As seen in (3.10), the applications of head movement and fusion to the complex verbal structure results in a single simple node containing the formal features of the entire structure. The VI instead looks directly at the target node in the usual way to discover whether that node can satisfy the "functional" or grammatical requirements of that VI. For example, the VI in (3.12) can be inserted into the node created in (3.10, copied as 3.13) because the features it is specified for are a subset of those appearing in the node.

(3.12) Vocabulary Entry for *ran*

√RUN [v] [past]	→	*ran* /ræn/

compatible for insertion

(3.13) Resulting form

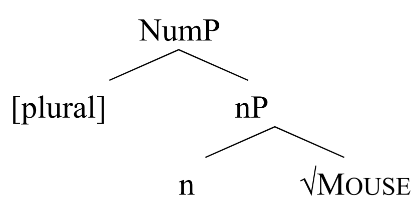


The example VI for *ran* requires two different functional features, [past] and [v]. While the specification for [v] identifies it as a verb, the specification for [past] sets it apart from *run*. Thus, similar to how *dog* will compete with *cat* for insertion into a node, *run* will compete with *ran* for insertion into (3.13), with *ran* winning the competition due to its more complete specification. Thus, the competition of roots enabled by Pfau's proposal of specific roots allows for another analysis of root allomorphy within DM.

3.1.2 Alternative Analysis of Root Allomorphy

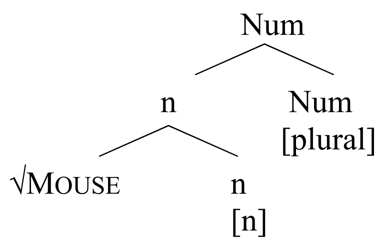
The processes of morphological merger and fusion described above apply to an account of root allomorphy. As an example, I show a derivation and insertion of the word *mice* using a fusion-based analysis. Recall the structure that results from the syntax above (3.1, repeated here as 3.14):

(3.14)



If we assume head-movement, all those heads move to one terminal node where they make up a complex head.

(3.15) Complex head resulting from head movement
(features contained in that head shown in square brackets)



The process of fusion is applied to complex heads such as that in (3.15) resulting in a simplex head. After the application of fusion, the resulting head is simplex and contains all the features previously in the complex head.

(3.16) [plural]
[n]
√MOUSE

This node, containing the root and several grammatical features, in now the target node for insertion. As we saw above with the VI for *ran* being specified for the feature [v] (indicating that it is a verb) and the feature [past], there is a specific VI for *mice* which is specified for the feature [past] and differs from the one for *mouse* in that it is more specified for number.

(3.17.a) Vocabulary entry for *mouse.*

√MOUSE	→	*mouse*
[n][13]		/maws/

[13] This specification means that *mouse* may only appear as a noun. *Mouse* may be underspecified for this feature if it can be used as a zero-derived verb. I use this specification for ease of presentation.

(3.17.b) Vocabulary entry for *mice*

√MOUSE	→	*mice*
[plural]		/majs/
[n]		

With the VI specifications in (3.17) we can show the following distributional patterns of the *mouse*/*mice* pairs: a) Since *mice* is the best specified of the two for the feature [plural] (*mouse* being underspecified for number), it will win competition into a node containing the [plural] feature (3.18.a). b) *Mice* would not win competition into a node containing [singular] since it is specified for the conflicting feature [plural], thus, *mouse*, being unspecified for number will be inserted into the singular environment (3.18.b). c) *Mice* would not win competition into one that contains neither feature (which hypothesis is the case in "general" constructions such as compounds like *mouse-trap*) since it is overspecified, thus again the underspecified *mouse* will win the competition (3.18.c)

(3.18.a) Competition for insertion into *two mice*

Target Node: √MOUSE [plural] [n]	
inserted candidate ◄------------	*mice:* √MOUSE [n] [plural]
underspecified ◄------------	*mouse:* √MOUSE [n]

(b) Competition for insertion into *a mouse*

Target Node: √MOUSE [singular] [n]	
overspecified ◄------------	*mice:* √MOUSE [n] [plural]
inserted candidate ◄------------	*mouse:* √MOUSE [n]

(c) Competition for insertion into *mousetrap*

Target Node:		
√MOUSE [n]	***overspecified*** ◄- - - - - - - - - -	*mice:* √MOUSE [n] [plural]
	inserted candidate ◄- - - - - - - - - -	*mouse:* √MOUSE [n]

3.1.3 Competition of Roots Revisited

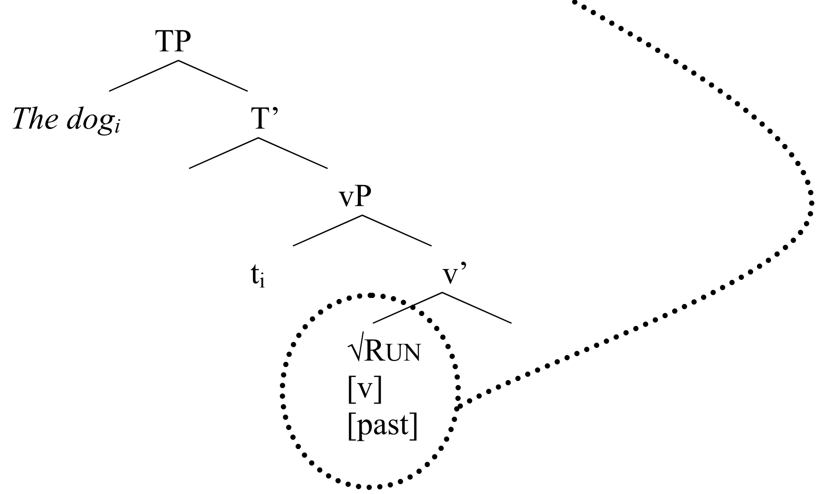
Suppose now that we had another word like *cat*, which is specified as a noun. Its entry (3.19) would be specified for the feature [n] rather than the feature [v], which would make it incompatible with the feature bundle found in (3.10, repeated as 3.20) because of both the specific root and the [n] feature.

(3.19) Vocabulary Entry for *cat*

√CAT [n] → *cat* /kæt/

(3.20) Resulting form

incompatible for insertion



Consider now the set of words *speech, speak,* and *spoke.* In previous incarnations of DM, the fact that these words were all linked to the same core meaning would be captured by postulating one VI that had a series of readjustment rules stipulated for it. Once we propose that root VIs are competing, there are three different VIs, one for each form. The feature specifications of all three VIs mention the same root, √SPEAK, but contain different functional material (3.21-3.23). Thus the common meaning shared among the three VIs is attributed to the fact that they are all specified for the same root rather than being phonological derivatives of one VI.

(3.21) Vocabulary Entry for *speak*

√SPEAK [v]	→	*speak* /spik/

(3.22) Vocabulary Entry for *spoke*

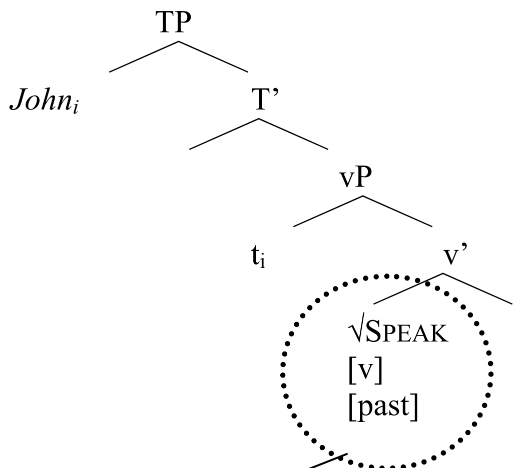
√SPEAK [v] [past]	→	*spoke* /spowk/

(3.23) Vocabulary Entry for *speech*

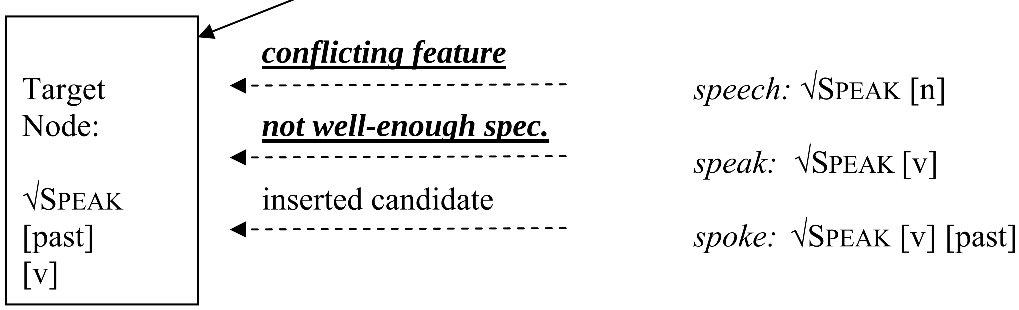
√SPEAK [n]	→	*speech* /spitʃ/

Since there are three different VIs, they can each participate in competition separately. In particular, they can all compete against each other. The VIs compete with each other for insertion into a target node, with the best-specified form winning the competition. For example, as seen in the sentence *John spoke* (3.24 and 3.25), *spoke* will win the competition into a past tense environment because it is better specified than *speak*. In the sentence *The speech began* (3.26 and 3.27), *speech* wins competition for insertion because both *speak* and *spoke* are specified for conflicting features ([v] and in the case of *spoke* [past]).

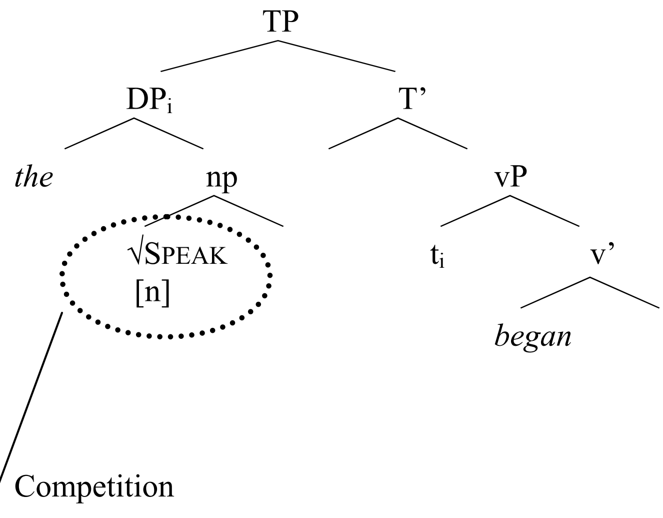
(3.24) John spoke.



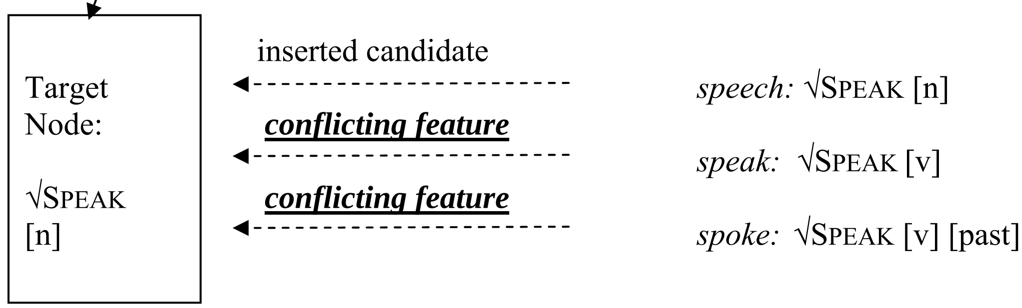
(3.25) Competition



(3.26) The speech began.



(3.27) Competition



As the example with *speak, spoke*, and *speech* shows, proposing that a) roots are contentful rather than general (Pfau 2000), b) functional nominalizing and verbalizing heads carry formal features, c) that those features come to adjoin to the root through head movement, and d) that those features fuse with the root to result in a simplex head with both content material and functional material allows us to argue that content morphemes such as *mouse* and *mice* compete with each other.

3.2. Differences between This and the Traditional DM Grammar.

The changes I have proposed above in section 3.1 make for a different model of DM than the one originally proposed by Halle and Marantz (1993, 1994). While the core fundamentals—i.e. syntactic structure of morphology, late insertion, the majority of morphological processes—remain unchanged, there are two key differences to this model that merit discussion. Proposing that *mouse* and *mice* are separate VIs eliminates the function of readjustment rules, whose job in traditional DM it is to change the unmarked form of a root VI into its marked form. In the model of DM sketched here, *mouse* and *mice* are separate VIs coincidentally linked to the same root. In addition, licensing a VI for insertion based on its "category" has historically been considered a part of that VIs *secondary exponence* (Harley and Noyer 2000). In the analysis here, that grammatical material is now part of a VI's *primary exponence*. In this section, I will detail the effects that this has on the model of the grammar and suggest that the exclusion of readjustment rules and the limitation of the scope of secondary exponence proposed here are positive adjustments to the model of the grammar.

3.2.1 Readjustment Rules

As discussed above, readjustment rules are no longer necessary in this model of the grammar. Since *mouse* and *mice* are separate VIs that compete with each other for insertion, there is now no need for there to be a rule that changes *mouse* into *mice*. An immediate upside to this treatment of root allomorphy is that it unifies the treatment of functional morphemes and content morphemes. In traditional DM, there are in essence two different grammars to account for allomorphy. Allomorphy of functional items (such as *el, la, los,* and *las* in Spanish) is captured through competition, the underlying assumption that a language's inventory of functional morphemes is entirely suppletive. On the other hand, root allomorphy is captured through the insertion of an unmarked VI whose phonology is readjusted given a particular conditioning environment. In the model of the grammar proposed here, all morphologically conditioned allomorphy is essentially suppletive and all VIs, even those realizing roots, compete in the same fashion for insertion. At the simplest level, the root allomorphy grammar proposed here and the one assumed by traditional DM differ in one key way: the grammar proposed here eliminates the need for the ubiquitous use of readjustment rules and relies only on competition, making it a simpler model of the grammar—one that contains fewer mechanisms and reduces the computational load.

As a side benefit, the particular aspect of the grammar that is discarded—readjustment rules—was already suspect. Though DM is a minimalist model of the grammar, and as such is a derivative of the GB tradition, readjustment rules are transformational. Further, despite the claim of some researchers such as Marantz (1997) and Halle (1997) (see also Embick and Noyer 2006) who claim that one rule can apply to a whole class of verbs, they are largely idiosyncratic and language specific. This means that a grammar has to have a major component that is composed of a long list of memorized readjustments. This adds to both the computational load of the grammar and to the memory load on the grammar. By proposing that each allomorphic pair is a set of memorized words, we don't alleviate the load on memory—each VI - readjustment rule pair is replaced with a VI-VI pair—but we do relieve the computational load of the grammar. In this model, there isn't the computational load of readjusting roots; instead, there is the already extant load of choosing the VI to be inserted.

However, this reintroduces the problem that Marantz (1997) appealed to in his justification of readjustment rules: learnability. Marantz proposed that readjustment was a solution to a central learnability concern with root allomorphy: how a learner of a language can learn that two VIs are linked to the same root. To put it into context, how does a learner ever learn that *mouse* and *mice* are linked to the same root? The mutual exclusivity constraint (proposed by Markman and Wachtel 1988) on the acquisition of language mandates associating new words with new concepts.   A learner's purposefully ignoring that constraint in the case of *mice/mouse* but not in the case of *car/truck* is difficult to explain. Marantz argued that suppletion was limited to functional morphemes, whose fundamental presence in the grammar is mandated by the language acquisition device itself—they do not have to be 'acquired' in the same way as roots. The learner can learn that in one context, 'be' is spelled out by the VI 'is' while in another it's the VI 'are' because the underlying grammatical structure has already informed them that this functional morpheme is present. This way the learner only has to memorize the formal specifications of a given functional morpheme. Never do they have to learn that two separate VIs are linked to the same root, avoiding that particular problem. Rather, the learner learns one VI corresponding to one root and then later acquires a readjustment rule to change that VI to its allomorphs.

This proposal immediately made the somewhat strange prediction, which Marantz acknowledges, that all clearly suppletive alternations must be of functional allomorphs—

they must not be linked to roots. Thus, it predicts that pairs like *go/went* and *bad/worse* are actually functional morphemes (not linked to a root) because they are instances of true suppletion and suppletion is limited to functional morphemes. Marantz claimed that all suppletive pairs that appear to be linked to roots would be limited to pairs such as *go/went* which aren't really "full verbs" but are better treated as light verbs, making them functional.

The existence of content morphemes that are undeniably suppletive would be counter evidence to Marantz's claim. One such form in English is *person/people* (cf. *mea* 'kill.sgO' ~ *sua* 'kill.plO' in Yaqui[14]; in Hopi, *tiyo* 'boy' (sg.)' ~ *tootim* 'boy' (pl.) and *wùuti* 'woman (sg.)' ~ *momoyam* 'woman (pl)'[15]). Thus, Marantz's claim that suppletion is only for functional morphemes, while ideal for addressing the learnability problem, is falsified by the data. Therefore, while the model of the grammar I propose here essentially removes Marantz's solution to that problem without proposing a new one, Marantz's solution already ran into the problem that there is in fact suppletion of content morphemes. This proposal, however, does have the effect of discarding the strange prediction that forms like *bad/worse* and *go/went* are not linked to roots. Rather, I propose here that suppletion is all that is necessary for all allomorphy, which reduces the number of mechanisms necessary to capture allomorphy.[16]

---

[14] Yaqui example provided by Heidi Harley

[15] Hopi example provided by Jason Haugen

[16] Another "function" readjustment rules is to capture the historic relationship between related forms. For example, the phonological process of i/j umlaut is what created the pairs *foot/feet* and *goose/geese*. Readjustment rules capture this relationship. However, a historical change such as i/j umlaut does not entail a synchronic change, removing that justification for the existence of readjustment rules.

3.2.2 Licensing

Recall from chapter 2 that in the pre-Pfau model of DM, VIs which belong to only one "class" (such as noun) are licensed by the c-commanding functional head according to the proposal by Harley and Noyer (2000). My proposal here evacuates the need for such licensing[17] as the licensing functional head fuses with the root and the root VI can select for its features. One of the inherent problems in using secondary exponence to license a root morpheme in an environment is that is strictly less local than it could be.

An example of words that have "lexical class" requirements that is the pair *speak/speech* as discussed above. Recall that if we hypothesize the following VIs for *speech* and *speak* licensing falls out naturally without requiring that the VI be able to look at the node above the target node for licensing.

(3.28) Vocabulary Entry for *speak*

√SPEAK [v]	→	*speak* /spik/

Vocabulary Entry for *speech*

√SPEAK [n]	→	*speech* /spitʃ/

Since the licensing of *speech* as a noun or *speak* as a verb in the model I propose here requires only looking at the node targeted for insertion rather than the part or all of the entire derivation, the model proposed here requires a significantly lighter computational

[17] Licensing itself is argued to be a part of *secondary exponence*, a feature of DM I allude to here (the selection of the zero morpheme for [plural] due to its adjoining to *mice* is an application of secondary exponence) and discuss in detail in Chapter 2. In suggesting that licensing of nouns, verbs, etc is no longer necessary, I do not mean to suggest that secondary exponence as a whole is not necessary. Rather, it is necessary for many reasons as discussed in Chapter 2.

load than secondary exponence model. For each content morpheme that has a distribution limited to its "lexical class", this model proposes a lighter computational load than a model of DM that requires licensing. Given the sheer magnitude of a language's inventory of content morphemes, the savings of computational load is drastic.

3.2.3 Null heads

A side effect of removing secondary exponence as the means for licensing class specific root morphemes is the drastic reduction of the number of null morphemes that DM is forced to propose. The original model of DM (Halle and Marantz 1993, 1994) proposes a large amount of null heads. Recall the derivation for *mice* (3.1) (repeated here as 3.29)

(3.29)
```
            NumP
           /    \
   [plural]      nP
                /  \
               n    √MOUSE
```

In DM, the structure after insertion is that in (3.30) wherein, √MOUSE is realized by *mice* (conditioned by the c-commanding elements), [plural] is realized by ø (an unpronounced head), and n is also realized by ø.

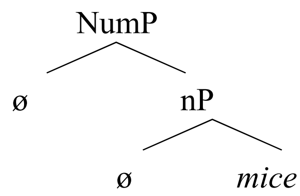
(3.30)
```
            NumP
           /    \
          ø      nP
                /  \
               ø    mice
```

In just this one little complex, for one overt VI, there are two null VIs. Such would be the case for all irregular forms similar to *mice* where there is no "overt" realization of an element that conditions a readjustment rule. This by itself is a large number of null morphemes. However, if you employ the traditional licensing through secondary exponence, those functional licensers are also overwhelmingly realized as null heads. Since the vast majority of "free" content morphemes would involve these null licensers (regardless of whether they are required by the VI or not) the number of these null morphemes in a given derivation is substantial and in the grammar as a whole would be enormous. DM is not unique in proposing null morphemes. What makes DM different is that it predicts a huge number of null morphemes in two cases where other theories of the grammar (such as Lexicalist models) don't have to – licensers (such as n) and morphemes such as [past] or [plural] that are realized as zero when a readjustment rule changes the root (as is the case with *mice* or *drove*).[18] While many models of grammar predict null heads, these particular null heads are unique to DM and create a perceived fault in the predictions of the grammar.

DM is susceptible to criticism because it predicts so many more null morphemes in a given derivation than do other models of grammar. The fusion analysis given here removes that particular criticism of DM as licensing heads and heads that condition readjustment rules are not realized by zero in this model but are instead fused with the root and are realized by the same VI that realizes the root. Recall (3.30) (repeated here)

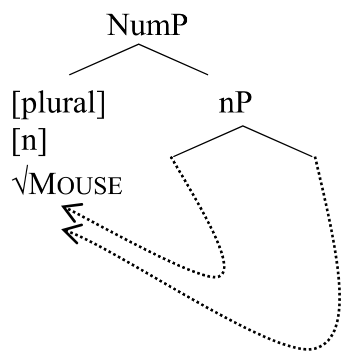
[18] Note that DM does not always change a functional head to zero when a readjustment rule is applied. In forms like *slept, better, went, houses,* etc, it is proposed that the root morpheme changes but the functional morpheme is still present and in its default form.

(3.30)



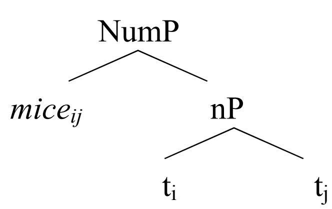
In the pre-Pfau model of DM, the licenser (n) and the head that conditions the application of the readjustment rule (num) are realized as zero. However, in the fusion account of root allomorphy proposed here, the root and [n] both move to [plural] and fuse (3.31)

(3.31)



Therefore, whereas the traditional DM account contains one overt head and two null morphemes, the fusion account sketched here contains only one overt head and two traces.

(3.32)



Thus, the fusion analysis simplifies the grammar by reducing the number of null morphemes in the inventory of a given language that has to be predicted by DM. Taken as a whole, the three central changes to DM—the rejection of readjustment rules, not using secondary exponence to license root VIs, and the reduction of zero morphemes—all drastically simplify the DM grammar. The reduction of both the computational load and the memory load[19] entailed by the grammar makes the grammar more economical. Further, the removal of readjustment rules discharges the last of the "transformational" rules in DM, bringing it more in line with recent GB tradition. Overall, the new model of DM proposed here is simpler and more attractive to the Minimalist community as a whole.

3.3 Expansion of the fusion analysis.

The combination of head movement, Bobaljik-style merger under adjacency, and fusion results in a tidy description of how the [past] feature fuses with the root √RUN to be realized by the word *ran*. However, even a cursory look at the details of the mechanics proposed in section 3.1.2 will show the prediction that formal features should always fuse with the root, ultimately resulting in only suppletion as a morphological tool. It seems that the current proposition actually removes the great strength of DM—that it predicts lexical decomposition as a function of the syntax.

For example, recall the different VIs presented for *speech, speak,* and *spoke* (repeated here as 3.33.a-c).

[19] The memory load is reduced in two ways: a) fewer null morphemes and b) two different types of memorization—readjustment rules and Vocabulary inventory—are reduced to just one—Vocabulary inventory.

(3.33.a) Vocabulary Entry for *speak*

√SPEAK	→	*speak*
[v]		/spik/

(3.33.b) Vocabulary Entry for *spoke*

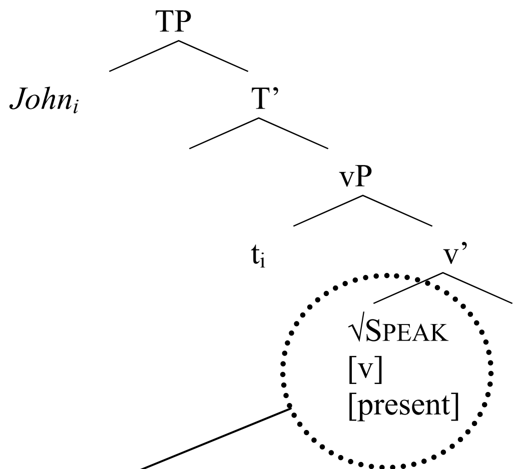
√SPEAK	→	*spoke*
[v]		/spowk/
[past]		

(3.33.c) Vocabulary Entry for *speech*

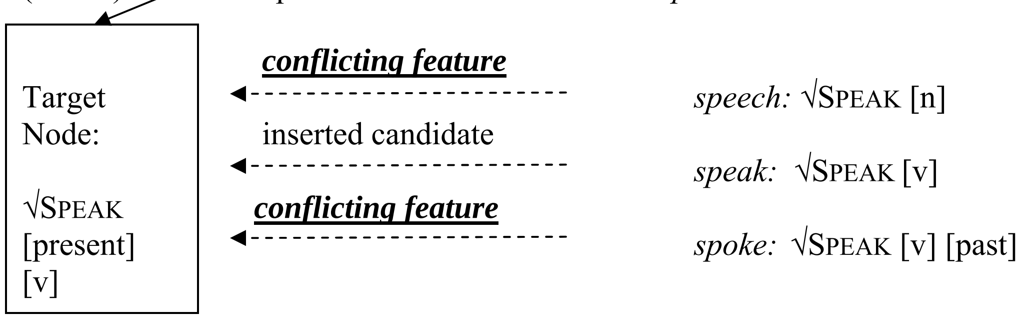
√SPEAK	→	*speech*
[n]		/spitʃ/

These three VIs very well predict that *speech* will appear in a nominal environment and *spoke* will appear in a past tense environment. However, it also predicts that *speak* will appear in *all other verbal environments*. In English, this is fine for most cases, except third person singular. For example, in the derivation of *He speaks,* the present tense and the phi-features are overtly spelled out by the affix *–s*. However, in an account where fusion applies to all complex heads, specifically fusing tense with the root, the tense should never be spelled out by its own head. The sentence should be **He speak.*

(3.34.a) John speaks.

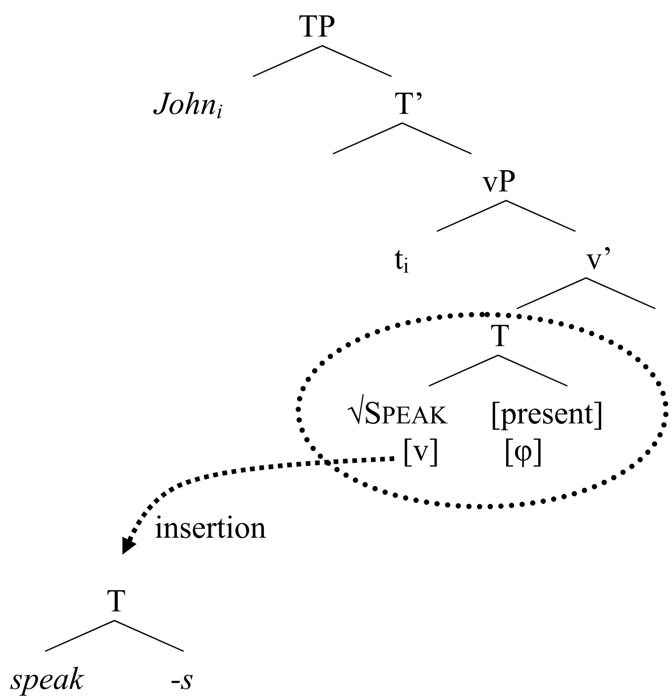


(3.34.b) Competition results in insertion of *speak*.



In order to maintain the decomposability of the form *speaks*—and for that matter all other concatenative morphology—the derivation must result in a complex head that has not been fused to a simplex form. Thus, the derivation for *John speaks* must actually be a tree such as seen in (3.35).

(3.35) John speaks.



If we make the reasonable assumption that there is always fusion, then we cannot have affixed forms such as *speaks*. Thus, we have to assume that complete fusion fails most of the time, which gives us complex forms. However, if we say that √SPEAK can fail to fuse with [present] in order to result in *speaks* we must also predict that it can fail to fuse with [past] and result in **speaked.* Thus, the application of fusion has to be blocked in order to trigger regular concatenative morphology, but must be allowed in order to trigger suppletion.

3.3.1 The ¬ specification.

The immediate problem described above is that, as outlined above, regular forms like *thrash* are going to be inserted into a node where the [past] feature had fused with the root, resulting in no overt realization of the [past] morpheme. I propose that the solution lies in specifying the VI for *thrash* for an incompatibility with the feature [past], ensuring that the VI will not be inserted into a node containing that feature.

(3.36) Vocabulary Entry for *thrash*

√THRASH	→	*thrash*
[v]		/θræʃ/
¬ [past]		

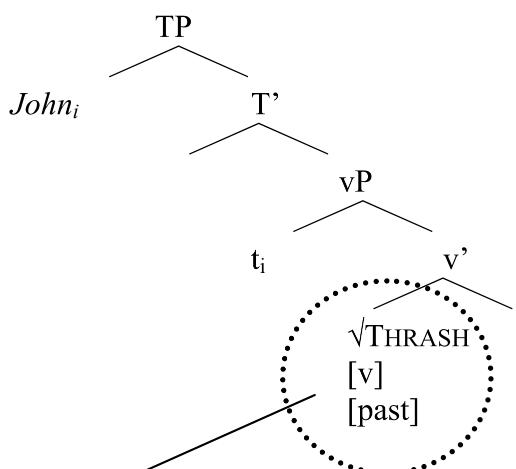
I specified *thrash* with the specification ¬ [past] (read: "not past"). I use this notation to indicate that the Vocabulary Item *thrash* is not compatible with the feature [past]. What this means is that the VI cannot be inserted into a node containing the feature [past] This specification is in many ways the inverse of the normal specification used by DM. For example, *thrash* as specified above for [v]. In terms of DM, this specification means that *thrash* must realize the feature [v]. *Thrash* is also specified as ¬ [past]. That means that *thrash* must not be inserted into a node containing the feature [past]. With the addition of the ¬ specification, a VI can now lose a competition for three reasons: 1) not being well enough specified (i.e. there is a better specified candidate), 2) containing a conflicting feature (i.e. the VI is specified for a feature that is not present in the node), and 3) being specified for incompatibility with a feature present in the target node. So this type of competition would occur as in (3.37), imagining a word *thras that was the nominalization of *thrash*.

(3.37) Competition

Target Node: √THRASH [v] [past]	***conflicting feature*** ◄- - - - - - - - - - - - - -	**thras:* √THRASH [n]
	blocked feature ◄- - - - - - - - - - - - - -	*thrash:* √THRASH [v] ¬ [past]

With this type of notation, the insertion of a regular word like *thrash* or *walk* cannot occur into a node that has fused with the past tense.

(3.38.a) John thrashed.



(3.38.b) Competition

Target Node: √THRASH [past] [v]	***blocked feature*** ◄- - - - - - - - - - - - - -	*thrash:* √THRASH [v] ¬ [past]

The only possible VI in the inventory of English that can realize the target node in (3.38.b) is the past tense affix *–ed, which* is specified only for the feature [past]. Crucially, in addition to discharging but not realizing the feature [v], inserting *–ed* also discharges the root without realizing it. What sets content VIs apart from functional VIs is that they realize a root. The "special"-ness of roots is that they that contribute the extra-grammatical meaning to an utterance. In other words, a sentence doesn't properly convey a message if all the roots are discharged by functional morphemes without overt realization. Thus this "special"-ness of roots is captured by a constraint on insertion that a root must be realized by the VI which discharges it.

Since there is no VI that is specified for the root √THRASH and is not blocked from the feature [past], no VI can be inserted into the derivation and thus the derivation crashes. The only derivation containing both the feature [past] and the root √THRASH that will converge will be one where the root and the [past] feature have failed to fuse, thus resulting in insertion of both *thrash* and *–ed.*

3.3.2 Enter MINIMIZE EXPONENCE

The constraint MINIMIZE EXPONENCE proposed in Chapter 1 enters play at this point in the analysis of root allomorphy. Up until now, fusion of the root with the functional material accounts nicely for root allomorphy and proposes a simpler model of DM. However, to this point I have not shown what motivates this fusion. Whereas the model of grammar proposed here is more economical in the sense that it reduces the computational load entailed by readjustment rules and secondary licensing, I have proposed the ubiquitous application of the morphological process of fusion, which makes the computational load of this grammar *heavier*—i.e. now a given derivation is littered with applications of fusion, which entail more computation for every derivation. I will show that this addition of the process of fusion to the derivation, while more computationally intensive actually satisfies a larger economy constraint at the cost of extra computation. That constraint, of course, is MINIMIZE EXPONENCE.

Returning to the *speak* trio, *speak*, like all verbs in English, is likely specified for incompatibility with some feature in the complex of [3rd person], [singular] and [present], since all verbs in most standard English dialects have third singular present overtly spelled out with the affix *–s* (i.e. there are no strong verbs where there is an irregular form for third person singular present). For ease of presentation, I will indicate that this incompatibility with the notation ¬ [3sg], despite the fact that this notation simplifies a more complex phenomenon. Thus, the three words are specified as follows:

(3.39.a) Vocabulary Entry for *speak*

√SPEAK	→	*speak*
[v]		/spik/
¬ [3sg]		

(3.39.b) Vocabulary Entry for *spoke*

√SPEAK	→	*spoke*
[v]		/spowk/
[past]		

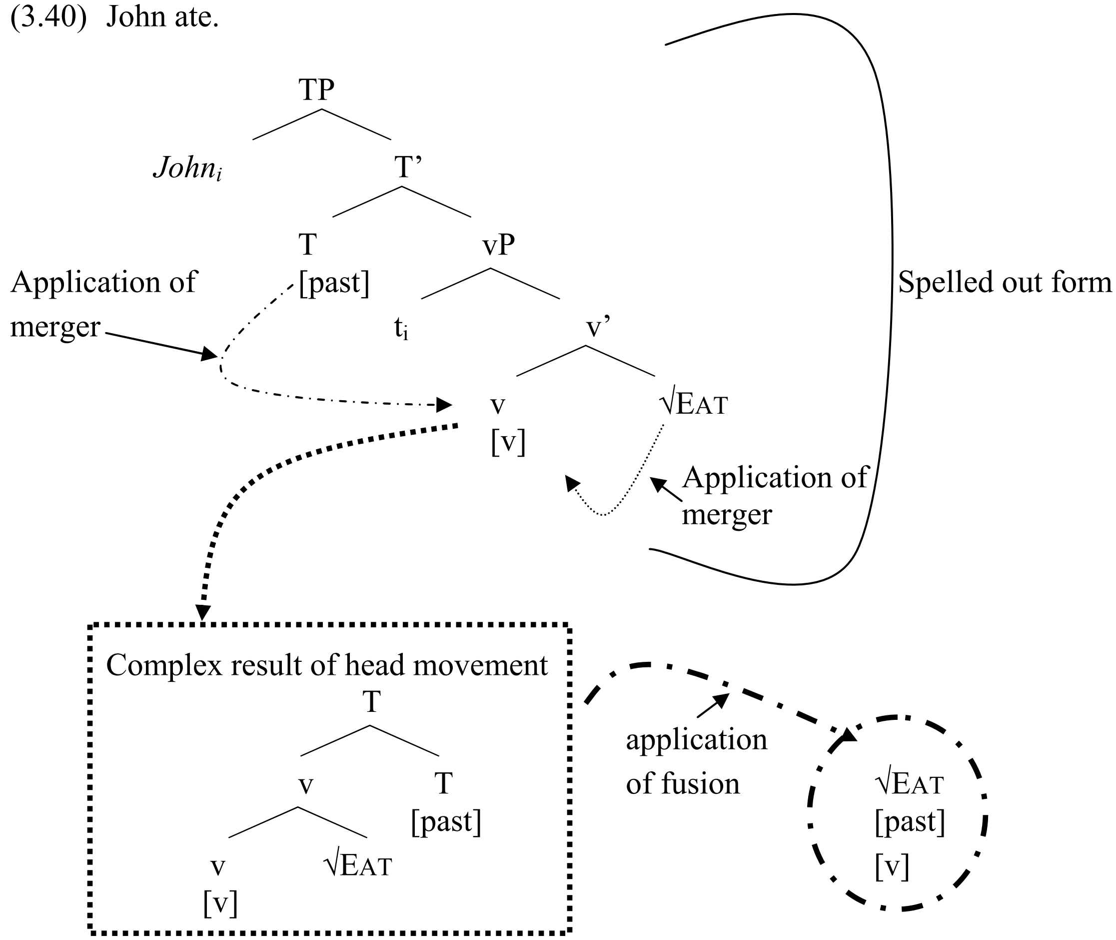
(3.39.c) Vocabulary Entry for *speech*

√SPEAK	→	*speech*
[n]		/spitʃ/

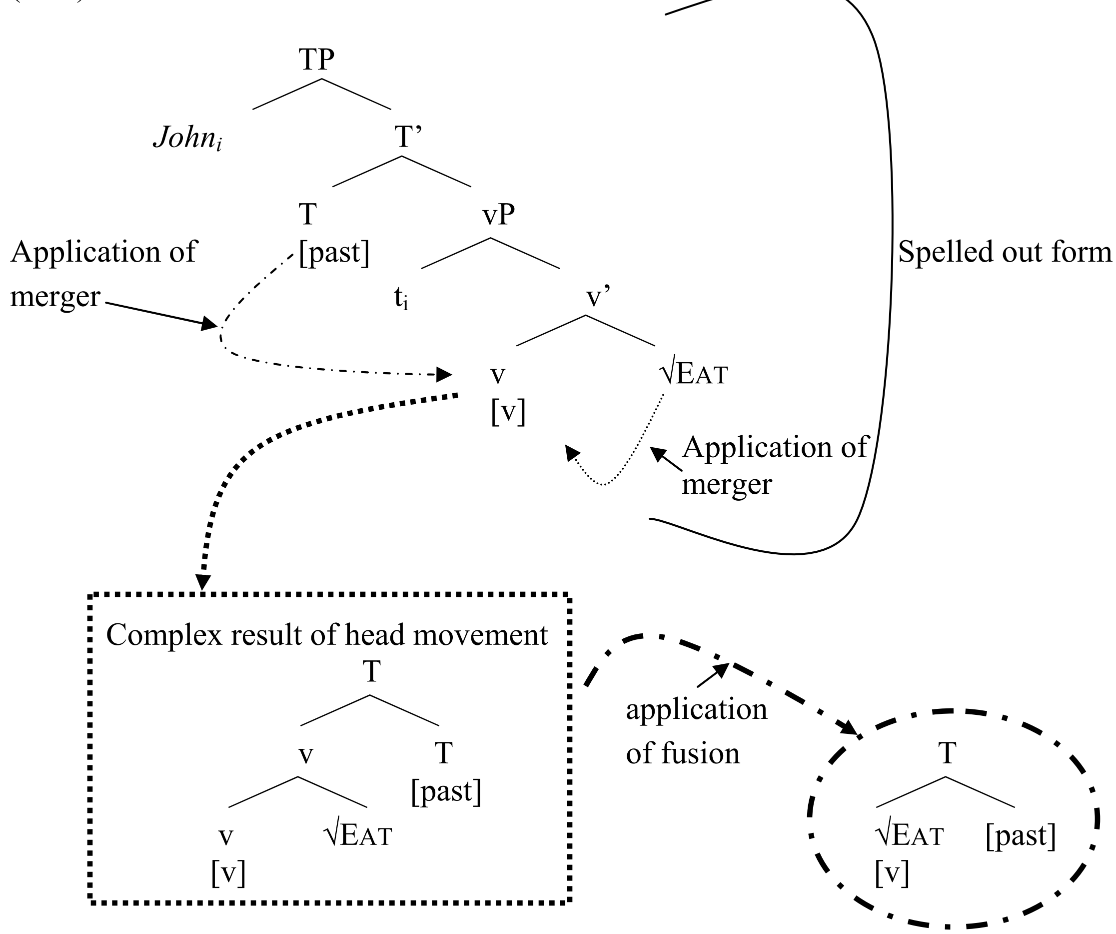
These specifications will result in exactly the distribution that we see with the three forms. However, this analysis relies entirely on there being two different derivations competing for convergence—the one where the [3sg] feature has fused and the one where it has failed to fuse. It follows then that there must be those two derivations, one fused and one unfused, for all utterances.

Imagine the two following possible derivations for *John ate.* In the derivation in (3.48), [past] has fused with the √EAT, meaning that the VI candidate that will win competition will be *ate*, which is specified for [past]. However, in (3.40), the node containing √EAT has not fused with the feature [past]. That means that even if *eat* is specified for ¬ [past], that specification will not stop *eat* from being inserted because the past tense is in another node. Both *eat* and *–ed* would be inserted, meaning that the same set of formal features could result in two possible utterances, **John eated* and *John ate.*

(3.40) John ate.


TP
$John_i$
T'
T
[past]
vP
t_i
v'
v
[v]
√EAT
Application of merger
Application of merger
Spelled out form
Complex result of head movement
T
v
T
[past]
v
[v]
√EAT
application of fusion
√EAT
[past]
[v]

(3.41) *John eated.



Since **John eated* is ungrammatical, it cannot be the case that two derivations for the same feature set exist in parallel. Since a derivation with √EAT fusing with [3sg] will crash because there is no VI compatible with the derivation, the trouble is that there is no as-yet-formulated way that explains why **John eated* crashes.

I propose that this is where MINIMIZE EXPONENCE, an economy constraint on the grammar that prescribes that the most economical derivation is the one that is realized by the fewest Vocabulary Items, is used to choose the most economical derivation (and thus the one that will converge) between the two possible derivations of the same set of formal features. In other words, the most economical utterance is the one that realizes all the formal features that need to be realized using the fewest words possible. Recall from Chapter 1, the definition of MINIMIZE EXPONENCE (repeated here as 3.42)

(3.42) MINIMIZE EXPONENCE: The most economical derivation will be the one that maximally realizes all the formal features of the derivation with the fewest morphemes.

The need to use as few words as possible and the need to express all the formal features are both satisfied by using fusion. In that way, more than one formal feature is realized by just VI, reducing the time and energy needed to utter one sentence. The logical extension of this is that the ideal language would maximize the number of forms like *ate* which capture both roots and formal features. However, that would mean a much larger inventory of stored words, which is also inefficient. Thus the compromise is to have fused forms for the most frequently used roots while leaving less frequent forms to regular morphological processes. We see this effect of MINIMIZE EXPONENCE on the Vocabulary cross-linguistically, lending credence to my proposal of its existence.

Thus, fusion is motivated by an effort to reduce an utterance to the fewest pronounced morphemes (null morphemes and morphemes with overt morphology are indistinguishable to this constraint). What blocks complete fusion of all functional material with the content material it c-commands is the limitations of the Vocabulary inventory of a given language as seen above and elaborated below.

To show Minimize Exponence in action, we return to the ungrammaticality of **eated.* The two derivations, *John ate* and **John eated*, can be evaluated for economy based on the constraint MINIMIZE EXPONENCE. In *John ate*, √EAT and [past] are realized by only one VI. In **John eated*, √EAT and [past] are realized by two VIs, *eat* and *–ed*.

Thus, as far as MINIMIZE EXPONENCE is concerned, *John ate* is the more economical derivation and the one that converges. On the other hand, consider the word *climb.* Since there is no **clomb*, a word that is specified for √CLIMB and [past], and *climb* is presumably specified for ¬ [past], *climbed* is the most economic derivation possible since one where √CLIMB and [past] are realized by fewer VIs (i.e. only one) is not possible.

In this way, MINIMIZE EXPONENCE will force fusion to occur wherever it can without resulting in a crash. Thus, the fusion we see in this chapter is actually driven by the need to make the utterance contain as few morphemes as possible. Though fusing the complex nodes that result from head movement involves more of a computational load, the utterances ultimately created by a grammar containing MINIMIZE EXPONENCE are more economical measured in the amount of energy consumed producing them rather than in deriving them.

3.4 Inflection in Compounds

The MINIMIZE EXPONENCE analysis presented here offers a new analysis of a classic observation about irregular root allomorphy. Jesperson (1909) observes that in addition to traditional synthetic and analytic compounds, English contains a variety of compounds that exhibit a unique behavior. Since the non-head member of these compounds is always a noun, these compounds are usually called *nominal compounds*.

The unique behavior of these nominal compounds is that, normally, inflection is not allowed in the non-head (left) member of the compound, even if the interpretation of the compound would require that inflection. For example in (3.43, adapted from Sproat

1985), a *coat-rack* is specifically designed to hold more than one coat, yet it cannot be called a **coats-rack*.

(3.43)	*dog-lover*	**dogs-lover*
	rat-chaser	**rats-chaser*
	log-cutter	**logs-cutter*
	hand-made	**hands-made*
	finger-bowl	**fingers-bowl*
	coat-rack	**coats-rack*

This seems to be the effect of some constraint on the grammar of English that disallows inflection in the non-head member of a nominal compound. However, a number of researchers have observed that when the non-head member of the compound is an irregular (i.e. it undergoes some sort of root allomorphy), the compounding of an inflected form is allowed. (see 3.44, adapted from discussion in Thomas-Flinders 1981 via Sproat 1985 and from Kiparsky 1982). For example, while a head infested with fleas must be *flea-infested* and never **fleas-infested*, the same head infested with lice can either be *louse-infested* or *lice-infested*. Similarly, a group of people jumping into a pool leading with their heads cannot be jumping **heads-first*, but those same people could jump *feet-first* (example adapted from Sproat 1985).

(3.44) *feet-first*
lice-infested
teeth-marks
alumni club
dice pool
people eater

Kiparsky's (1982) analysis for this data was couched within his model of Lexical Phonology and Morphology (LPM). He proposed that morphological operations happen at several levels. The first incarnation of his proposal included a Lexical level and a grammatical level. His recent revisiting of LPM made it compatible with Optimality Theory (LPM-OT) and proposed three different levels: the stem, the word, and the phrase (Kiparsky 2003).

In both of his theories, this pattern where irregulars behave one way and regulars behave another is captured by level-ordering. In LPM, Kiparsky proposed that the rule that inflected irregular forms occurred before the compounding rule, which itself happened before the rule inflecting regular forms.

(3.45) Ordering of operations:

1. Inflect irregular forms
2. Compound
3. Inflect regular forms.

Since *louse* is inflected to *lice* before compounding, *lice-infested* is permissible. However, since compounding happens before regular forms are inflected, compounding bleeds the regular inflection. Once *rat* has compounded with *infested* to form *rat-infested*, *rat* is no longer a valid target for the operation that adds *–s*. This means that **rats-infested* is blocked but *lice-infested* is permitted.

In LPM's modern incarnation, LPM-OT, which doesn't use rules but rather uses constraint hierarchies, the prohibition of regular inflection in compound but the permission of irregular inflection arises because the processes occur at different levels of the grammar, which have different constraint rankings.

There have been many objections to the use of level ordering analysis of this phenomenon, especially given that many models of the grammar reject either the idea of level ordering or the particular levels that Kiparsky proposed. One such objection is that regular inflection is not always prohibited from these structures. Hammond (1984) observed that the inflection is allowed in compounds where the plural form is interpreted as a group meaning (3.46, adapted from Hammond 1984). For example, an *admissions committee* isn't necessarily in charge of several *admissions*, but rather the process of *admissions*.

(3.46) *systems analyst*
parks department
admissions committee
numbers racket
reservations desk

Similarly, *pluria-tantum* forms permit inflection of the non-head member of the compound (3.47, adapted from discussion in Sproat 1985). For example, in addition to *pant-pocket* being a legal compound, so is *pants-pocket*.

(3.47) *pants-pocket*
alms-giver
odds taking

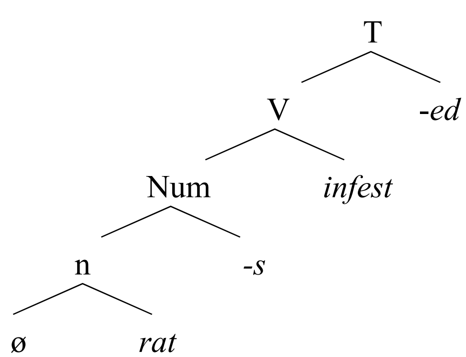
To complete the picture of the data, it is important to show that while regular inflectional morphology is blocked in English nominal compounds, regular derivational morphology is allowed.

(3.48) *grammaticality judgment*
grading session
participation grade
marketing suggestion
cooler unit
shifter knob
copier service
unhappiness factor

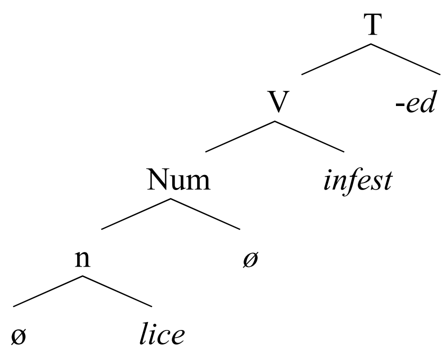
An analysis of this pattern within the framework of DM is difficult for a number of reasons. The first is that DM contains no generative lexicon. Thus, the analysis where by compounds are created in a level-ordered lexicon is unavailable. By DM principles, if one structure (regular inflection) is disallowed and another (irregular inflection) is allowed, that must be indicative of different syntactic structure in the two forms.

However, in DM, the structures for **rats-infested* and for *lice-infested* are largely identical. The first difference between the two is that in **rats-infested* the plural is realized by an overt VI while in lice-infested the plural morpheme is realized by a null morpheme. According to the tenets of DM, a VI without phonology is no different from one with overt phonology. Thus, this cannot contribute to a difference in grammaticality. The only other difference is that, in *lice-infested,* there has been the application of a readjustment rule. Again, the application of a readjustment rule should not affect the grammaticality of a syntactic structure.

(3.49.a) **Rats-infested*



(3.49.b) *Lice-infested*



The next important problem for the framework of DM is that it makes no distinction between "inflectional" affixes and "derivational" ones. From the point of view of DM, both are just the overt realization of terminal syntactic nodes with VIs that happen to be bound rather than free. Since DM doesn't recognize a difference between these two different types of affixation, referring to this difference is not a possible way to explain the grammaticality difference between derivational morphology being allowed in these structures but inflectional morphology being disallowed. Again, with the exception of the specific functional heads involved, the structure of *cooler-unit* and **rats-trap* should be identical.

The inclusion of the economy constraint MINIMIZE EXPONENCE to the framework of DM makes possible an analysis of the banning of inflection within English nominal compounds. In section 3.4.1, below, I detail that analysis.

3.4.1 MINIMIZE EXPONENCE analysis

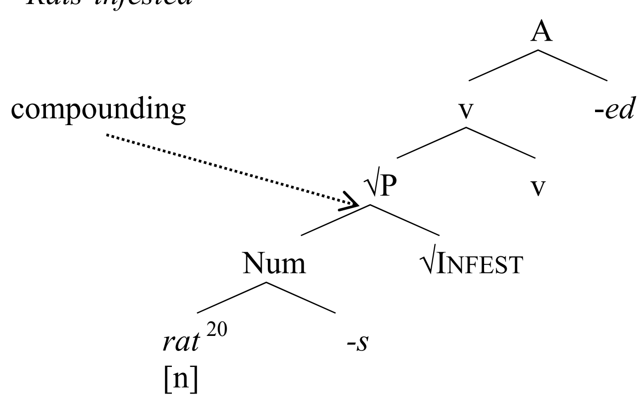
Compounding, in all of its forms, is an application of Merge in the syntax that creates a phrase that is later adjoined into one "word" by an application of morphological merger. In particular, compounding is an application of morphological merger to a pair of nodes α and β, where α is a phrase (X^n, n>0) and β is a root, dominated by the phrase (or √P). The √P resulting from Merge can then be sister to functional heads such as little-n or little-v in order for the compound as a whole to participate in the syntax and receive argument structure or undergo affixation.

The difference between the phrases that undergo compounding and those that undergo affixation is the difference of labeling in the syntax proper (before spellout). In particular, what will become compounds after spellout are dominated by a root phrase whereas affixation is dominated by the functional head. Thus the application of Merge that generates a compound merges X^n (n>0) to any √ and projects √ whereas affixation merges $\surd^n$ to X^0 and projects X where X is a functional head.

I argue that nominal compounds are an application of morphological merger in English that adjoins a noun (n^n, n>0) to a root under a projection of that root. More specifically, nominal compounds are the joining of the feature [n] to a root. Since the feature [n] is imbedded in the case of a regularly inflected form but, as a result of fusion, is not imbedded in irregular forms, morphological merger can target an irregularly inflected form but not a regular.

If nominal compounds are the result of an application of morphological merger to two adjacent heads where one is a root and the other bears the feature [n], we can see why regular inflection is blocked in forms like **rats-infested* (3.50).

(3.50) **Rats-infested*



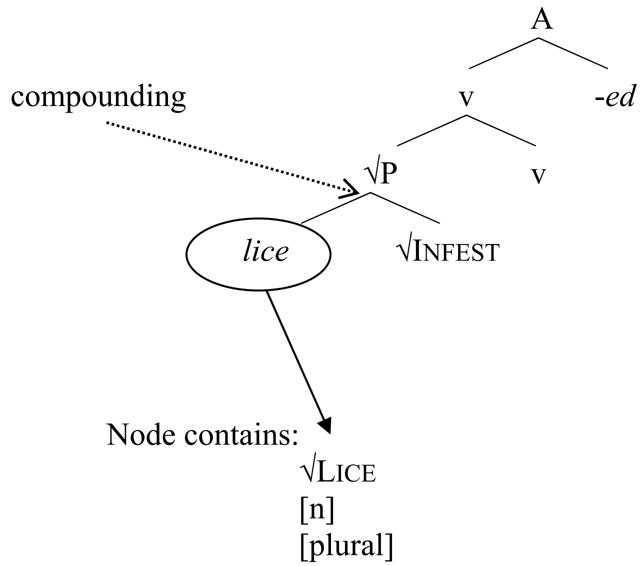
In such forms, even after fusion, the feature [n] is imbedded below the inflection (in this case Num). Since feature [n] is not adjacent to the root, compounding cannot target the two nodes and merge them into one form (the root is adjacent to the feature [plural] in (3.50)).

However, in forms with irregular inflection such as *lice-infested*, the root √LICE, the n-head, and the num-head have all fused to make one node. This means that the [n] feature is no longer embedded and is now adjacent to the root √INFEST. This adjacency of the [n] feature to the root allows compounding to target that structure.

---

[20] Shown after fusion.

(3.51) *Lice -infested*



Regular forms such as *rats* are two VIs composing a phrase dominated by *–s,* whereas irregular forms like *lice* are only one VI (with no complex structure after fusion). The VI for *lice* realizes both the root and the feature [n]. Thus, before insertion, the feature [n] is adjacent to the dominating root.

Similarly, pluria-tantum cases and cases like *admissions committee* and *systems analyst* are the effect of a root that has fused with an [n] feature compounding with another root. For example, on this analysis, *pants, scissors, odds,* and *alms* are each really just one Vocabulary Item (they refer to one thing) despite what appears to be surface morphology. That is the VI for *scissors* is (3.52).

(3.52) Vocabulary Entry for *scissors*

√CUTTING APPARATUS[21]	→	*scissors*
[n]		/sɪzr̩z/
[plural] [22]		

Similarly, an *admissions committee* is in charge of the process of admissions not several admissions, the *numbers racket* involves the process of gambling, not dealing with several numbers. Nouns like *admissions* and *numbers* are really simplex forms, not complex ones containing inflection. Some may not even be linked to the same root as the apparent stem (such as is the case with *numbers racket*). Rather, they are just one node with just one corresponding VI.

(3.53.a) Vocabulary Entry for *numbers*

√GAMBLING	→	*numbers*
[n]		/nʌmbr̩z/
[plural]		

(3.53.b) Vocabulary Entry for *scissors*

√CUTTING APPARATUS	→	*scissors*
[n]		/sɪzr̩z/
[plural]		

[21] I use this root to illustrate that the root is really about the referent not about the word. An unfortunate effect of the language being studied and the metalanguage being the same is that it is sometime not clear that the concept and the word are not the same thing.

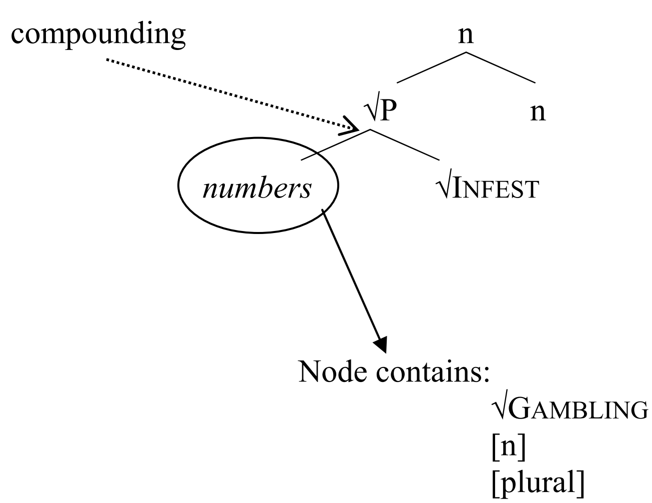
[22] we can assume that all these forms and *pluria tantum* forms are specified for [plural] because they cause plural agreement.

(3.53.c) Vocabulary Entry for *admissions*

√ADMISSIONS	→	*admissions*
[n]		/ædmɪʃn̩z/
[plural]		

Since *pluria-tantum* forms such as *scissors* and “group” forms such as *admissions* are one VI, they are inserted into a node where the root has fused with all the dominating functional material, including [n] and [plural], above them just as *lice* is inserted into a node where the root has fused with those features. That being the case, again in forms like *pants-pocket* and *numbers-racket* the feature [n] is adjacent to the dominant root in the compound so the application of morphological merger is legal (3.54).

(3.54) *Numbers racket*



Having addressed the banning of regular inflection in nominal compounds and the admission of irregular inflection, pluria-tantum forms, and “group” nouns, we can address why forms with derivational morphology are allowed despite the ban of having morphology dominating the root. Recall that the targeting restriction of nominal compounds is that the application of compounding must target the feature [n]. Since all nominalizing heads contain the feature [n], we expect the derivational morphology should be allowed in compounding exactly when *the most dominant derivational morpheme is a nominalizing head*. When the non-head member of a nominal compound is dominated by a nominalizing head, the maximal projection of that head, and thus the feature [n] is directly adjacent to the root and thus can be the target of morphological merger.

This is exactly what we see in the data (3.48 repeated here as 3.55)

(3.55) *grammaticality judgment*
grading session
participation grade
marketing suggestion
cooler unit
shifter knob
copier service
unhappiness factor

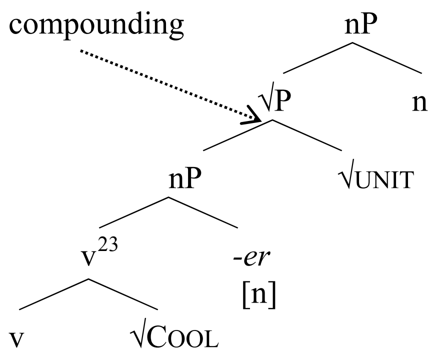
As shown in (3.55), the derivational morphology allowed in the non-head position is not unbounded. Rather, the dominant derivational morpheme is always a nominalizing element, as seen in (3.56).

(3.56)
-ity
-ing
-ion
-er
-ness

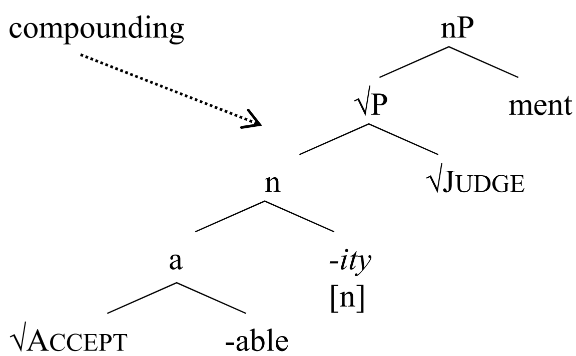
Since each affix is a nominalizer, each realizes the feature [n]. Thus, each of the derived words in the non-head position of a nominal compound is dominated by the feature [n].

This allows them to be the target of nominal compounding (since the feature [n] is not embedded).

(3.57) *cooler unit*



(3.66) *acceptability judgment*



To summarize, nominal compounding is an application of morphological merger that targets a root and the feature [n] and adjoins them. In the case of regular inflection in the non-head member (such as **rats-catcher*), that feature [n] is inaccessible to the application of merger because the [n] feature is imbedded. However, as a result of the fusion driven by the economy constraint MINIMIZE EXPONENCE, in irregular forms such

[23] This head here probably fuses. I have shown it unfused so the derivation is clear.

as *teeth-marks*, *pluria-tantum* forms such as *pants-pocket*, and "group noun" forms such as *parks service,* the [n] feature is no longer embedded and is now adjacent to the dominating root. Compounding can now apply, resulting in grammaticality. Finally, in the case of derivational morphology, the compounding is grammatical because, even without fusion, the [n] feature is already adjacent to the root targeted for compounding.

Recall the two reasons that the blocking of regular inflection in nominal compounds was problematic for Distributed Morphology: 1) the structures of regular inflection and irregular inflection were identical and 2) there is no difference between derivational morphology and inflectional, so there was no good way to block one and not the other. The analysis here solves both those problems and provides an analysis for this phenomenon within DM. 1) I propose here that the structures of **rats-infested* and *lice-infested* are in fact not identical. Where *rats* has failed to fully fuse (thus the overt realization of [plural]), *lice* has fully fused. 2) The blocking of **rats-catcher* versus the grammaticality of *cooler unit* is not due the difference between derivational morphology and inflectional morphology, but rather is due to the embeddedness of the feature [n]. Thus, DM can maintain that there is no difference between derivational and inflectional affixes.

3.5 Summary

The purpose of this chapter was to present an alternate analysis to lexical categories and root allomorphy within the framework of Distributed Morphology. The traditional analysis of lexical categories in DM is to have the content VI licensed for insertion through secondary licensing (targeting the immediately c-commanding functional head). The traditional DM analysis of root allomorphy employs the use of readjustment rules (again conditioned by c-commanding functional heads) to change the phonology of on VI to another VI. In the analysis presented in this chapter, I argued that licensing is a local process of targeting features actually in the node targeted for insertion. I argued that the root and the formal features come to occupy the same terminal node through the application of the process of fusion. I extended this analysis to an analysis of root allomorphy. This extension allows words like *eat* and *ate* to be separate VIs that compete with each other for insertion rather than one being the result of the application of a readjustment rules.

This proposal has a number of effects on the DM model of grammar. The first is that it limits application of secondary licensing, removing lexical categorization from the list of responsibilities of secondary exponence. Since secondary licensing is inherently less local and thus less efficient than primary licensing, this is taken to be a strength of this proposal. The other important effect of this proposal was to evacuate the need for readjustment rules in the DM grammar. Since VIs linked to the same root now compete with each other for insertion, readjustment rules are no longer needed to alter the phonology of irregular forms. As readjustment rules only serve the purpose of changing the phonology of irregular morphemes, since they aren't needed for that function any longer, the DM grammar doesn't need them at all. Since readjustment rules entail both long distance relationships and extra computational load, another strength of this dissertation is that it proposes a model of DM without readjustment rules.

In the current model of DM, functional morphemes and content morphemes in many ways participate in two different grammars. Whereas functional VIs participate in competition, content VIs did not, relying on readjustment rules and secondary exponence (never primary exponence) to license their insertion. The proposal here, in abolishing the need for both secondary licensing and readjustment rules and in elaborating on a theory of competition of roots, shows that it is possible for the model of the grammar to use just one process for insertion of both functional and content VIs. As this simplifies the model of the grammar, this is again taken as a strength.

This chapter also served to showcase MINIMIZE EXPONENCE. As proposed in chapter 1, the purpose of MINIMIZE EXPONENCE is to capture the conflict between the need to be maximally contrastive and maximally efficient at the same time. The realization of MINIMIZE EXPONENCE here is the ubiquitous fusion of functional heads in complex head arrangements. The effect of this fusion is that the features of those heads are realized without each needing to be realized by its own VI. The efficiency in the grammar comes from there being fewer VI items to be pronounced without loss of contrast (any ambiguity is then only a result of a given language's inventory). Thus, fusion is one possible tactic available to the grammar to satisfy this central conflict. As an added effect, a large portion functional heads that are realized as null morphemes in DM instead fuse with other heads to be realized by overt morphemes. This drastically reduces the number of null heads predicted by DM.

Finally, I proposed an analysis for the blocking of regular inflection in nominal compounds in English (e.g. **rats-infested*). For DM, this phenomenon has been difficult to analyze due to the tenets of the framework. There were two central problems for DM: 1) in DM the structures *lice-infested* and **rats-infested* are identical with the exception of the application of a readjustment rule; and 2) DM recognizes no difference between inflectional and derivational morphology (and derivational morphology is allowed in such structures). The analysis shows here that the structures of **rats-infested* and *mice-infested* are not identical: one involves more fusion driven by MINIMIZE EXPONENCE. The grammaticality of the constructions reduces to the embeddedness of the feature [n] (which meant that derivational morphology is allowed as long as a nominalizer was the dominant morpheme in the non-head member).

Taken as a whole, this chapter does three things. It showed an alternative analysis for root allomorphy. It proposed refinements to DM. It showed that that an analysis of nominal compounds is possible in DM.

CHAPTER 4: SUBCATEGORIZATION WITH MINIMIZE EXPONENCE

Chapter 3 discusses the effects the constraint MINIMIZE EXPONENCE has on an analysis of root allomorphy and offers a novel analysis for the unique behavior of inflection in English nominal compounds. Chapter 4 extends the survey of the effects of MINIMIZE EXPONENCE from the realm of morphological phenomena to the realm of syntax. In this chapter, I explore an analysis of subcategorization that results from the claim that roots fuse with functional heads merged above them in order to best satisfy MINIMIZE EXPONENCE.

Chapter 4 is organized as follows: Section 1 offers a brief survey of the phenomenon of subcategorization (or argument selection) and a description of the only extant treatment of subcategorization within the framework of DM—Harley and Noyer (2000). Section 2 provides a sketch of the effects fusion has on an analysis of the argument selection of a verb. Sections 3 through 6 detail some of the predictions about the behavior of verbs that a model based on the MINIMIZE EXPONENCE makes including the polysemy of verbs, structural coercion, and dative alternations.

4.1. Subcategorization

A tremendous amount of work has been done in syntactic theory on the very simple observation that many verbs tend to appear with a variety of different argument structures. (Grimshaw 1979, Chomsky 1981, Borer 1994, 2004, Pesetsky 1995, Ramchand 1998, Baker 1988, and others). The aim of this work is to capture a number of different behaviors such as polysemy (a single verb is found in different syntactic and semantic environments).

(4.1). a. Harry kept the bird in the cage (*keep=detain)*
b. Susan kept the money (*keep=retain*)
c. Sam kept the crowd happy (*keep=maintain*) .
d. Let's keep the trip on a Saturday (*keep=maintain*)
(Jackendoff 1996, definitions for *keep* are mine)

e. I ran.
f. I ran the dog.
g. I ran the dog to Julie
h. The water ran.
i. The water ran into the sink

Above, in examples (4.1.a-d), the verb *keep* appears in the same syntactic configuration in all four examples, but its meaning changes to account for the different meanings of its object. In older lexical syntactic theory, what meanings that a verb selected for and how that selection affected the meaning of the verb was called *S-selection* (semantic selection, Grimshaw 1979). In DM, this specification is for the root √KEEP and is analyzed as extra-linguistic information found in the Encyclopedia.

However, in examples (e-i) we see the same verb appear in different *syntactic* configurations. The verb *run* appears with only an agent (e), an agent and a patient (f), an agent, a patient, and a location (g), only a patient (h), and a patient and a location (f). Explaining this variability in argument structure has been a theoretical goal of several different proposals (see Woolford 1984, Pesetsky 1995 for example).

A straightforward solution to the type of behavior seen in *run* and *keep* (4.1) is to stipulate that the argument-selecting behavior of verbs is variable or there are multiple instances of one verb, each with different selectional requirements. However, for many verbs, certain semantic or syntactic frames are not permitted.

(4.2) a. John put the paper on the shelf.
b. *John put on the shelf
c. *John put the paper.
d. *John put. (Pesetsky 1995)

e. The ship arrived.
f. *The captain arrived the ship.

g. I wondered whether the fate of the world was in my hands.
h. *I wondered the fate of the world.

In (4.2) we see a number of verbs that block certain structures. *Put* strictly requires an agent, a patient, and a location. Any syntactic configuration with fewer than those arguments is ungrammatical. Opposite to *put*, which requires a minimum number of arguments, *arrive* blocks an "extra" argument. (4.2.f) is ungrammatical because of the presence of an "extra" agent argument.

Some verbs seem to be sensitive to the category of their arguments, not only their presence. (4.2.g, h) show that *wonder* is sensitive to the syntactic category of its objects—it requires a CP. That some verbs are sensitive to the category of their arguments is the part of subcategorization called *C-selection* (category selection, Grimshaw 1979). This phenomenon is loosely called *subcategorization* (Grimshaw 1979), which is said to contain S-Selection and C-selection as well as idiosyncratic argument structure.

To capture this behavior, work in syntactic theory has claimed that a verb is lexically specified in some way for the environments that it is licensed to appear with (Grimshaw 1979 and following). The scope of this type of analysis is not limited to a simple licensing of a verb in some contexts but not in others. Theories of licensing also must capture frame alternations such as the double object alternation (or dative shift).

(4.3) a. Julie gave Ripley a bone.
b. Julie gave a bone to Ripley.

c. Julie delivered the scratching post to Gimli.
d. *Julie delivered Gimli the scratching post.
e. Jack asked Julie a question
f. *Jack asked a question to Julie.

Compounding studies of argument requirements of particular verbs is the phenomenon of *structural coercion* (Gleitman 1990, Lidz 1998). Structural coercion is the phenomenon whereby the specific argument requirements (and the corresponding semantic and categorical restrictions) of a verb are violated, but the violation of those requirements does not result in ungrammaticality. Rather, we are able to interpret the utterance.

(4.4) a. The giant ham sandwich ate the dog (cf. *The ham sandwich over-tipped*[24])
b. John thought the book to Mary
c. Fast-forward eats the tape.
d. The aliens glassed the planet
e. The bridge exploded the engineers.

The sentences in 4.4 are all grammatical. However, they are "odd" in one of two different senses. Examples (a, c, and e) are semantically odd. All of them involve a

[24] Notice that the interpretation of *the giant ham sandwich ate the dog* literally involves a sandwich capable of eating a dog while *the ham sandwich over-tipped* only involves a person who ordered a sandwich.

violation of the s-selection of the verb. In (a), we are forced to imagine a giant dog-eating ham sandwich in order to make it compatible with its role as the agent of *eat*. On the other hand, in (c) we change the meaning of *eat* to something like "ruin through catching in gears" in order to accommodate *fast-forward* as its agent. In the realm of Minimalist syntax, these remain a problem. For DM however, these effects are extra-linguistic and are "coerced" in the encyclopedia.

Examples (b and d) are crucially different. Both examples involve a verb appearing with extra arguments (in the case of *glass,* appearing with any arguments at all). Despite the extra arguments, both are interpretable. *John thought the book to Mary* seems to mean that either he used telekinesis to move the book to Mary or he memorized the book and used telepathy to transfer its contents. *The aliens glassed the plant* means the aliens bombarded the planet so thoroughly that the surface of the planet was turned to glass.

In the realm of construction grammar (Lakoff 1977, Langacker 1987, Goldberg 1995) this effect is attributed to the fact that syntactic constructions themselves contain meaning and that structural coercion is just a matter of concatenation of the meaning of the verb, the arguments, and the structure. In Minimalist syntax, the problem is more perplexing because the structure itself carries no meaning. The coercion must result from the presence of interpretable features. In Lexicalist Minimalism, the phenomenon is truly troubling since a violation of a Lexical entry's requirements should result in a crash of the derivation due to the incompatibility of features in the syntax with those that need to be checked in the verb.

Most of the work that has been done in this area of syntax and lexical semantics has been done within Lexicalist models of syntax (Grimshaw 1979, Woolford 1984, and others). As a result, much of this behavior has been captured by specifying a verb for the types and number of arguments that it takes in its lexical entry and then allowing for productive alterations of these specifications in a generative lexical component of the grammar. In a model of syntax without a generative lexical component that precedes the syntax, such as DM, this has to be captured differently.

4.1.1 Subcategorization within DM

The Lexicalist type of analysis is incompatible with late-insertion models such as DM because the concept of the lexicon in DM is different from that of Lexicalism. In DM, the lexicon as a generative mechanism does not exist, *per se*. Rather, the work of the lexicon is distributed among many different modules within the syntax. Crucially, in Lexicalist theories a verb can specify what types of elements it can be merged with and then project the structure it appears in, since fully formed words are the atoms that are manipulated by the syntax. However, since words are inserted into a fully formed structure in DM, the words themselves cannot dictate the structure of the sentence. Rather the word must be licensed for insertion into some derivations but banned from insertion into others.

Harley and Noyer (2000) provide an analysis to capture the licensing of a VI in a set of possible structures. In their analysis, the insertion of a VI is licensed via the secondary exponence features of that VI. Harley and Noyer (2000) propose licensing (discussed above in chapters 2 and 3) whereby particular verbs are licensed for insertion by the features of elements that c-command them or that they c-command.

(4.5)	Phonology	Licensing environment
	a. *sink*	[±v], [+DP], [±cause]
	b. *big*	[–v], [+DP]
	c. *open*	[±v], [+DP], [±cause]
	d. *destroy*	[+v], [+DP], [+cause]
	e. *arrive*	[+v], [+DP], [–cause]
	f. *grow*	[+v], [+DP], [±cause]

For example, the VI for *sink* is specified for the following: a) it may or may not be c-commanded by a little-v head (*the sink* vs. *The boat sank*) and that little-v head may or may not be causal (*The boat sank* vs. *I sank the boat*). Its object must be a DP. On the other hand, *destroy* must be c-commanded by a causal little-v (**The city destroyed* vs. *The barbarians destroyed the city*) and its object must be a DP, while *arrive* must be c-commanded by a non-causal little-v (**The captain arrived the ship* vs. *The ship arrived*). Harley and Noyer's analysis crucially relies upon secondary licensing in order to account for subcategorization.

However, as shown above in chapter 3, the model of DM I propose uses fusion and the MINIMIZE EXPONENCE constraint to eliminate the need for secondary licensing to account for irregular morphology. Those same features of the proposed grammar eliminate the possibility of using secondary licensing in this case as well since MINIMIZE EXPONENCE will trigger the fusion of c-commanding elements with the root. Thus, another explanation of subcategorization within DM is necessary. Section 4.2 details my proposal for capturing subcategorization within DM making use of the fusion that is driven by the economy constraint MINIMIZE EXPONENCE.

4.2 MINIMIZE EXPONENCE Based Account of Subcategorization.

As we saw in Chapter 3, employing MINIMIZE EXPONENCE eliminates the use of secondary licensing in order to capture a VI's lexical class because the root fuses with the c-commanding functional heads in order to best satisfy the constraint. Since in any given derivation, fusion of roots with the functional heads projected above them allows the roots to gain the features of the functional those heads, corresponding Vocabulary Items may be specified for those features gained through fusion.

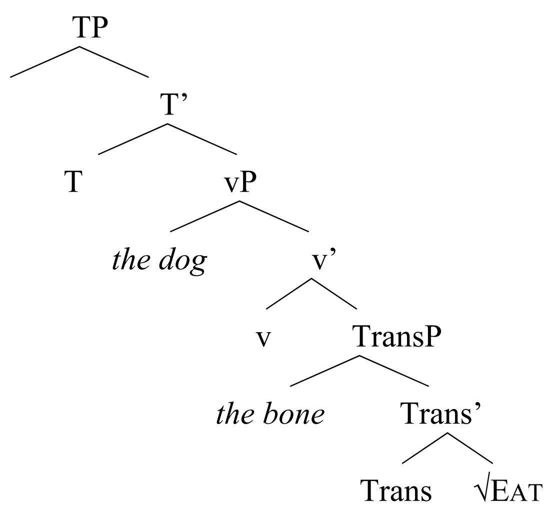
Following the claim that every argument is introduced to the syntax by a particular functional head, the fusion analysis easily extends to include subcategorization, allowing for an analysis that is not dependent on secondary licensing. This new analysis shows another application of the fusion driven by MINIMIZE EXPONENCE.

Marantz (1984), Larson (1995), Kratzer (1994, 1996) and many others have argued that agent arguments are not selected for by verbs. As I have assumed in my trees to this point, the agent argument is selected by a functional head called "light verb" (little-v) or VOICE that is generated above VP (Kratzer 1994). Further work in event structure has argued the light verb head to be the locus of the meaning CAUSE (see Kratzer 1994, 1996, Ramchand 1997, Harley 1995, and related).

Following the "severing" of the external argument from the verb, Borer (1994), Jelinek (1998), Ramchand (1997, 1998) and others have also endeavored to sever the theme argument from the verb as well. They have argued that the theme argument is also projected by a functional head. The name of this head varies from author to author. For the sake of clarity and ease of presentation, I will adopt Jelinek's (1998) name for that head, TRANS, so that it will easily be differentiated from the other verbal head, little-v.

Since this functional head takes the root as its sister, the theme argument must be projected in specifier position. Under such an analysis, the structure of a transitive sentence looks like (4.6) (shown before movement):

(4.6) The dog ate the bone.



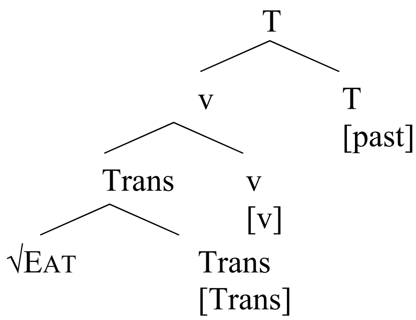
As a matter of parsimony, we can also assume that functional heads that are merged above the root (Pesetsky 1995) project all arguments including instruments, goals, experiencers, locatives, etc. I assume this throughout this dissertation. For the purposes of this dissertation, I give simple names to the argument heads that I show. The head that projects a dative object I call G and the head that projects locatives I call L.

In Chapter 3, I propose that all functional heads carry formal features that are identifiable for competition. Here, I extend that proposal to include the functional heads that project arguments. The features (or feature bundles) carried by these heads bear a meaning that roughly corresponds with the theta role of the argument that the head licenses. For example, the type of "v" that licenses agent arguments carries the feature [v], which has the meaning CAUSE (Kratzer 1994, Ramchand 1997, 1998). The head that projects themes, Trans, carries the feature [Trans].[25]

Like all other features, these "argument features" move up the tree through normal head movement. Recall that head movement (an instantiation of morphological merger) will cause the root to move up the tree and collect the c-commanding formal features into a complex head. In the example above (4.6), √EAT moves to v through Trans. As it moves, the features of each head √EAT attaches to are added to the complex head structure through head adjunction. Again, we assume a Bobaljik-style "merger under adjacency" analysis of the marking of tense on verbs in English (see Bobaljik 1994). The combined applications of merger under adjacency and head adjunction result in the complex head show in 4.7.

[25] At this point in this dissertation, I am treating [Trans] as one feature. However, it is likely that these are actually a bundle of features. For now, the difference is moot. I will return to this distinction below when it becomes important.

(4.7) Complex head created by derivation in (4.6).



Again, as shown in Chapter 3, the process of fusion applies to the complex head resulting in the simplex head in (4.8), which carries the formal features [v], [Trans], and [past] as well as the root.

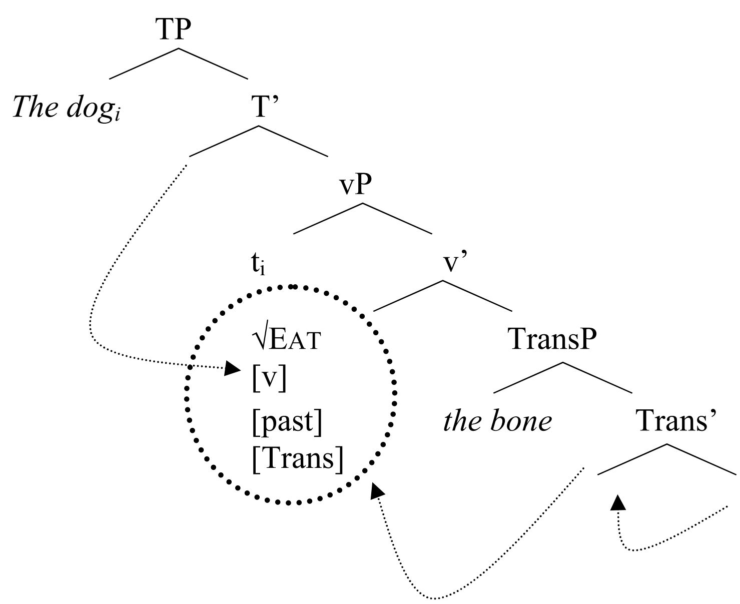
(4.8) √EAT
[v]
[V]
[past]

This simplex node is now a candidate for vocabulary insertion and the discharge of its features by just one VI rather than the four that would have been necessary to discharge the features of the four different heads in the complex structure. Thus, root VIs can be specified for these features just as they can be specified for any other features. The features of a root VI can be determined by examining the possible and impossible environments for that word and finding feature specification that is compatible with the possible environments and incompatible with the impossible ones. As an example, below I show how we can arrive at the feature specifications for *ate*.

4.2.1 Determining the Feature Specification of a Vocabulary Item

To show how to determine the specifications of a VI, I continue with the example of *ate*. Since the VI *ate* can be inserted into the node created in (4.6) above, *The dog ate the bone,* we can conclude from *ate*'s insertion in this context that it is specified for some subset of the features contained in that node, up to and including all of them. If it had any more features, insertion would be blocked. In (4.9.a), I show a derivation of *The dog ate the bone* that results in a fused node containing the features [v], [past], and [Trans]. In (4.9.b), I show all the possible specifications of *ate* and how we can determine the specifications for *ate* since we know it is allowed in the node created in (4.9.a). Since we know that *ate* is allowed in such a derivation, we know that the specifications in (4.9.c) are the maximum specifications for a VI for *ate* (an overspecified candidate would be disallowed).

(4.9) a. *The dog ate the bone.*



b. Possible specifications for *ate*

Node that *ate* is allowed in:	Specification	Vocabulary Item
√EAT [past] [v] [Trans]	**not a possible specification (*overspecified*)** ◄------	*ate:* √EAT [v] [past] [Trans] [G]
	maximum possible specification ◄------	*ate:* √EAT [v] [past] [Trans]
	possible specification ◄------	*ate:* √EAT [v] [past]
	not a possible specification (*conflicting feature*) ◄------	*eat:* √EAT [v] ¬ [past]

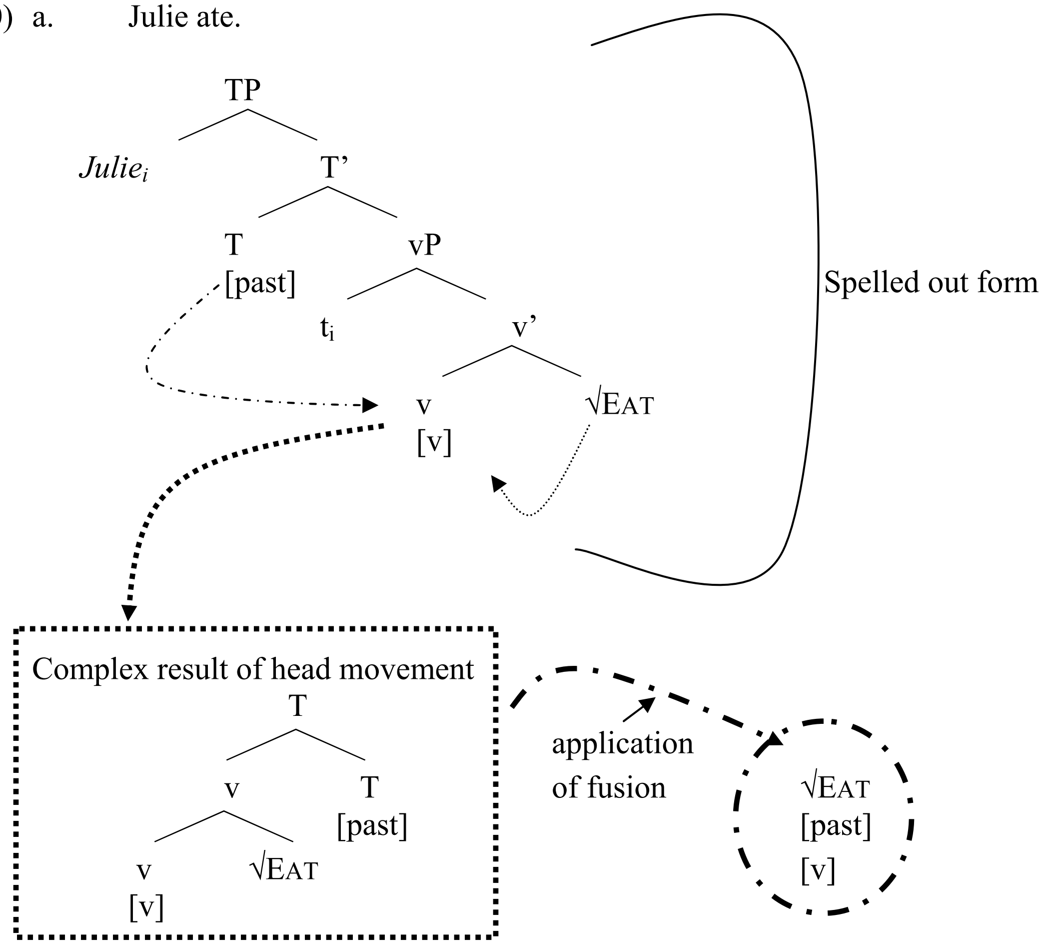
c. Maximum specification for Vocabulary Item for *ate*

√EAT [v] [Trans] [past]	→	*ate* /ejt/

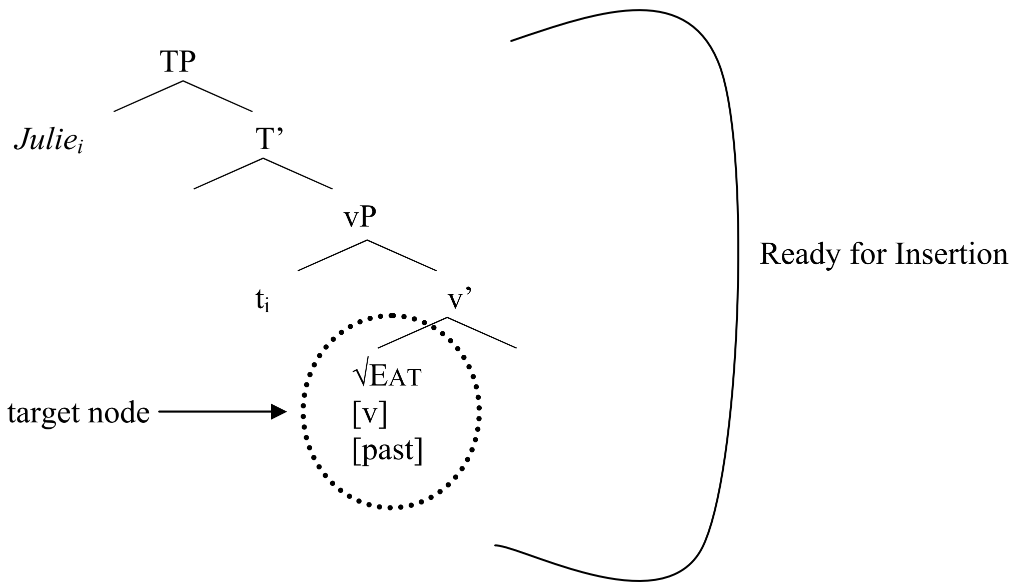
Just examining the maximum features necessary for insertion into *The dog ate the bone* is not sufficient to determine the specification of *ate.* Since this is an underspecification model of subcategorization, a VI that is underspecified for one or both of [Trans] and [v] are still possible VIs (we know that [past] is required to distinguish *ate* from *eat*).

We can precisely determine the specification for *ate* by examining other environments it can appear in, as well as environments in which it loses the competition to another VI. *Ate* can appear in a derivation that does not have an object such as in sentences like *Julie ate.* If we assume a structure for *Julie ate* where the object is not base-generated rather than an elision account of the missing object, its structure is (4.10).

(4.10) a. Julie ate.

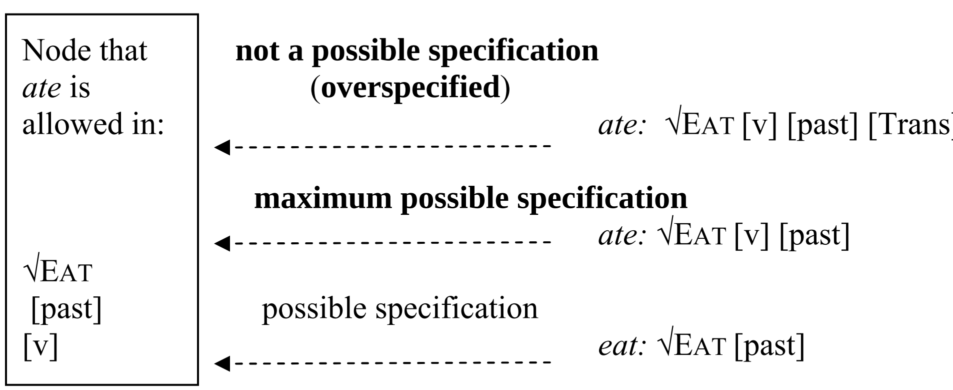


b. Resulting form



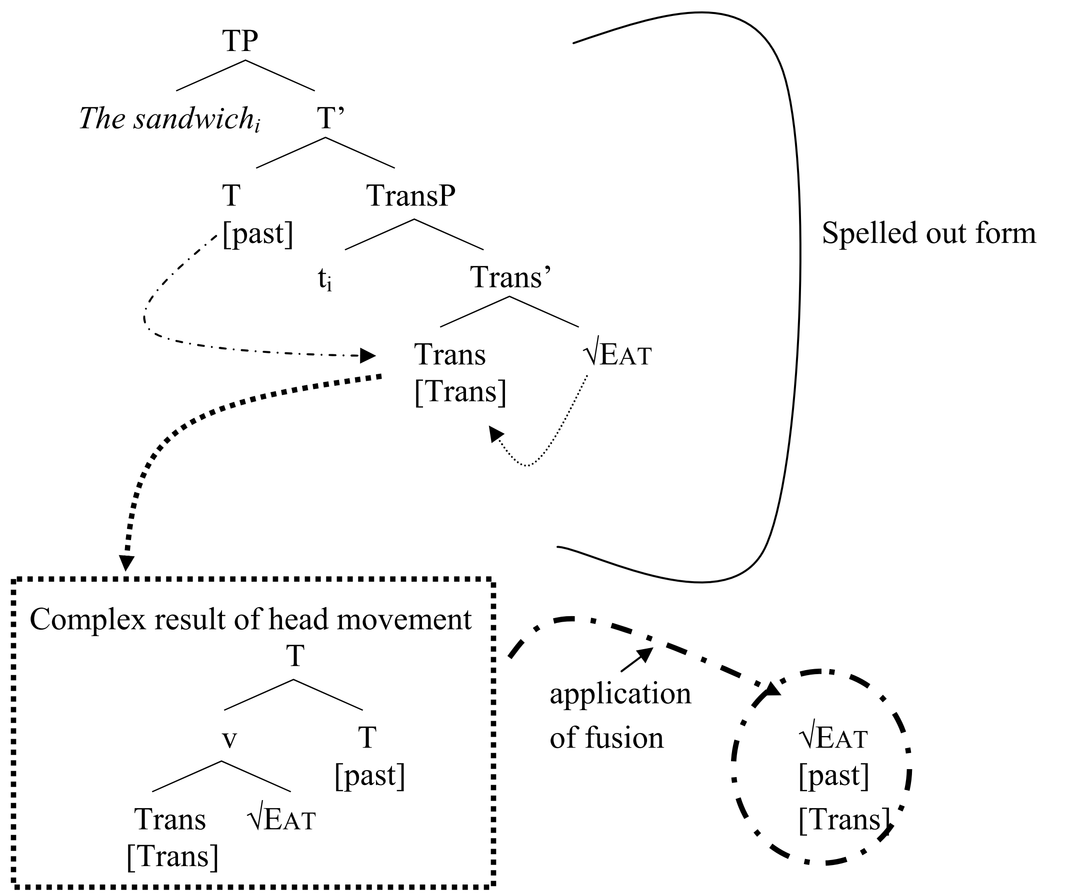
In the structure in (4.10), there is no [Trans] feature to be fused to √EAT. It is clear then, since *ate* can still be inserted into the node that the VI for *ate* must not be specified for [Trans].

(4.11) Possible specifications for *ate*

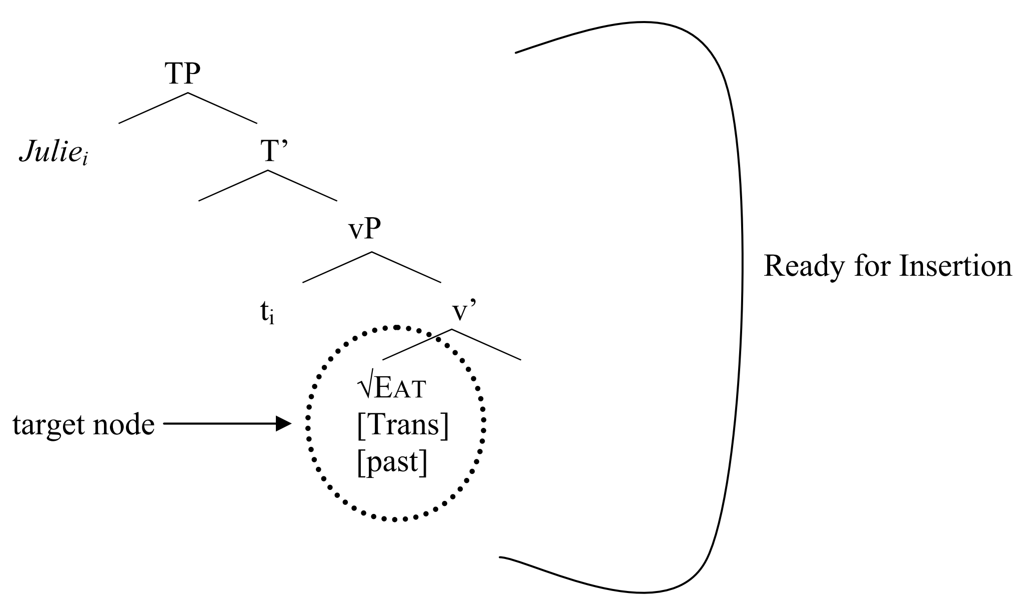


However, since *ate* cannot appear as an unaccusative or a zero-derived nominal, the VI for *ate* must be specified with the feature [v].

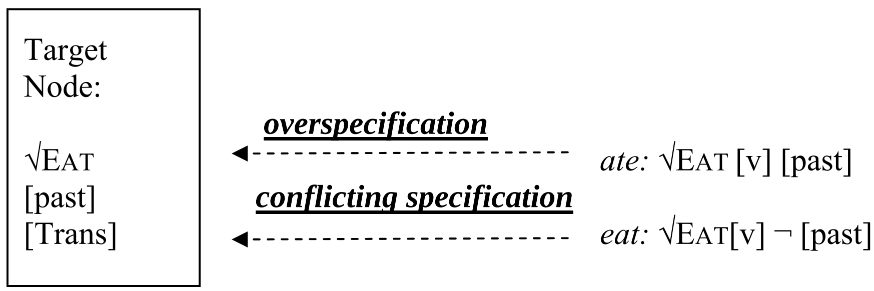
(4.12) a. **The sandwich ate*. (where the sandwich is devoured)



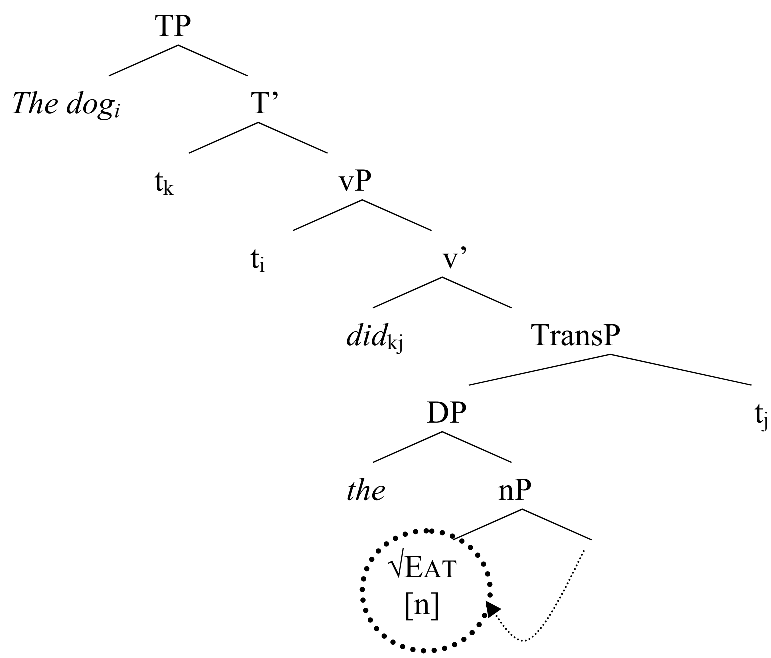
b. Resulting form



c. Competition (no winning candidate...both *eat* and *ate* require [v])



(4.13) a. **The dog did the ate..*



b. Competition (no winning candidate)

Target Node:	***overspecification*** ←-------------------	*ate:* √Eat [v] [past]
√Eat [n]	***overspecification*** ←-------------------	*eat:* √Eat[v] ¬ [past]

The competitions in (4.12.c) and (4.13.b) show us that in order to capture the ungrammaticality of **The sandwich ate* and **The dog did the ate, ate* must be specified for the [v] feature. Thus the relevant features that the VI *ate* is minimally specified for are [past] and [v], which can be interpreted as the VI requiring that it a) realize the root √Eat, b) have an agent, and c) be past tense. Since this is the same as the maximum specification we determined to be possible in (4.1), *Julie ate*, we can conclude that (4.14) is the specification for *ate*. Conveniently, this means that we arrive at the same VI we assumed in Chapter 3.

(4.14) Vocabulary Item for *ate*

√Eat	→	*ate*
[v]		/ejt/
[past]		

The VI in (4.14) predicts that *ate* will be able to appear in any environment as long as the utterance is in the past tense and as long as there is an agent argument. The VI is indifferent to whether there is a theme present.

4.2.2 Differences among Subcategorization Models

The predictions above are different from the predictions made by Lexicalist theories of subcategorization such as those proposed by Grimshaw (1979), Chomsky (1981), Pesetsky (1995), Ramchand (1997,1998), Baker (1988), and others in Lexicalist models. This model proposes that there is only one instantiation of a VI and that that VI will appear in a multitude of different environments as long as those environments include at least the minimum number of arguments. In this way, this is an underspecification analysis. Lexicalist analyses tend to include lexical entries that are fully specified for their arguments or that contain optional elements (see Woolford's 1984 analysis).

However, within the realm of DM, the Harley & Noyer (2000) analysis (repeated as 4.15) can easily be made to make underspecification predictions in the same way that this analysis does if the optional elements (such as [±v]) are instead treated as unspecified (4.16).

(4.15) Harley and Noyer (2000) licensing

Phonology	Licensing environment		
a. *sink*	[±v]	[+DP]	[±cause]
b. *big*	[–v]	[+DP]	
c. *open*	[±v]	[+DP]	[±cause]
d. *destroy*	[+v]	[+DP]	[+cause]
e. *arrive*	[+v]	[+DP]	[–cause]
f. *grow*	[+v]	[+DP]	[±cause]

(4.16) Harley and Noyer (2000) licensing adjusted to an underspecification analysis.

Phonology	Licensing environment		
a. *sink*		[+DP]	
b. *big*	[–v]	[+DP]	N/A
c. *open*		[+DP]	
d. *destroy*	[+v]	[+DP]	[+cause]
e. *arrive*	[+v]	[+DP]	[–cause]
f. *grow*	[+v]	[+DP]	

The availability of an underspecification analysis is a benefit of the competition aspect of DM. The subset principle dictates that VIs are often inserted into environments with more features than the VIs are specified for. Thus, VIs are already underspecified for the features they realize in the sense that they discharge all the features in a node, not only the ones they are specified for.

The aspect of this analysis that sets it apart from the Harley and Noyer (2000) analysis is not the predictions that it makes, but rather the locality of the effect that licenses insertion of the VI. Since the Harley and Noyer (2000) analysis uses secondary licensing, the same concerns about licensing we discussed in Chapter 3 apply to the Harley and Noyer (2000) analysis. That is, the constraints on insertion are inherently non-local: The VI must not only check the node for compatibility but also the rest of the VP shell. In the fusion analysis, the VI only checks features in the node it is inserted into, a process that is more local and is thus more economical.

Having now sketched the fundamentals of the fusion-based account of thematic licensing of verbs, I dedicate the rest of Chapter 4 to the finer aspects of this analysis including what behaviors are predicted by this model.

4.2.3 Evidence for the Feature [v] in English

The proposal here rests on the assumption that functional heads that select for arguments contain the features such as [Trans] and [G] that I propose above. Since the complex heads containing these features undergo fusion, the features themselves are realized by the same VI that realizes the root. Therefore, the features that I propose here are often not realized by overt morphology. It can be argued that verbalizing morphemes such as –*ize* and -*ify*, realize that feature.

However, if the features such as [Trans] can be selected by Vocabulary Items, it follows that there should be minimal pairs that are differentiated by those features. Just as *eat* and *ate* are different only in that *ate* selects for [past] and *speech* and *speak* differ only in that *speech* selects for [n], there should be pairs that are different only in that they select for [v]. An example of such an alternation is the *rise*/*raise* alternation.

(4.17) a. My hand rose from the table.
b. I raised my hand off the table.

The form *rise* appears as an unergative while *raise* appears as a transitive verb. Under this analysis, the VI for *raise* must be specified for the [v] feature, requiring it to be inserted into a derivation with an agent argument. Since there is an alternation for transitivity here, this means that the VI for *rise* must not be specified for [v] (also evidenced by the fact that *rise* appears without an agent).

(4.18) Vocabulary Item for *rise*[26]

√RISE → *rise*
/rajz/

(4.19) Vocabulary Item for *raise*

√RISE → *raise*
[v] /rejz/
[Trans]

Both VIs are linked to √RISE. In this way, the alternation of *rise* and *raise* is treated as parallel to the alternation of *eat* and *ate*. Thus, the alternation of *rise* and *raise* can be taken to be evidence that the features I propose here do exist in the grammar of English—the difference between the two words is the presence of the feature [v] (and, coincidentally, [Trans]).

4.2.4 Blocking "Extra" Arguments

As I have set this proposal out so far, it is apparently too strong. It entails that any verb can appear with any number of arguments beyond its minimum, as shown by *run* in Section 2.2 of this chapter. However, there are verbs that are ungrammatical in the specific constructions where they appear with more arguments than its "default" conditions (4.20).

(4.20) a. The ship arrived.
b. *The captain arrived the ship.
c. Ripley fell down the stairs.
d. *Julie fell Ripley down the stairs.

---

[26] Note that *rise* is not actually specified for [Trans]. *Rise* not only appears as a zero-derived nominal, but it can also be argued that the intransitive forms it surfaces in can either be agentive (*I rise early from bed*) or non-agentive (*The X-wing rose from the swamp*).

In (4.20b), the "extra" agent argument is blocked for *arrive*. What is necessary, then, is to design a mechanism within the framework of Distributed Morphology to block the insertion of VIs into nodes that they otherwise qualify for. We return then to the ¬ specification proposed in chapter 3. By specifying a Vocabulary Item for incompatibility with a feature carried by one of the functional heads that licenses an argument, we ensure that that VI cannot be inserted into a node that has that particular root and that particular feature together. This feature blocking will prevent a derivation containing that root and that feature from being grammatical unless that feature is otherwise realized by another VI. Ultimately, this means that there is an incompatibility of the root the VI realizes and the blocked feature. Since the head is what licenses the merging of the argument, the argument and the VI will not co-occur unless there is another overt morpheme, such as a transitivizer, to license the argument.

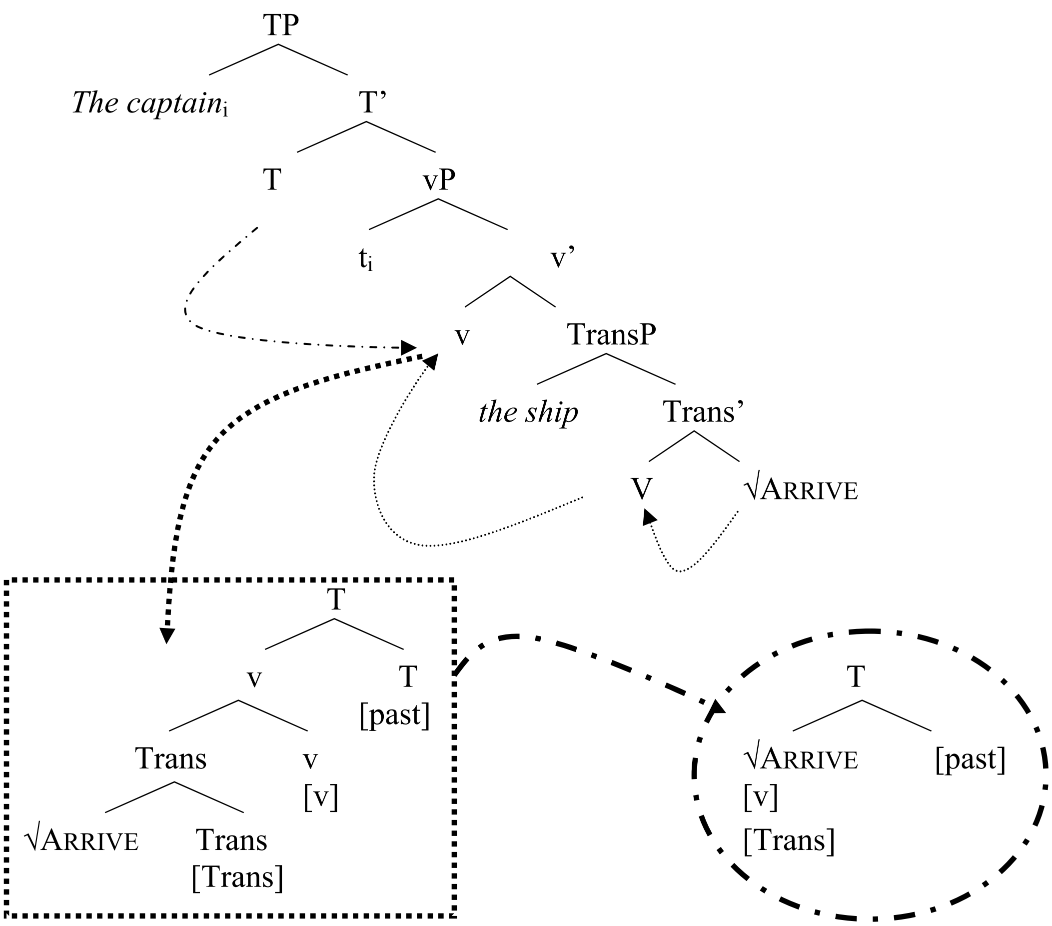
To see how this works, let's return to the example of *arrive*. *Arrive* is incompatible with an agent argument, as we see with the ungrammaticality of sentences like **The engineer arrived the train*. Such behavior is indicative of the VI for *arrive* being incompatible with [v], the feature from the little-v head which licenses the agent argument. The VI would then be (4.21, ignoring specifically morphological specification like ¬ [past]).

(4.21) Vocabulary Item for *arrive*

√ARRIVE	→	arrive
¬ [v]		/ərajv/
[Trans]		

A VI specified in this way is blocked from insertion into the derivation of **The captain arrived the ship* because the ¬ [v] it is specified for is not compatible with the [v] in the fused head it is inserted into in (4.22).

(4.22) a. *The captain arrived the ship.

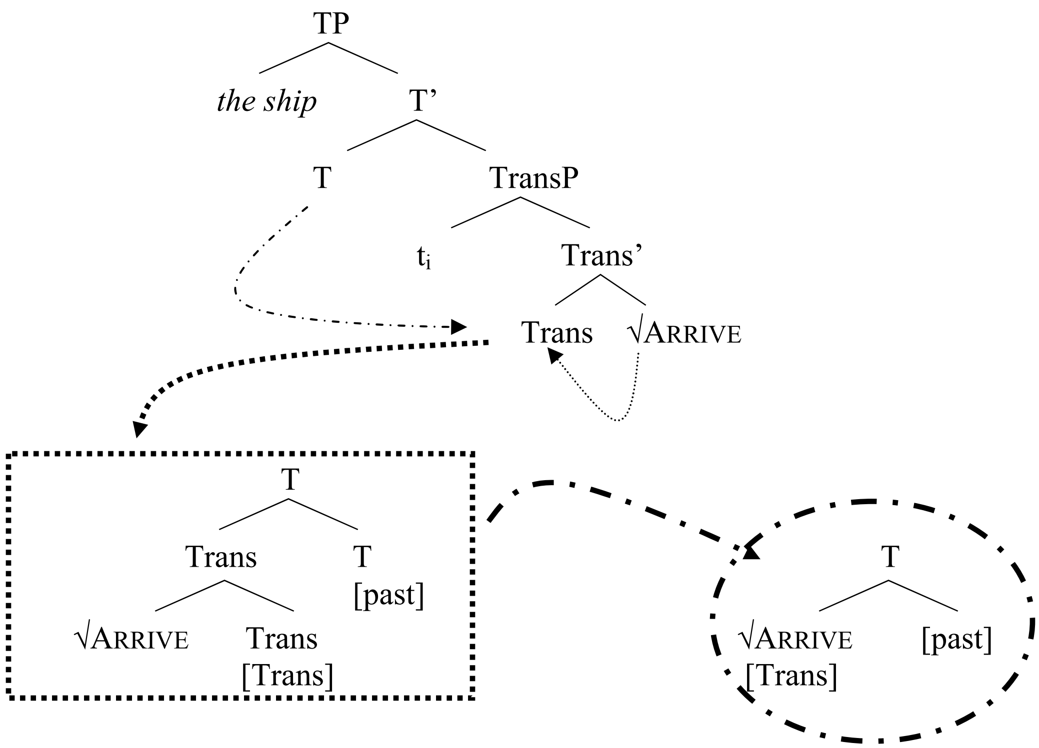


b. Competition (no winning candidate)

Target Node: √ARRIVE [Trans] [v]	***<u>conflicting specification</u>*** ◄------	*arrive:* √ARRIVE ¬ [v] [Trans]

Of course, the VI for *arrive* in (4.21) is compatible with a derivation that does not include a little v. Thus, *arrive* is easily inserted into a derivation such as *The ship arrived* (4.23).

(4.23) The ship arrived.



b. Competition

Target Node:

√ARRIVE
[Trans]

inserted candidate ←------------------ *arrive:* √ARRIVE ¬ [v] [Trans]

A VI specified in this way captures the grammatical incompatibility of *arrive* with the light verb. Crucially, there is no claim in this formalization that the meaning of *arrive* is incompatible with the meaning CAUSE. We know that the meaning of *arrive* is compatible with the meaning of CAUSE because things in the real world can certainly arrive with some sort of external causation. This type of blocking shows the ill-formedness of the otherwise meaningful **The captain arrived the ship* to be purely grammatical. In this way, the model proposed here presents an elegant account for the blocking of a verb into particular environments where there is an "extra" argument.

4.3 Rampant Polysemy

A strong prediction of this analysis is that a verb will be able to appear in any environment that has at least the minimum number and types of arguments that it is specified for and none that it bears a ¬ feature for. Since a VI such as *ate* is only specified for the features [v] and [past] it must be able to appear in any environment as long as an agent argument is present. This is the exact behavior this analysis sets forth to capture. In the paradigm in (4.24), we see the same root appear in a number of different semantic and thematic environments.

(4.24)	a.	I run.	e.	I run Mary the paper
	b.	I run the race.	f.	I run the paper to Mary.
	c.	I run the dog.	g.	John went on the run
	d.	The water runs.	h.	John built a run for the dog.

A lexicalist approach to this behavior would need to generate at least five different lexical entries for *run* based on syntactic structure alone. Acknowledging the fact that (4.24.b) and (4.24.c) have the same structure but the interpretations are slightly different, the number of lexical entries grows substantially. Excluding the possibility of a transformation within the lexicon for the double object construction or a Woolford (1984) style optional theta-grid, a strong lexicalist theory needs to generate eight different lexical entries to capture the above behavior, each with its own s-selection and linking.

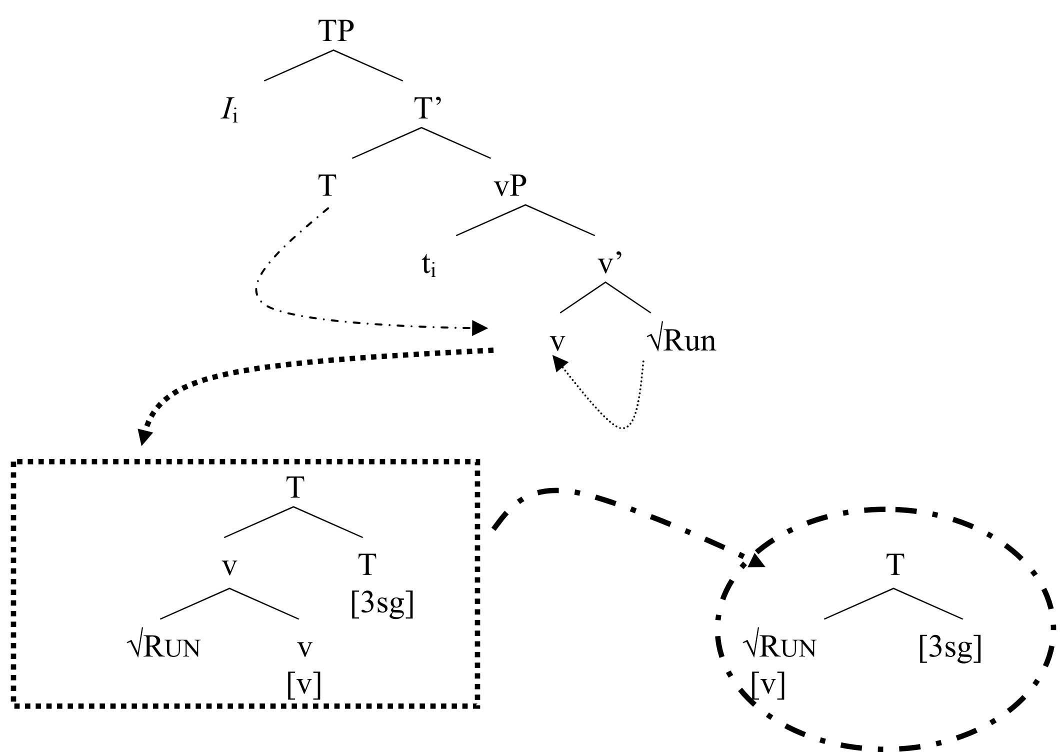
However, an underspecified analysis of this behavior as proposed here captures this behavior with just one VI.

(4.25) Vocabulary Item for *run*

√Run	→	run
¬ [3sg]	/rʌn/	

This VI can be inserted into any of the structures in (4.24) as its requirements—a root only—are a subset of the features of the fused nodes in all eight derivations, including the zero derived nominals, as illustrated in (4.26-28).

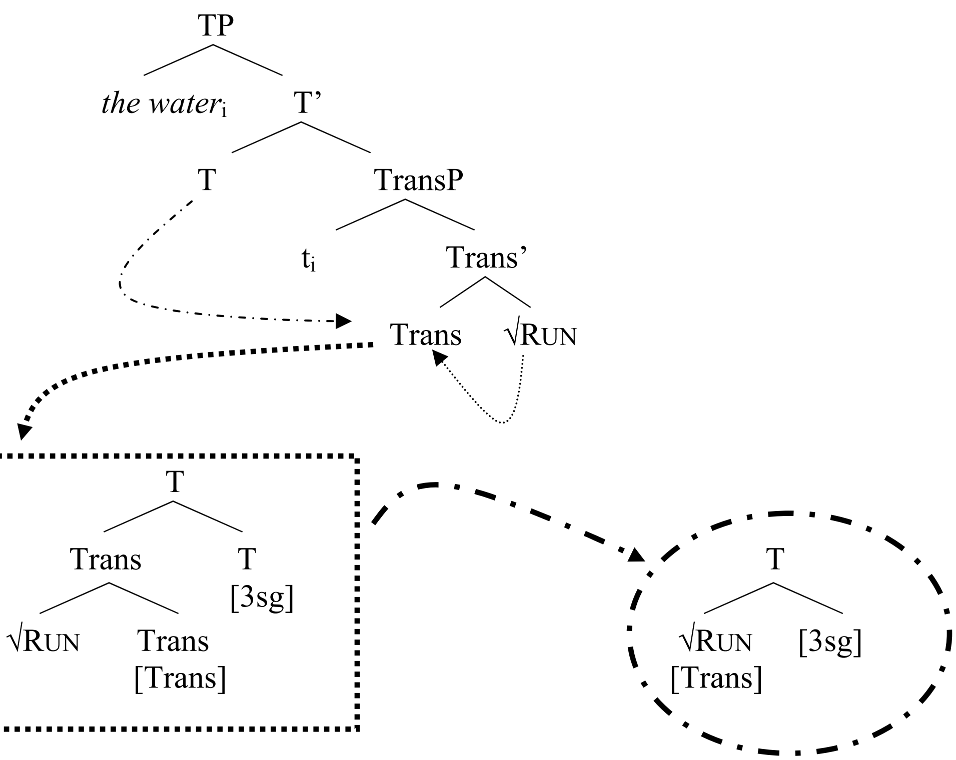
(4.26) I run.



b. Competition

Target Node: √RUN [v]	inserted candidate ←- - - - - - - -	*run:* √RUN ¬ [3sg]

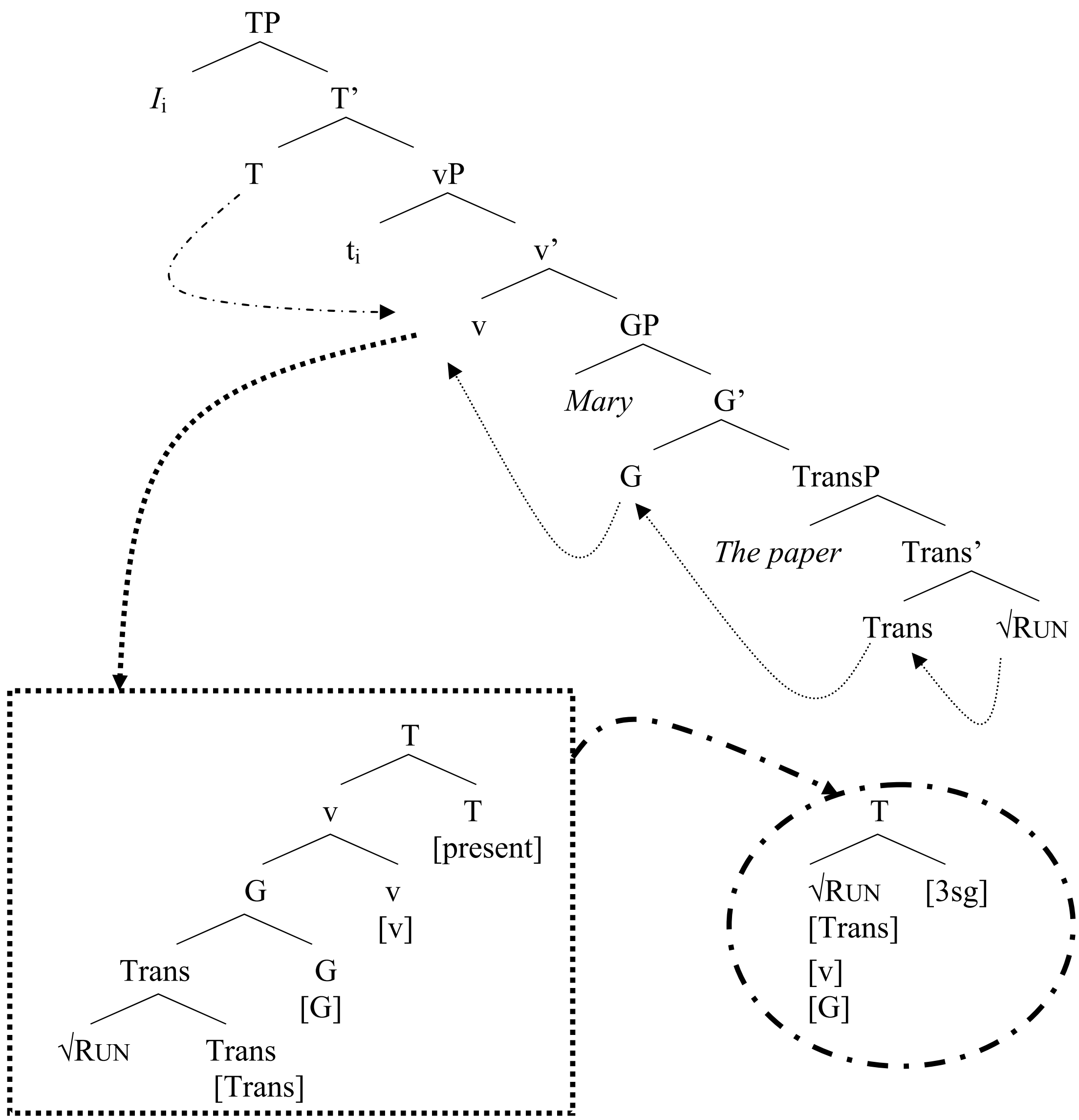
(4.27) The water runs.



b. Competition

Target Node: √RUN [Trans]	inserted candidate ←- - - - - - - -	*run:* √RUN ¬ [3sg]

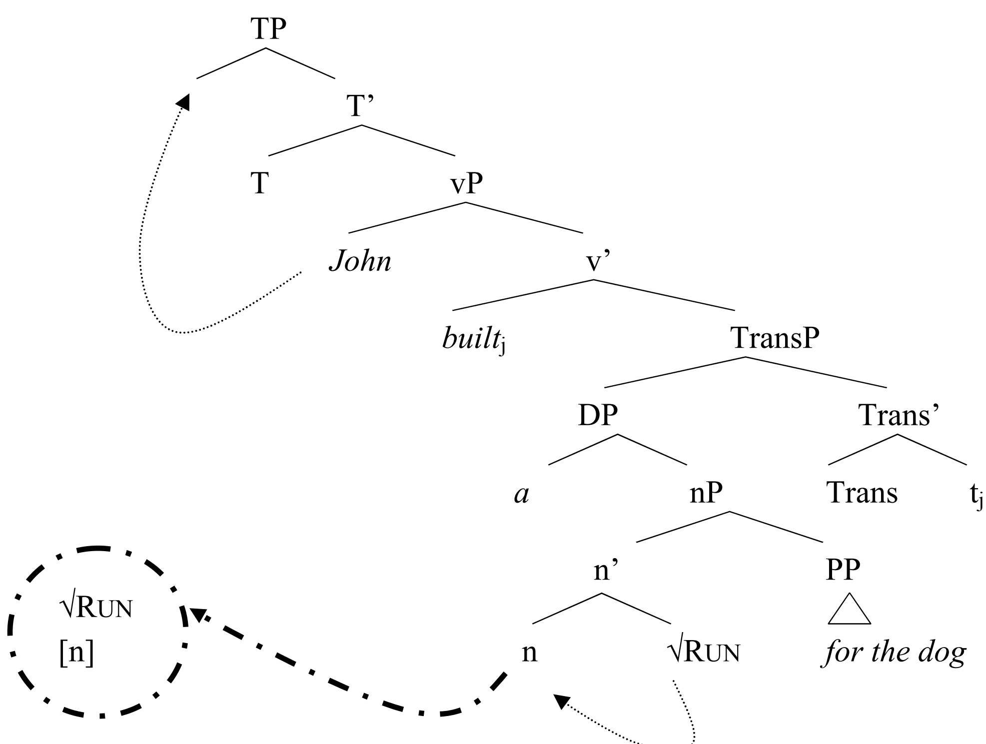
(4.28) I run Mary the paper.



b. Competition

Target Node: √RUN [Trans] [v] [G]	inserted candidate ◄- - - - - - - - - - - -	*run:* √RUN ¬ [3sg]

(4.29) John built a run for the dog.



b. Competition

Target Node: √RUN [n]	inserted candidate ◄- - - - - - - - - - - -	*run:* √RUN ¬ [3sg]

Lexicalist theories are often dependent on rigid s-selectional or thematic properties of lexical entries. When a verb behaves the way that *run* does above, a lexical theory of this sort is forced into the precarious prediction: They are committed to the existence of a vast number of lexical entries which are mostly synonymous (i.e. linked to the same concept) and completely homophonous. These "homonyms" differ only by the syntactic environments they are licensed for. The underspecification analysis of a verb's licensing conditions that is presented in this paper captures why verbs appear to be polysemous without asserting a lexicon full of homonyms.

4.4 Structural Coercion

The crux of this proposal is that roots are underspecified for the environment in which they can appear. It follows that the grammar will produced structures in which a root appears with more arguments than it normally does. A root may even appear in an environment that our real world knowledge is not compatible with. For example, the utterance #*The ham sandwich ate the dog* is produced grammatically by the syntax. (see syntactic bootstrapping literature: Lidz 1998, Naigles 1993, Gleitman 1990; see also Lidz and Gleitman 2004 for rebuttal) The Encyclopedia marks this sentence as ill-formed because ham sandwiches do not make good eaters. However, we can adjust our real world knowledge to fit the sentence by interpreting *ham sandwich* in such a way that a ham sandwich is somehow capable of eating a dog. We can also adjust the meaning of *eat* rather than the meaning of its arguments. This is of course how we get sentences such as *Fast forward eats the tape*

A verb that appears with "extra" arguments that are "incompatible" with our real world knowledge forces us to coerce our real world knowledge to fit the utterance, especially if, as listeners, we are to assume that the speaker strictly adheres to Grice's maxims. In this way, when we are presented a sentence such as *John thought the book to Mary*, we are able to interpret it.

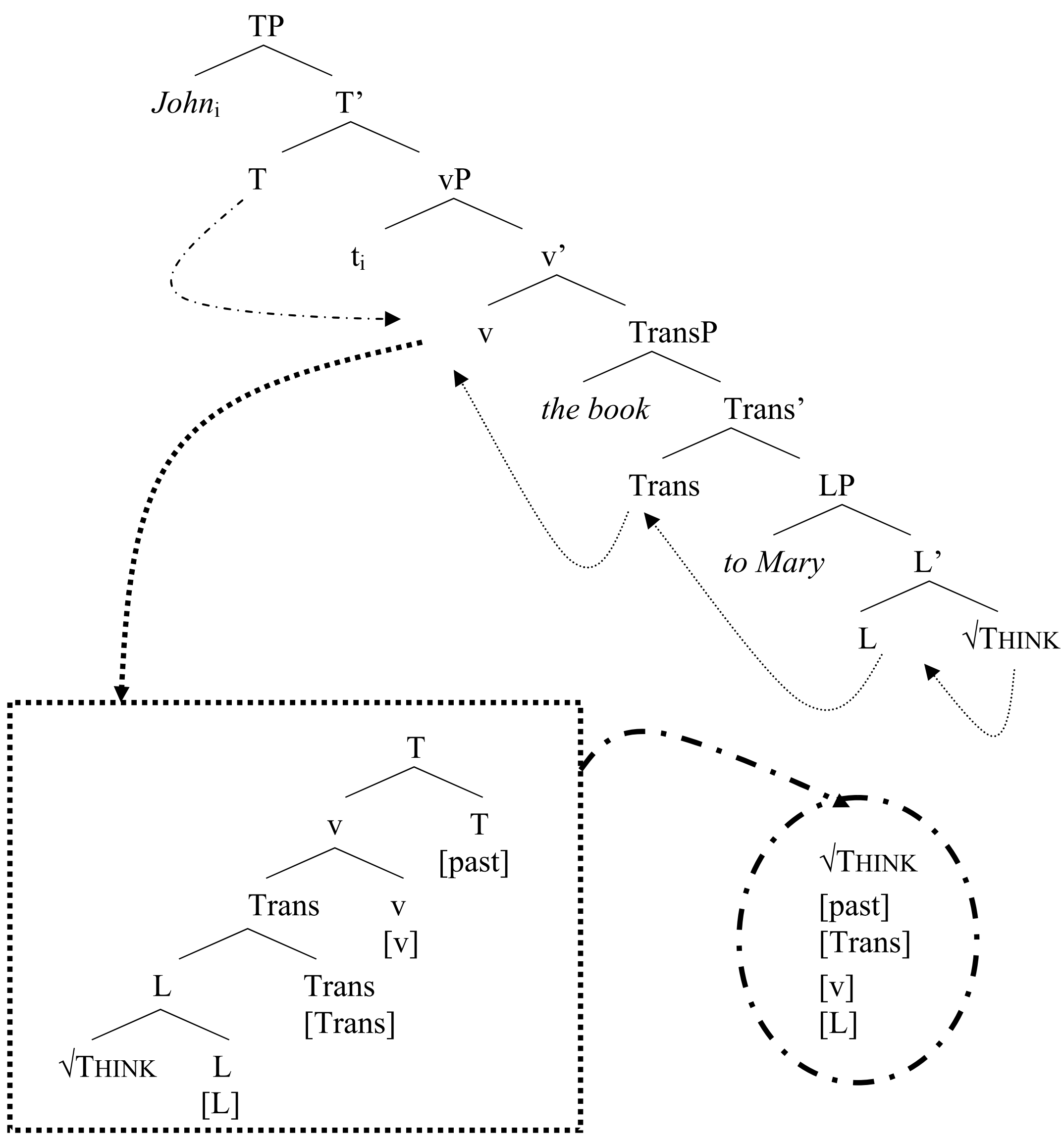
Not only does the framework proposed here account for structural coercion, it predicts it. For example, since *thought* typically appears in unergative structures (or in CP complement transitive structures) we can preliminarily assume that the VI for *thought* is specified only for [v] (and the past tense).

(4.30) Vocabulary Item for *thought*

√THINK	→	*thought*
[v]		/θat/
[past]		

Since *thought* is so underspecified, it easily appears in an environment that is specified more fully for arguments, such as the environment in #*John thought the book to Mary.* In that environment, the fused node that the VI is inserted into (4.30) is a superset of the features the VI contains.

(4.31) John thought the book to Mary



Competition

Target Node: √THINK [past] [Trans] [v] [L]	← inserted candidate	*thought:* √THINK [past] [v]

4.5 Dative Alternations

One of the more difficult alternations that subcategorization theories have tried to capture is the dative alternation. Historically, the pattern seen in (4.32) was a classic example of a transformation where the structure with an oblique indirect argument was transformed into a double object construction with the same meaning (Chomsky 1965). More recent work (Oehrle 1976 and following) has shown the two different structures to be different derivations without a transformational connection. In the realm of work within subcategorization, the onus then fell on explaining why, especially if the sentences conveyed the same information, some verbs allow both structures while others prohibit one structure and allow the other (4.32).

(4.32) a. Julie gave Ripley a bone.
b. Julie gave a bone to Ripley.
c. Julie delivered the scratching post to Gimli.
d. *Julie delivered Gimli the scratching post.
e. Jack asked Julie a question
f. *Jack asked a question to Julie.

Oehrle (1976) observed that the two structures do not actually entail the same relationship between the arguments to the verb. Rather than a goal argument being present in both the "to dative" construction and the double object construction, the "to dative" contains a locative argument (instead of a goal argument) and a thus a locative head that projects it.

The fact that the "to dative" corresponds to location rather than goal arguments is exemplified by the word *sent* (4.33).

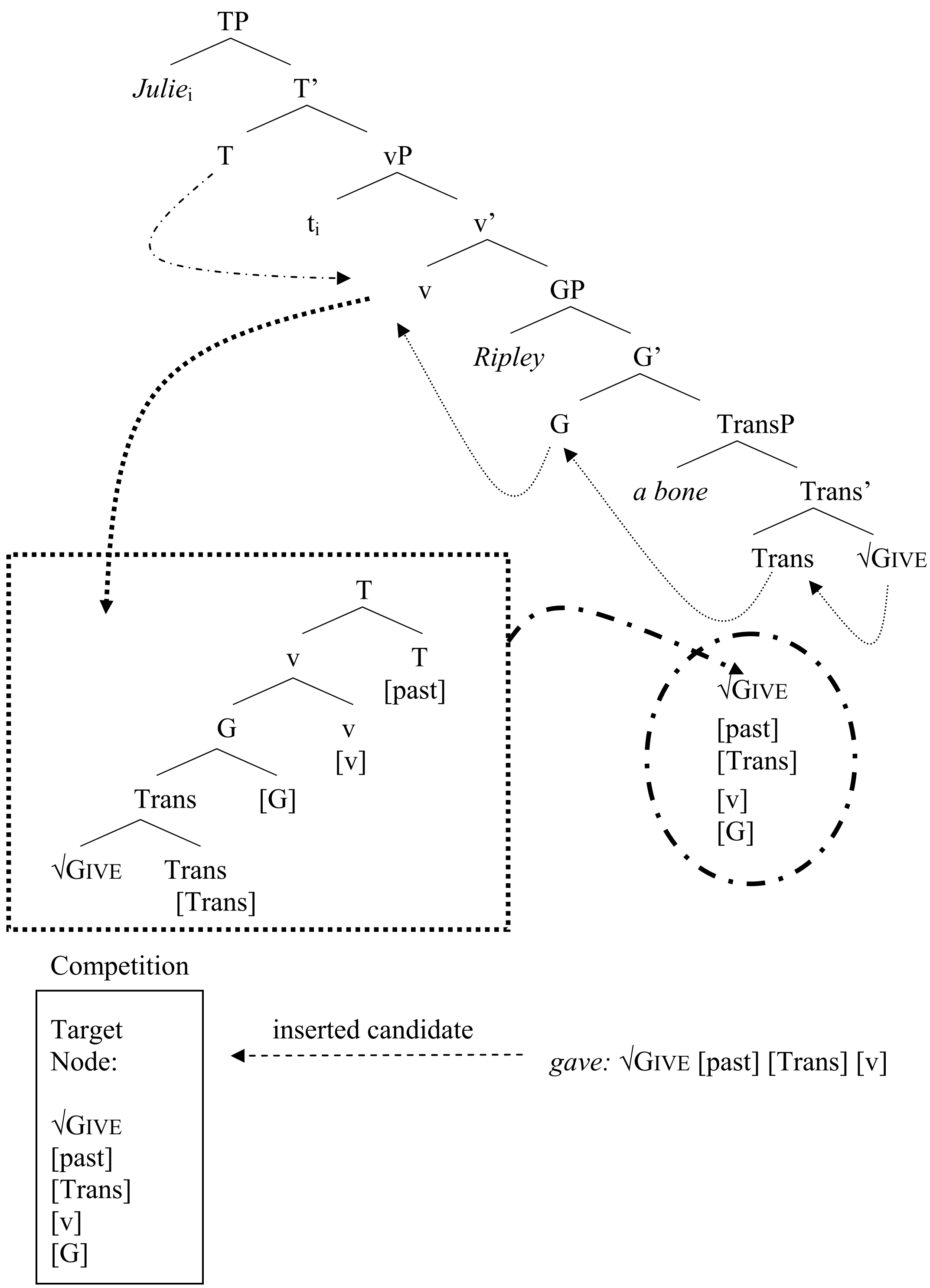
(4.33) a. Jack sent Julie a message.
b. Jack sent a message to Julie.

When a person is the DP in the locatum/goal alternation such as in (4.33), the interpretations of both sentences are so close that they give rise to the intuition that the sentences are related. However, if the locatum/goal DP in the alternation is a place rather than a person as seen in (4.34), it becomes clear that there are two different roles being associated with the positions because only one of the structures is permitted. This shows that they are not the same argument: a person is a "good" goal and a "good" location. A place, however, is not a good goal, only a good location.

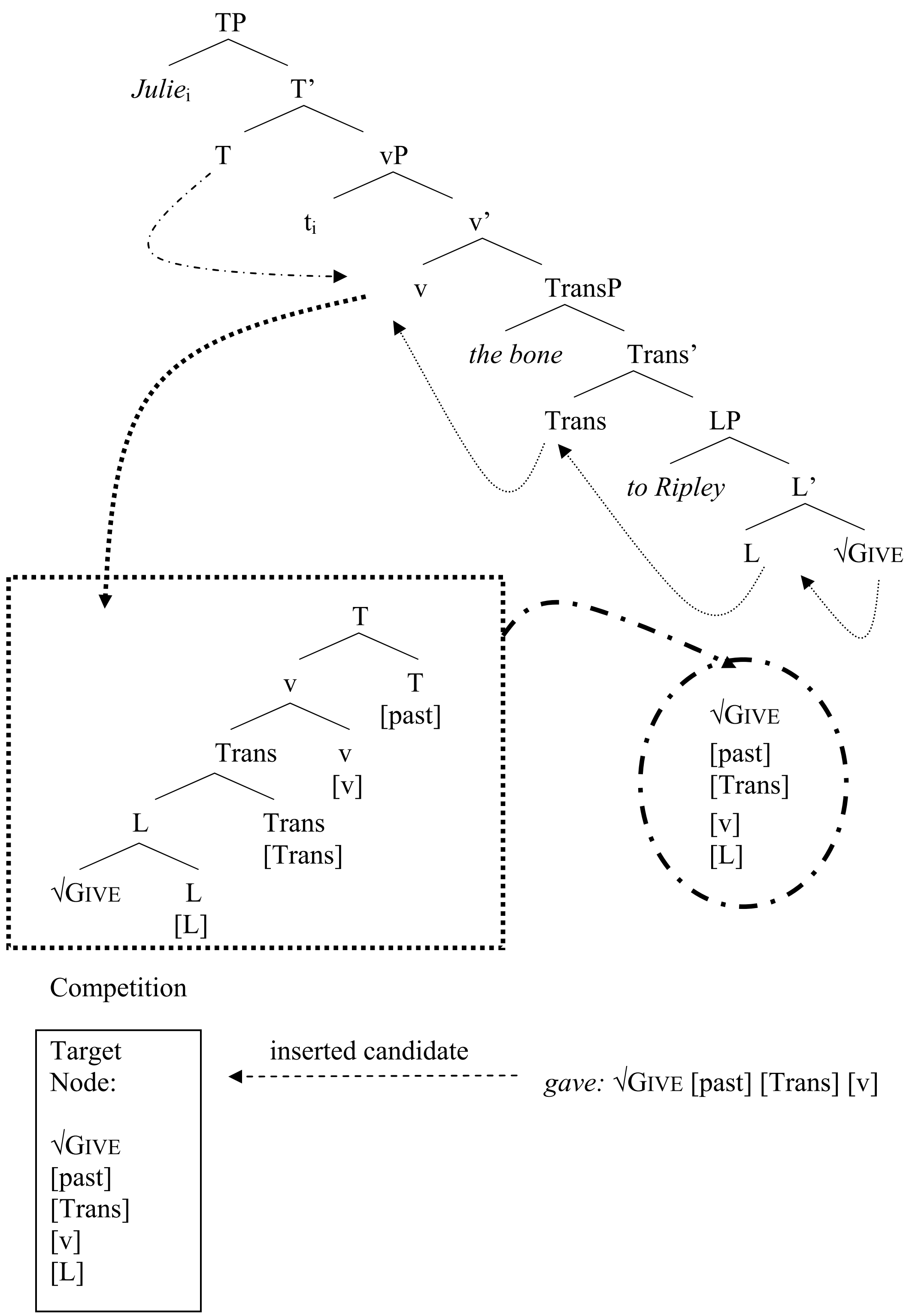
(4.34) a. Julie sent the package to France
b. #Julie sent France a package.

In the derivations above, I have been calling the head that projects the goal argument G and its corresponding feature [G]. I will continue that practice here, differentiating it from the locative, which I call L. With the assumption of two different heads comes two different derivations for the locative (4.35) and the dative (4.36). Applying the fusion analysis described above results in two different fused heads that are subject to VI insertion as illustrated below.

(4.35) Julie gave Ripley a bone.



(4.36) Julie gave a bone to Ripley.



This type of analysis of dative constructions allows us to generalize about the specifications for VIs of different classes of verbs. For example, words like *deliver* and *donate* allow locatives and themes optionally but strictly prohibit goals. Such a class of verbs would have a VI such as in (4.37).

(4.37) Vocabulary Item for *deliver*

√DELIVER	→	deliver
[v]		/dəlɪvr/
¬ [G]		

Such a VI allows insertion into a derivation with or without a locative argument and with or without a theme, but expressly bans its insertion into a derivation with a goal argument.

(4.38)
a. John delivered the pizza.
b. John delivered the pizza to Mary.
c. *John delivered Mary the pizza.
d. John hasn't delivered to Mary yet.

Similarly, the class of verbs that behaves like *ask* would be underspecified for most arguments but blocked in the case of [L]. Verbs such as *give* and *run* would be truly underspecified in terms of L and G. In such a way, a MINIMIZE EXPONENCE-based fusion approach to dative alterations that captures "dative shift" as a difference in a verb's selecting for or blocking locative arguments and goal arguments.

4.6 C-selection

Recall from (4.2, repeated here as 4.39) that there are words like *wonder* which select for the category that arguments must have.

(4.39) a. I wondered whether the fate of the world was in my hands.
b. *I wondered the fate of the world.

Wonder requires that its theme must be a CP. As discussed above, this aspect of subcategorization has historically been called category selection. This appears to be a problem for the analysis shown here because the specification of *wonder*—which by hypothesis would require the feature [Trans]—mandates only that the root be c-commanded by the Trans head, which introduces theme arguments. That specification has no way of determining that the specifier of Trans be a CP.

Harley and Noyer (2000) proposed that the little-v head comes in different varieties. If we extend that claim, all the heads that license arguments have different varieties. Thus, there is a Trans head that licenses a DP in its specifier, another that licenses a CP, and yet another that licenses a TP. Similarly, the claim that some vPs license the meaning CAUSE and others don't can be extended to Trans to apply to the difference between affected and non-affected objects: one Trans head licenses the meaning BECOME, another doesn't. The one that does not license BECOME would be the Trans head present in sentences with verbs like *push, wipe, hit* etc.

Throughout this dissertation, I have claimed that little-v carries one feature [v] and similarly Trans carries one feature [Trans]. Given that there is a variety of those heads, the heads are not composed of just one feature, but rather a bundle of features and that the different types of these heads are a result of subtle differences of the contents of those bundles. Words like *wonder* which specifically select for the Trans head that allows for CP specifiers and words like *push* that specifically select for the Trans head that does not carry a feature corresponding to BECOME can be assumed to be better specified than VIs that select for just [Trans]. That is, all VIs that require themes have been marked here as selecting for the feature [Trans], which is shorthand for the bundle of features contained in any given Trans head. Those VIs that are specific about which Trans heads they are compatible with must be even better specified for the features from the [Trans] head they must realize. They will be specified for a feature in a particular Trans head that sets it apart from the other Trans heads. Thus, *wonder* is only compatible with CP objects because the VI is only compatible with the features contained in the specific Trans head that licenses CP arguments.

4.7 Summary

The ultimate purpose of this chapter was to extend the predictions of the ubiquitous fusion motivated by MINIMIZE EXPONENCE beyond the realm of morphological phenomena to the realm of syntactic phenomena. To this effect, I showed that a fusion analysis can explain subcategorization. I argue here that the selectional behavior of a given verb is an effect of the functional features that the VI is specified for. Arguments are selected for by functional heads. Those functional heads fuse with the root, giving the target node both the root and the relevant formal features. This allows the root VI to select for those features.

Since DM is an underspecification model, this led to the prediction that roots should be compatible with arguments in excess of their core requirements, which can be seen in normal verbal behavior and extends to structural coercion. I also extend the feature blocking system described in chapter 3 to account for the blocking of arguments

(**The engineer arrived the train*). The end result is to show that the fusion motivated by MINIMIZE EXPONENCE also drives the selectional behavior of verbs.

CHAPTER 5: REMAINING ISSUES

Chapters 1, 3, and 4 represent the central proposals of this dissertation. The purpose of this chapter is to examine some other issues that result from MINIMIZE EXPONENCE. By its very nature, this chapter comes in the form of a list of several unconnected discussions on avenues of further research.

Section 1 considers the possibility of an interface between syntax and neo-Davidsonian event semantics that arises as a result of both the ubiquitous fusion driven by MINIMIZE EXPONENCE and the subcategorization analysis proposed in Chapter 4. Section 1 first details the model of semantics based on Parsons 1990, then its apparent incompatibility with modern syntactic theory. Then I examine how the proposals in this dissertation point to a possible interface between Minimalist syntax and neo-Davidsonian event semantics.

Since the majority of this dissertation focuses on English data, section 2 of this chapter is included to present some of the typological predictions that this model makes. In a similar vein, Section 3 details the predictions made by this model about the nature of verb classes in light of the specifications detailed in Chapter 4. Section 4 provides a discussion on the nature of specification given the proposal of the ¬ specification proposed in Chapter 3. In particular, Section 4 discusses the nature of unmarked forms and the Elsewhere Principle.

5.1 Event Semantics

In Chapter 4, I present a model of syntax based upon the proposal by Jelinek (1988). She proposes that there is a functional head Trans, which (parallel to little-v) is the head that is responsible for licensing the presence of a theme argument in the syntax. She argues that this head is often realized cross-linguistically as a transitivizing morpheme. In adopting her analysis, I essentially rejected the view that the theme argument is the complement of the verb (or the root). This assumption was crucial to my model of subcategorization. In order for the root to fuse with it and gain the [Trans] feature, the Trans head must be present in the syntax. Otherwise, no verb could be specified for requiring a theme argument.

Had this dissertation come before Jelinek's proposal, it would *predict* that model. In particular, I claim that VIs are specified for both content and formal features. Those formal features determine the argument structure of the verb. That means that there must be a formal feature that corresponds to the presence of the theme argument. It follows then that there must be a functional head associated with the theme argument. Thus, Jelinek's model of the syntax is now confirmed by an analysis that predicts the existence of that head.

Further, I claim that roots themselves have no inherent argument structure. Rather, the specifications of corresponding Vocabulary Items determine what argument structures a particular root is compatible with. This effectively reduces ungrammaticality of sentences such as **The engineer arrived the train* to a simple problem: The Vocabulary does not have a word with which it can realize the features generated by the syntax. There is nothing inherently ungrammatical with combining the root √ARRIVE with the features [Trans] and [v]; we just don't have a word to realize that particular combination of features.

The logical inference from this is that √ARRIVE itself does not have any requirements that the combinatorial system of syntax needs to satisfy—its only job is to provide content to what would otherwise be a derivation composed entirely of formal features. Roots enter the syntax already saturated (or more precisely, there is nothing to saturate). All the thematic relations of a particular root are structural, rather than inherent. This is consistent with the observation that, especially in English, roots such as *hit*, *table,* and *chair* so easily appear as both verbs and nouns. It also predicts the extreme productivity we find in English where verbs are zero-derived into nouns and vice-versa. Finally, it predicts the variety of grammatical argument structures and structural coercion.

Parsons (1990) proposes a theory formal semantics which is often today called Compositional neo-Davidsonian Event Semantics (Bayer 1997). In this model, sentences are treated as quantification over events. Specifically, Parsons treats every argument of a verb, even the verb itself, not as a function that needs to be saturated, but rather as an element that is conjoined to an event. In neo-Davidsonianism, all "words" entering into the semantics are themselves already saturated and all arguments are introduced by an independent function on the event.

The model of syntax proposed in Chapter 4 predicts both that verbs enter the syntax already saturated and that arguments are a function of the combinatorial system rather than the verb themselves. In this way, the model of syntax proposed here mirrors the Parsonian model of semantics.

5.1.1 Parsons (1990)

Parsons (1990) defended an observation first made by Davidson (1967)—hence the name, neo-Davidsonianism—that individual verbs can appear with a variable set of arguments and modifiers.

(5.1) a. Brutus stabbed Caesar in the back with a knife.
b. Brutus stabbed Caesar in the back.
c. Brutus stabbed Caesar with a knife.
d. Brutus stabbed Caesar.

For the sake of completeness, I'll add the following sentences:

(5.1) e. Brutus stabbed.
f. Caesar was stabbed.
g. Caesar was stabbed with a knife.
h. Caesar was stabbed in the back.
i. Caesar was stabbed in the back with a knife.

There are a number of important observations about the sets of sentences in (5.1) that Parsons (1990) makes.

1. (5.1.a.) entails all the other sentences in the two sets. In particular, the meaning of any one of the other sentences is present in (5.1.a) along with additional meaning. Importantly, this entailment is (often) unidirectional. (a) entails (b) and (c), but neither (b) nor (c) entail (a). (b) and (c) both entail (d) which entails neither

2. (5.1.a) entails the conjunction of any two of the other sentences (allowing for overlap). So *Brutus stabbed Caesar in the back with a knife* entails the conjunction of *Brutus stabbed Caesar in the back* and *Brutus stabbed Caesar with a knife*.

Extending those observations to the discussion in Chapter 4, it is important to add one more crucial observation.

3. There is nothing about *stabbed* that requires any of the modifiers or arguments in any of the sentence. *Stabbed* appears grammatically with all of them.[27]

The modifiers, including the arguments, in (5.1) are called intersective modifiers. Kadmon and Landman (1993) identify the following the following two important principles of these types of modifiers: a) they allow permutation, and b) they allow dropping. The example provided by Bayer (1997) is quite good at showing these processes. Examine the sentence in (5.2) from Bayer (1997).

(5.2) John is a forty-year-old, blond, blue-eyed American dressed in a suit, with a beard, in his midlife crisis.

The modifiers within (5.2) can be reordered (within the restrictions of the syntax) without changing the truth conditions of the sentence.

(5.3) John is a blond, blue-eyed forty-year-old, American with a beard, dressed in a suit, in his midlife crisis.

In addition, any of these modifiers can be dropped without changing the core meaning of the utterance.

Since these types of intersective modifiers can occur over both verbs (as seen in 5.1) and nouns (5.2), they must be quantifying over some other element. Parson, following Davidson (1967), concludes that they are (independently) quantifying over events. This leads Parsons to propose the formal semantic structure of an utterance such

[27] *Stabbed* cannot appear without any arguments, however, a fact attributable to the presence of some verbal head (and, perhaps, the EPP) in any given derivation.

as *Brutus stabbed Caesar in the back with a knife* as being a conjunction of several functions on one event.

(5.4) *Brutus stabbed Caesar in the back with a knife*

(∃e)[Stabbing (e) & Subj (e, Brutus) & Object (e, Caesar) & In (e, back) & With (e, knife)][28]

This structure perfectly captures the ability of the modifiers to be dropped to result in smaller sentences.

(5.5) *Brutus stabbed Caesar in the back with a knife*

(∃e)[Stabbing (e) & Subj (e, Brutus) & Object (e, Caesar) & With (e, knife)]

(5.6) *Brutus stabbed Caesar in the back*

(∃e)[Stabbing (e) & Subj (e, Brutus) & Object (e, Caesar) & In (e, back)]

Parson's (1990) proposal is crucially different from classical semantic theory in two ways: (a) there is only one operation: conjunction and (b) none of those operations are on *stab*. Rather *stab* is a function on the event and is not a function of *Brutus* on *Caesar*. Crucially, *stab* is indifferent to the presence of the arguments and other modifiers, contrary to classical formal semantics, where *stab* is a function whose variables (object and subject) need to be realized.

Thus the crucial features of neo-Davidsonian event semantics are the following: (a) individual words are not functions over their arguments, rather they are functions over an event; (b) the arguments of a verb are not functions on the verb, nor do they saturate

[28] This notation should be read as follows: there exists an event; that event is an event of stabbing, the subject of that event is Brutus, the object of that event is Caesar, that event was in such a manner that it was in the back, and that event used the instrument a knife.

the variables of the verb; they are functions on the event; and (c) there is only one function on the event: conjunction.

The combination of all the features of Parsonian event semantics has some interesting effects on the model of syntax. Assuming that the syntax's "job" is to introduce all the formal features necessary for LF to be interpretable, a model of the syntax that will be compatible with Parsonian event semantics will have to have several features.

Minimalist syntax shows a sentence to be built in some way or another around the verb (i.e. the verb is a member of the initial application of Merge to begin a derivation). In a model compatible with Parsons, the derivation would begin with the merger of the event with some other element, not the merger of the verb and its complement. Some functional head must introduce this event so that it can be interpreted at LF (and so all the syntax can be build on it).

The event is the element of the syntax that is consistently operated upon (functional heads would now serve the job of conjoining modifiers (including arguments) to the event head). Content items themselves must not be providing any syntactic information. Rather, since each modification is something like Subj (e, Brutus), there must be a functional head for each of those that conjoins the modifier to the event.

The next section will elaborate on the model of syntax that is compatible with Parsons' proposal.

5.1.2 Neo-Davidsonianism Meets Syntax

In the last decade, there have been a number of proposals that have mirrored neo-Davidsonian event semantics in the syntax. Some of note are Borer (1994, 2004), Marantz (1997), Ramchand (1997, 2002), Ritter and Rosen (1998), and Travis (1994). Ramchand (1997) is one of the more important of the proposals in that it proposes that Aspect is the locus for the introduction of the event element in the syntax. Ramchand argues that, in many cases, verbs themselves are complex—being composed of a content morpheme and a functional head. She argues that the Asp head introduces the event and also introduces the object.

Ramchand (2002) elaborates that schema to show that three functional heads compose a verb and its argument structure: one introducing the causer, one introducing the "process" meaning, and one introducing the "result". Borer (1994, 2004) also proposes a similar breakdown where all semantic relations with the event are introduced with a corresponding functional head, though hers are different from Ramchand's. Especially in regard to arguments and theta roles, the Parsonian syntactic analyses all tend to have one thing in common: most, if not all, arguments are introduced by a functional head rather than by the verb.

Ramchand (1997) is one of the exceptions to this rule. For Ramchand, unlike Borer's (1994, 2004) and Marantz's (1997) proposal, the verb itself is still responsible for some arguments—in particular, those that get weak case (i.e. genitive). For Ramchand, the arguments projected by the root (she calls it the verbal noun) are different from arguments projected by functional heads in two ways: a) in the semantics, they are interpreted as modifying the verb, not the event, and b) their interpretation is imperfective (rather than perfect, which Asp introduces).

Ultimately, models of the argument structure that are based on Parsonian event semantics are split into two camps. The difference between the two camps is whether the verb itself licenses an argument. Models such as Borer's and Marantz's show all argument structure to be syntactic while the models that follow Ramchand assume some arguments licensed by the verb itself. For Ramchand, the difference is motivated by the overt presence of an Aspect head, which triggers different argument structure and case. In order to account for Ramchand's data, the model I have espoused here (Borer's and Marantz's) predicts an unpronounced head that licenses what Ramchand calls "weak case".

The syntax that I propose in Chapter 4 contributes to this debate. As I mention above, the MINIMIZE EXPONENCE-driven fusion analysis of morphology and subcategorization, which strives to capture the same data the analyses I have discussed here do, predicts (and requires) a functional head for each argument in order for root to gain the relevant interpretable features. Further, there is an account for the "missing head" that Ramchand observes. An analysis of Ramchand's data that is compatible with my proposal would show the imperfective head to have fused with the root while the perfective head failed to fuse, causing overt spellout. In this way, the predictions I have made here act as a kind of tiebreaker between Borer and Ramchand in that for other reasons (specifically, in order to license the insertion of the Vocabulary Item), I predict that all arguments are licensed by functional heads.

Ramchand's model of event semantics is clearly a hybrid of Parsonian semantics and more classical semantics (which is why she calls it Post-Davidsonian rather than neo-Davidsonian) in that not all modifiers modify the event—some modify the verb. Borer's model is still consistent with the observation that argument structure is independent of the root (which was one of the motivations behind the proposals made by both Davidson 1967 and Parsons 1990).

5.1.3 Remaining Questions:

In Section 1.2 above, I describe some of the ways that the predictions of the neo-Davidsonian model of semantics are transferred to the realm of syntax. This transference serves both to show a mirror between the syntax and the semantics and provide a means for the semantics and the syntax to interface. To be more specific, if the logical form of a sentence is such as we saw in (5.4, repeated here as 5.7), we want the syntax to send to LF something that looks like that.

(5.7) *Brutus stabbed Caesar in the back with a knife*

(∃e)[Stabbing (e) & Subj (e, Brutus) & Object (e, Caesar) & In (e, back) & With (e, knife)

Section 1.2 discusses a few of the things that would have to be true of the syntax in order for it to produce such a logical form. The first is that the syntax would have to have functional projections for each modifier whose job was to conjoin the argument and the event in a meaningful way. We see that realized in the proposal here and those by Borer, Ramchand, and others. The next is that the individual roots would have to be completely saturated (i.e. they themselves are not functions with empty slots, but the targets of the conjunction introduced by the functional heads). In the proposal in Chapter 4, we see that to be predicted as well.

However, a number of issues impede a total mirroring of the logical form predicted by Parsonian event semantics and the LF that would result from even the syntax predicted here. The first is the event variable. According to the LF predicted by Parsons, each modifier is a conjunction of the argument with the event variable. Ramchand proposes that the event variable is introduced by the Aspect head, but that accounts for only one of the event variables. The simplest solution is to argue, expanding upon Ramchand, that in fact, each functional head introduces the event variable that it conjoins to its specifier. This immediately accounts for the fact that for each argument, there is a corresponding event variable.

This however leads inexorably to the next problem with a matching of syntax to the LF predicted by Parsonian semantics: hierarchical structure versus flat structure. Parson's model has a flat structure. Each argument function such as Subj (e, Brutus) is just conjoined one after the next in a series of conjunctions. This property is one of the key reasons that this type of LF so easily allows for permutation and dropping. There is no hierarchical structure at all (except in the case of complex phrases, which have their own corresponding event variable).

One tenet of the Minimalist Program is that syntax has complex hierarchical structure. In particular, Merge is construed as *recursive* set formation (Chomsky 1995). The application of Merge (except at the very bottom of the derivation) takes one or more extant sets and constructs a set that contains both sets. This results in complex hierarchical structure. That structure is completely absent from the Parsonian model of event semantics – it is just a series of conjunctions of an event and an individual. Therefore, while Parsons' mechanisms create sets, they are often sets of individuals. Thus, Parsons' mechanisms don't show the embedding assumed in Minimalist syntax. This is complicated by the fact that hierarchical sentence structure is crucially ordered. Parsons' mechanism, as it employs only conjunction, results in an unordered LF.

There are two possible ways to account for this apparent incompatibility. The first is to stipulate that the functional heads in the core syntax have selectional requirements of their complements. While the core job of the functional head is a function between its specifier and the event, it also selects for where it can appear in the derivation. Specifically, it selects what it can be merged with to get it into the derivation. This stipulation would mean that the ordering (and some of the hierarchical structure) is a property only of the core syntax and that it has no effect on the LF. Its function would be linearization. The selection of the heads would permit the pronunciation of the derivation, since we speak one dimensionally (excluding time) and an LF representation is two-dimensional. Another, solution to the problem is to show that Merge is itself not actually a function that makes a set out of other sets, but is instead an operation that only targets individuals.[29]

Ultimately, while the purpose of this section is only to show that the remaining apparent incompatibilities between syntax and Parsons' model are not insoluble, the

[29] See forthcoming work done by David Medieros.

argument that syntax is hierarchical while LF is not is more satisfactory. The basics of the recursivity mandated by the Minimalist model of syntax do not entail that LF need be recursive. Rather, just as there as a linearization process at PF, there could be a similar flattening at LF for interpretation or the hierarchical nature of syntax could only serve the purpose of linearization.

5.2 Typological Predictions

MINIMIZE EXPONENCE as proposed here is intended to be a PF constraint on UG, not an English-specific constraint. Since the majority of this dissertation revolves around English data, the purpose of this section is to elaborate on some of the typological predictions of this model.

I argue the fusion that is driven by MINIMIZE EXPONENCE is mandatory. In order to realize the largest amount of formal features with the fewest morphemes, the solution proposed here is to fuse functional material whenever possible (fusion being only applicable to complex head structure, thus it will never reduce a sentence to just one word, nor will it ever cause two roots to fuse together[30]). This is not the only way to satisfy MINIMIZE EXPONENCE. If a language had access to a process to discharge features from a derivation other than insertion (one possible explanation of pro-drop, I suppose) that would also satisfy MINIMIZE EXPONENCE.

[30] The only possible time two roots could fuse together is if we hypothesized root-root compounding to be the result of head movement or some other process that creates a complex head rather than a syntactic phrase. In my description of compounding in chapter 3, I assumed that compounding results in a syntactic phrase, not a complex head, so in the model I assume the issue is moot. In a syntactic theory where compounds are complex roots, the constraint on insertion mentioned earlier that roots must be realized by a VI, would disallow fusion of two roots.

Furthermore, whether or not a node fuses in the system sketched above is entirely dependent upon the specification of the VI[31] that will be inserted into that node. For example, since *walk* is specified a ¬ [past], then the fusion of the root √WALK with the [past] feature is effectively blocked. Since *walk* is indifferent to [3sg], the fusion is not blocked. The amount of fusion in a given language is entirely dependant on Vocabulary inventory of that language. In that way, many of the morphological and syntactic differences of a language can arise entirely from the specifications of that language's VIs. For example, in English, nominative and accusative cases always fuse with the root (which is why the only overt realization of these cases is in suppletive forms like pronouns and *whom*). This means that no words in English contain ¬ specifications for nominative or accusative. Conversely, in a language like Latin, where nominative and accusative are regularly spelled out with an affix (5.8), the "default" for a VI would be to have a ¬ specification for nominative and accusative.

(5.8) Latin nominal inflection

stem		nominative	accusative
puella	'thing'	*-ø*	*-m*
deu-	'god'	*-s*	*-m*
manu-	'hand'	*-s*	*-m*
re-	'thing'	*-s*	*-m*

On a larger scale, this difference in specifications could give rise to the difference between fusional and agglutinative/isolating languages. Agglutinative/isolating languages such as English or Mandarin have an inventory of VIs with a proportionally

[31] In a roundabout way... what is actually happening is that there is a relationship between the amount of fusion in the derivation selected to converge and the VIs inserted into that derivation.

high number of ¬ specifications. These ¬ specifications impede fusion more often, leading to more features being realized as independent morphemes (either bound in the case of English or free in the case of Mandarin). On the other hand, the inventory of more fusional languages (such as Spanish) contain proportionally fewer ¬ specifications (and correspondingly, VIs contain more features that they realize) leading to morphemes which realize comparatively more features. For example, in Spanish the verbal affix *–o,* in *hablo* 'I speak', realizes a large set of features—1st person, indicative mood, singular, present tense, and perfect aspect.

Similarly, the amount of syncretism in a given language is also function of that language's ¬ specifications. A language whose VIs contain a large number of ¬ specifications will undergo less fusion, requiring an greater inventory of functional morphemes while a language with fewer ¬ specifications will undergo more fusion allowing it to employ fewer functional morphemes to realize the same amount of interpretable features.

These differences among languages is not parametric but is rather an effect of the specifications in their VIs. So-called fusional languages are languages where the inventory of the Vocabulary contains a large amount of VIs that are highly specified for features they realize. Conversely, a non-fusional language contains a large amount of VIs that are specified for what they must not realize, forcing those features to be otherwise discharged. The difference between an agglutinating and an isolating language is a function of how much the functional morphemes in the language's inventory are free—an agglutinating language contains more bound functional morphemes. It follows then that polysynthetic languages may not be a matter of parametric variation (Baker 1988) but rather just a matter of VI content.

5.3 Verb Classes

So much work has been done on verb classes throughout the field of linguistics (see Levin 1993 for a detailed approach) that they are often assumed, especially in the realms of inflectional morphology and subcategorization, realms that this dissertation addresses in detail. The basic hypothesis is that verbs (and nouns for that matter) are organized in the Lexicon and that the morphology and syntax are sensitive to that organization. For example, the general assumption in the study of classical Latin is that its verbs belong to five classes (or conjugations). Membership to these classes determines the inflectional paradigm that the verb employs. Similarly, for argument structure, the general argument (Levin 1993) is that verbs belong to certain classes and that those classes determine that verb's argument structure.

The first concern about the idea of verb classes comes from the very structure of the Distributed Morphology grammar. Since there are no paradigms in DM, class membership would have to be a realization of some interpretable feature in the syntax. Returning to Latin conjugations, according to DM, if there is an affix that is somehow sensitive to the class membership (e.g. *–amus, -imus, -emus* for first person plural present tense—1st conj., 2nd conj, and 3rd conj. respectively), that must mean that there is something in the syntax that contributes that meaning. That syntactic feature would be analogous to gender. That feature would be realized by its own affix—in the Latin case –

a-, -i-, or *–e-*, since those sounds appear in almost every member of their respective paradigms.[32]

Since in DM, paradigmatic morphology is reduced to a set of Vocabulary Items that happen to have shared features in their specifications—which explains cross paradigmatic syncretisms (*-mus* in the example above)—the existence of verb classes as targets of paradigmatic morphology is disallowed by the mechanics of the framework. In Chapter 4, I propose that the selectional behavior of a verb is also just a matter of its feature specification. Thus the fact that *throw* and *fax* (example from Bresnan and Nikitina 2003) have the same selectional behavior is not due to their belonging to any particular verb class, but rather is attributable to their having identical (or near identical) specifications for abstract (non-root) formal features.

In order to account for the sum of verbal behavior in English, Levin had to propose a large number of verb classes, many of which have a very small membership. This large number of verb classes is a prediction of the analysis proposed in Chapter 4. The apparent large number of verb classes is attributable to the fact that VIs can be specified for any combination of the relevant features. Since there are several of the argument features proposed in Chapter 4, and a VI can select for not only any set of those features but also for the corresponding ¬ specification, there is a large but finite set of possible specifications. The magnitude of the set of possible specifications gives rise to the large number of verb classes proposed by Levin. The finiteness of that set is what

[32] Alternatively , the VI for the verb could be in someway specified for class membership. Then that class membership could be a secondary exponent of the affix. Then however it is curious as to what feature is being realized by *–a-, -i-,* and *–e-*.

gives rise to the appearance of verb classes in general. Thus, verb classes are an artifact of the finite set of functional heads that project arguments and the finite set of possible feature specifications. Verb classes reduce to the sets of VIs that have identical feature specification.

This argument helps to account for why speakers can disagree on the selectional behavior of a verb (e.g. whether *I pondered whether the world is round* is grammatical). In the model developed here, the difference across speakers is attributable only to the presence or absence of a feature specification for that verb. In a model with verb classes, that difference would have to be attributable to the verb being a member of a different class. Since class membership is in part driven by semantics, it is not entirely clear how speakers who agree on the meaning of a word could differ on the class to which it belongs.

5.4 On Feature Specification and the Elsewhere Principle

There is a curious result of the model predicted here as to the structure of the Vocabulary (or lexicon). Recall from Chapter 3 that regular morphology is triggered by the ¬ specification. So the past tense of *thrash* is *thrashed* because *thrash* is specified as ¬ [past] (3.43, repeated here as 5.8)

(5.8) Vocabulary Entry for *thrash*

√THRASH	→	*thrash*
[v]		/θræʃ/
¬ [past]		

It follows then that all words which are regularly inflected and any stem that requires derivational morphology (i.e. does not zero derive) must be contain a ¬ specification for each feature realized by an affix it is compatible with. This means that unmarked forms such as *thrash, walk, black, arrive,* etc which participate in the all normal regular morphological processes of a language, while they are relatively unspecified for the formal features they must realize (cf. *ran, ate, mice* etc which are relatively more specified) are heavily specified for the features the must not realize. I.e. while they have a small amount of traditional feature specifications, they have a large amount of ¬ specifications. Returning to *thrash*: *thrash* is relatively unspecified in traditional terms (realizing only a root and the feature [v]). However, since it gets regular inflection for past tense, progressive aspect, third person singular present, etc., it is heavily specified with ¬ specifications.

One of the fundamentals of Distributed Morphology is the Elsewhere Principle. The Elsewhere Principle mandates that the VI with the smallest feature specification will appear in the most environments (i.e. is the "default"). However, in this model the "default" forms in a language are likely to be heavily specified with ¬ specifications. In terms of formal morphological theory, this doesn't have any drastic effect on the model of the grammar. Rather, it is just counterintuitive that the most underspecified forms tend to be highly specified on another axis.

Where this might have some effect on the model of the grammar is in learnability theory. A child overgeneralizes regular morphology (e.g. **goed*). In the model proposed here, that indicates that when a learner of the language hears a word, they assume that it is heavily specified with ¬ specifications in addition to its traditional specifications. For example, a learner might assume the following VIs for *go* and *went*, giving rise to **goed.*

(5.9.a) Speculative Vocabulary Item for *go*

√GO	→	*go*
[v]		/gow/
¬ [3sg]		
¬ [past]		

(5.9.b) Speculative Vocabulary Item for *went*

√WEN	→	*went*
[v]		/wɛnt/
[past]		

5.5 Summary

The purpose of this chapter was to provide a quick list of future avenues of research in light of the MINIMIZE EXPONENCE-driven fusion that I proposed in Chapter 3. One of those avenues I discussed here was the interface of the syntax outlined in Chapters 3 and 4 with the neo-Davidsonian model of event semantics. As I discussed here, it seems that the next obvious step in that direction is to outline the reasons that syntax is hierarchical while neo-Davidsonian event semantics is not.

I also described some of the typological predictions of this model. Obviously, the next step down that avenue of research is to describe the behavior of irregular forms and inflectional patterns in other languages, such as Spanish, Latin, or Mandarin as surveyed above, in light of the fusion predicted here. A further step is to describe the argument structures of other languages as a function of fusion. One interesting avenue to pursue in particular are languages that have overt heads that introduce arguments such as the Uto-

Aztecan language family (Jelinek 1988, Haugen 2004) or the Salishan language family (Bischoff 2004). In that case, the argument would be that the root has failed to fuse with Trans (since Trans is realized by an overt morpheme), so the prediction from this model is that verbs could not select for themes.

Another avenue of research stemming from the discussion in this chapter would be to compile a list of the specifications of verb VIs. From the predictions here, the sets of VIs with identical specifications should somewhat align to the verb classes predicted by Levin (1993). Finally, the final obvious avenue of research is examining the ramifications the ¬ specification has on the Elsewhere Principle and analyses that have relied upon it. The corollary to this is to see if there is a way to test the abstract claims of this model against child acquisition data to see if the realization of the Elsewhere Principle is realized as the positing of ¬ specifications in language acquisition.

CHAPTER 6: CONCLUSIONS

The purpose of this dissertation is to address conflict in the grammar whereby on one hand, we have the need to convey a message and the need for that message to be as clear as possible but on the other, we need our message to be as efficient as possible. While this conflict is prevalent in much of the linguistic literature, in the realm of morphosyntax, this conflict is largely ignored by theoreticians. In syntax, one such realization of this conflict is the following: a maximally contrastive grammar involves pronouncing every single formal feature in a maximally contrastive manner whereas a maximally efficient grammar sums all the features of an utterance into fewest words.

The Minimalist tradition is largely a study of what the system is and isn't capable of, not how that system is put to use. However, this results in a certain loss of predictive power. I proposed an economy constraint that captures these competing forces on the grammar—in particular the balance necessary in pronouncing all the interpretable features of a given derivation in the most efficient way possible (1.1, repeated here as 6.1).

(6.1) MINIMIZE EXPONENCE

> The most economical derivation will be the one that maximally realizes all the formal features of the derivation with the fewest morphemes.

The gist of this constraint is that the best utterance is the one that conveys the most amount of information with the least effort (measured in number of morphemes that have to be pronounced). In terms of the production of an utterance, this constraint captures the struggle between the need to be maximally contrastive and the need to be maximally efficient.

This dissertation focused on exploring the role that this type of constraint would have on the model of the grammar by specifically looking at the effects of MINIMIZE EXPONENCE on analyses of familiar morphosyntactic phenomena.

To this end, I presented an alternate analysis to lexical categories and root allomorphy within the framework of Distributed Morphology. The traditional analysis of lexical categories in DM is to have the content VI licensed for insertion through secondary licensing (targeting the immediately c-commanding functional head). The traditional DM analysis of root allomorphy employs the use of readjustment rules (again conditioned by c-commanding functional heads) to change the phonology of on VI to another VI. In the analysis presented in Chapter3, I argued that licensing is a local process of targeting features actually in the node targeted for insertion. I argued that the root and the formal features come to occupy the same terminal node through the application of the process of fusion. I extended this analysis to an analysis of root allomorphy. This extension allows words like *eat* and *ate* to be separate VIs that compete with each other for insertion rather than one being the result of the application of a readjustment rules.

This proposal has a number of effects on the DM model of grammar. The first is that it limits application of secondary licensing, removing lexical categorization from the list of responsibilities of secondary exponence. Since secondary licensing is inherently less local and thus less efficient than primary licensing, this is taken to be a strength of this proposal. The other important effect of this proposal was to remove the need for readjustment rules in the DM grammar. Since VIs linked to the same root now compete with each other for insertion, readjustment rules are no longer need to alter the phonology of irregular forms. As readjustment rules only serve the purpose of changing the phonology of irregular morphemes, since they aren't needed for that function any longer, the DM grammar doesn't need them at all. Since readjustment rules entail both long distance relationships and extra computational load, another strength of this dissertation is that it proposes a model of DM without readjustment rules.

In the current model of DM, functional morphemes and content morphemes in many ways participate in two different grammars. Whereas functional VIs participate in competition, content VIs did not, relying on readjustment rules and secondary exponence (never primary exponence) to license their insertion. The proposal here, in abolishing the need for both secondary licensing and readjustment rules and in elaborating on a theory of competition of roots, shows that it is possible for the model of the grammar to use just one process for insertion of both functional and content VIs. As this simplifies the model of the grammar, this is again taken as a strength.

The realization of MINIMIZE EXPONENCE here is the ubiquitous fusion of functional heads in complex head arrangements. The effect of this fusion is that the features of those heads are realized without each needing to be realized by its own VI. The efficiency in the grammar comes from there being fewer VI items to be pronounced without loss of contrast (any ambiguity is then only a result of a given language's inventory). Thus, fusion is one possible tactic available to the grammar to satisfy this central conflict. As an added effect, a large portion functional heads that are realized as null morphemes in DM instead fuse with other heads to be realized by overt morphemes. This drastically reduces the number of null heads predicted by DM.

I proposed an analysis for the blocking of regular inflection in nominal compounds in English (e.g. **rats-infested*). For DM, this phenomenon has been difficult to analyze due to the tenets of the framework. There were two central problems for DM: 1) in DM the structures *lice-infested* and **rats-infested* are identical with the exception of the application of a readjustment rule; and 2) DM recognizes no difference between inflectional and derivational morphology (and derivational morphology is allowed in such structures). The analysis shows here that the structures of **rats-infested* and *mice-infested* are not identical: one involves more fusion driven by MINIMIZE EXPONENCE. The grammaticality of the constructions reduces to the embeddedness of the feature [n] (which meant that derivational morphology is allowed as long as a nominalizer was the dominant morpheme in the non-head member).

In Chapter 4, I extended the predictions of the ubiquitous fusion motivated by MINIMIZE EXPONENCE beyond the realm of morphological phenomena to the realm of syntactic phenomena. I showed that a fusion analysis explains subcategorization effects in the grammar. I argue here that the selectional behavior of a given verb is an effect of the functional features that the VI is specified for. Arguments are selected for by functional heads. Those functional heads fuse with the root, giving the target node both the root and the relevant formal features. This allows the root VI to select for those features. Since DM is an underspecification model, this led to the prediction that roots should be compatible with arguments in excess of their core requirements, which can be seen in normal verbal behavior and extends to structural coercion. The end result is to show that the fusion motivated by MINIMIZE EXPONENCE also drives the selectional behavior of verbs.

In Chapter 3, I proposed a system of ¬ specifications that allows VIs to block their insertion into nodes containing features they are incompatible with. This set up a system that accomplished several things: first and foremost, it provides DM with a feature blocking mechanism. Second, it triggers regular morphology. I also extended the feature blocking system to account for the blocking of a verb's arguments (**The engineer arrived the train*), since in the proposal here, a verb's argument structure is an artifact of the specification of its VI.

In Chapter 5, I provided a quick list of future avenues of research in light of the MINIMIZE EXPONENCE-driven fusion that I propose. One of those avenues I discussed here was the interface of the syntax outlined in Chapters 3 and 4 with the neo-Davidsonian model of event semantics. I also described some of the typological predictions of this model, discussed a possible theory of verb classes within DM, and provided an overview of the effects of my proposal on the Elsewhere Principle.

6.1 Some Final Thoughts

MINIMIZE EXPONENCE is PF constraint on the grammar which essentially instructs the grammar to produce utterances that perfectly balance the need to be contrastive with the need to be energy efficient. I wonder, then, what other kinds of PF constraints are on the grammar that affect the morphosyntax. For example, linearization might be the effect of a constraint that instructs the grammar to produce utterances that are pronounceable, given the constraints on using language.

One such area that I think these types of PF constraints might extend into is homophony-avoidance. Since homophony-avoidance in the realm of morphosyntax is gradient (in that we try to try to create maximally contrastive utterances but still end up with structural ambiguity) it follows that this is again a balance struck by the grammar (please see a survey of morphosyntactic homophony avoidance in Haugen and Siddiqi 2006) . This balance might arise from a violable constraint that prohibits syntactically derived ambiguity (Flack 2006) or morphologically derived ambiguity (Urbanczyk 1999, Siddiqi 2004), given a model of the morphosyntax driven by constraint hierarchies (such as OT-LFG, Bresnan 1996, 2000, Johnson 1998). This balance could also arise from a PF economy constraint similar to MINIMIZE EXPONENCE that requires the grammar to produce utterances with the minimal amount of morphosyntactic ambiguity.

I think that MINIMIZE EXPONENCE is likely one of a family of economy constraints on the PF branch of a derivation whose purposes could range from maximizing efficiency of the utterance to maximizing its contrastiveness and even to maximizing its interpretability. I believe this realm of morphosyntax (constraints and operations that act upon a derivation between spellout and phonological processes) is an exciting avenue of future research and I believe that the framework of DM is a perfect framework to pursue those questions.

REFERENCES

Adger, D. 1994 The Licensing of Quasi-Arguments. Proceedings of CONSOLE 1 ed. by P. Ackema & M. Schoorlemmer, 1-18. The Hague: Holland Academic Graphics.

Aronoff, M. 1994. Morphology by itself. MIT Press: Cambridge.

Baker, M. 1988. Incorporation: a Theory of Grammatical Function Changing. University of Chicago Press: Chicago.

Bayer, J. 1997. CP-Extraposition as Argument Shift. Rightward Movement, ed. by D. Beerman, D. LeBlanc, and H. C. van Riemsdijk, 37–58. John Benjamins: Amsterdam.

Bischoff, S. 2006. Basic Clause Structure in Coeur d'Alene: A Preliminary Working Paper. MIT Working Papers in Linguistics. Working Papers on Endangered and Less Familiar Languages 7, Studies in Salishan ed. by S. Bischoff, L. Butler, P. Norquist, and D. Siddiqi, 159-190 Cambridge: MIT Working Papers in Linguistics.

Bobaljik, J. 1994. What does adjacency do? MIT Working Papers in Linguistics 22: The Morphology-Syntax Connection, ed by H. Harley and C. Phillips, 1-32.

Bobaljik, J. 2002. A-chains at the PF-interface: Copies and 'Covert' Movement. Natural Language and Linguistic Theory 20.2:197-267.

Borer, H. 1994 The Projection of Arguments. Functional Projections, University of Massachusetts Occasional Papers 17, ed. by E. Benedicto and J. Runner. Amherst Massachusetts: GLSA, University of Massachusetts.

Borer, H. 2004 Structuring Sense (two book manuscript), Oxford University Press

Bresnan, J. 1996. Morphology Competes with Syntax: Explaining Typological Variation in Weak Crossover Effects Is the Best GoodEnough?, edited by Barbosa et al. MIT Press.

Bresnan, J. 2000. Pidgin Genesis and Optimality Theory. Processes of Language Contact: Case Studies from Australia and the Pacific, ed. by J. Siegel. Montreal: Les Editions Fides, 145--173.

Bresnan, J. 2001. Lexical-Functional Syntax. Oxford: Blackwell

Bresnan, J. and T. Nikitina. 2003. On the Gradience of the Dative Alternation. [ms.]

Chomsky, N. 1957. Syntactic Structures. The Hague: Mouton.

Chomsky, N. 1965. Aspects of the Theory of Syntax. Cambridge: MIT Press.

Chomsky, N. 1970. Remarks on Nominalization. Reprinted in Davidson and Harman. pp. 262

Chomsky, N. 1981. Lectures on Government and Binding. Foris, Dordrecht.

Chomsky, N. 1989. Some Notes on Economy of Derivation and Representation. MIT Working Papers in Linguistics 10 ed. by I. Laka and A. Mahajan. Cambridge: MIT Working Papers in Linguistics

Chomsky, N. 1994 Bare Phrase Structure. Occasional Paper #5, Cambridge: MIT Working Papers in Linguistics.

Chomsky, N. 1995. The Minimalist Program. Cambridge: MIT Press.

Chomsky, N, and M. Halle. 1968.The Sound Pattern of English. New York: Harper and Row.

Davidson, D. 1967. Truth and Meaning. Synthese 17. 304-323

Embick, D. 2003. Locality, Listedness, And Morphological Identity Studia Linguistica,

Volume 57, Number 3 (December 01, 2003), pp. 143-169
Embick, D. and A. Marantz. 2006. Architecture and blocking. University of Pennsylvaniaand MIT. [ms.]
Embick, D. and R. Noyer. 2001. Movement Operations after Syntax. Linguistic Inquiry 32:4, 555–595.
Embick, D. and M. Halle. 2005. On the status of stems in morphological theory. Proceedings of Going Romance 2003 ed. by T. Geerts and H. Jacobs. John Benjamins, Amsterdam/Philadelphia, 59–88.
Falk, Y. 2001. Lexical-Functional Grammar: An Introduction to Parallel Constraint-Based Syntax. Standford: CSLI Publications
Flack, K. 2006. OT contrast preservation and disambiguation in Japanese scrambling. Paper presented at the Annual Meeting of the Linguistic Society of America. Albuquerque, NM: January 6, 2006.
Fodor, J. 1998. Concepts. Oxford: Clarendon
Gleitman, L. 1990. The Structural Sources of Verb Meaning. Language Acquisition 1. 3-55.
Gleitman, L., H. Gleitman, C. Miller, and R. Ostrin. 1996. "'Similar' and Similar Concepts". Cognition 55. 321-376.
Goldberg, A.. 1995. A Construction Grammar Approach to Argument Structure. Chicago: University of Chicago Press.
Grimshaw, J. 1979. Complement Selection and the Lexicon. Linguistic Inquiry Cambridge, MA 1979; 10: 279-326
Grimshaw, J. 1990. Argument Structure MIT Press Cambridge MA
Gropen, J. 1989. The Learnability and Acquisition of the Dative Alternation in English Language: Journal of the Linguistic Society of America, June; 65 (2): 203-57.
Hale, K., and S. Keyser. 1993. On Argument Structure and the Lexical Representation of Syntactic Relations. The View from Building 20 ed. by Samuel J. Keyser & Kenneth Hale. Cambridge: MIT Press. 53-109.
Halle, M. 1997. Distributed Morphology: Impoverishment and Fission. MIT Working Papers in Linguistics: PF: Papers at the Interface. B. Bruening, Y. Kang, M. McGinnis, 425-49. Cambridge: Department of Linguistics, MIT;
Halle, M. and A. Marantz. 1993. Distributed Morphology and the Pieces of Inflection. The View from Building 20 ed. by S. Keyser and K. Hale,111-176.. Cambridge: MIT Press.
Halle, M. and A. Marantz. 1994. Some Key Features of Distributed Morphology. MIT Working Papers in Linguistics 21: Papers on Phonology and Morphology. Ed. by A. Carnie, H. Harley, and T. Bures, 275-288. Cambridge: MIT Working Papers in Linguistics.
Hammond, M. 1984. Level Ordering, Inflection and the Righthand Head Rule. MIT Working Papers in Linguistics vol. 7
Harley, H. 1995. Subjects, Events and Licensing. PhD dissertation, MIT.
Harley, H. and R. Noyer. 1999. State-of-the-Article: Distributed Morphology. Glot International 4.4, 3-9
Harley, H. and R. Noyer. 2000. Formal versus Encyclopedic Properties of Vocabulary:

Evidence from Nominalisations. The Lexicon-Encyclopedia Interface, edited by Bert Peeters, Elsevier Press.

Haugen, J. 2004. Issues in Comparative Uto-Aztecan Morphosyntax. Ph.D. dissertation, U. Arizona

Haugen J. and D. Siddiqi. 2006. From Semantic Contrast to Phonological Contrast: Homophony Avoidance in Morphology. Paper presented at the Annual Meeting of the Linguistic Society of America. Albuquerque, NM: January 7, 2006.

Hayes, B. 1980. Aklan stress: disjunctive ordering or metrical feet? Cahiers linguistiques d'Ottawa 9,in Papers from the Tenth Meeting of the North East Linguistic Society, University of Ottawa.

Hornstein, N. 1999 Movement and Control. Linguistic Inquiry, 69-96.

Jackendoff, R. 1996 "Semantics and Cognition" in S. Lappin (ed.), Handbook of Semantics, Oxford: Basil Blackwell Inc.

Jelinek, E. 1998 "Voice and Transitivity as Functional Projections. The Projection of Arguments: Lexical and Compositional Factors in Yaqui ed. by M. Butt and W. Geuder. Stanford, CA: Center for the Study of Language and Information viii

Jesperson, O. 1909-1949. A Modern English Grammar on Historical Principles. London: George, Allen and Unwin.

Johnson, M. 1998. Optimality-theoretic Lexical Functional Grammar. Presented to the 11th Annual CUNY Conference on Human Sentence Processing, Special Session on the Lexical Basis of Syntactic Processing: Formal and Computational Issues, Rutgers University, March 20, 1998.

Kadmon, N. and F. Landman 1993. Any. Linguistics and Philosophy, 16: 353-422.

Kiparsky, P. 1982. Lexical phonology and morphology. In Linguistics in the Morning Calm, ed. I. S. Yang, pp. 3-91. Seoul: Hanshin.

Kiparsky, P. 1997. Remarks on denominal verbs. Complex Predicates ed. by A. Alsina, J. Bresnan and P. Sells, 473-499. Palo Alto, CA: CSLI Publications.

Kiparsky, P. 2003. Syllables and moras in Arabic. The Syllable in Optimality Theory, ed. by C. Féry and R. van de Vijver, 147-82. Cambridge: Cambridge University Press.

Kratzer, A. 2002. Telecity and the Meaning of Objective Case The Syntax of Time ed. by J. Guéron, and J. Lecarme, 389-423. Cambridge, MA: MIT Press

Kratzer, A. 1996. Severing the External Argument from its Verb. Phrase Structure and the Lexicon ed. by J. Rooryck and L. Zaring, 109-37. Dordrecht: Kluwer Acad.

Lakoff, G. 1977. Linguistic gestalts. CLS 13:236-87.

Langacker, R. W. (1987) Foundations of cognitive grammar: Theoretical Prerequisites. Stanford, CA: Stanford University Press.

Larson, R. 1995. Semantics. Invitation to Cognitive Science Vol. 1 Language ed. by L. Gleitman. Cambridge: MIT Press

Levin, B. 1993.English verb classes and alternations: A preliminary investigation. Chicago: University of Chicago Press

Lidz, J. 1998. Constraints on the Syntactic Bootstrapping Procedure for Verb Learning. Proceedings of the 22nd Annual Boston University Conference on Language Development, I-II ed. by A. Greenhill, M. Hughes, H. Littlefield, and H. Walsh, 488-

98. Somerville, MA: Cascadilla

Lidz, J. & Gleitman, L.R. 2004. Yes, we still need universal grammar; reply to Goldberg, A. (2004). Cognition.

Marantz, A. 1984. On the Nature of Grammatical Relations. Cambridge: MIT Press

Marantz, A. 1988. Clitics, morphological merger, and the mapping to phonological structure. In Theoretical morphology, ed. by M. Hammond and M. Noonan, 253–270. New York: Academic Press.

Marantz, A. 1997. 'Stem suppletion, or the arbitrariness of the sign.' Talk given at the Universite' de Paris VIII

Marantz, A. 1997b. No Escape from Syntax. Don't Try Morphological Analysis in the Privacy of Your Own Lexicon Proceedings of the 21st Annual Penn Linguistics Colloquium (U. Penn Working Papers in Linguistics 4:2) ed. by A. Dimitriadis, L. Siegel, C. Surek-Clark, and A. Williams, 201-225.

Naigles, L., H. Gleitman, L. Gleitman. 1993. Syntactic Bootstrapping and Verb Acquisition. Cognition and Language. A Developmental Perspective ed. by Esther Dromi. Norwood: Ablex.

Noyer, R. 1997. Features, Positions and Affixes in Autonomous Morphological Structure. New York: Garland.

Oehrle, R. 1976. The Grammatical Status of the English Dative Alternation, Ph.D dissertation, MIT.

Parsons, T. 1990. Events in the Semantics of English. Cambridge, MA: MIT Press.

Pesetsky, D. 1995. Zero syntax: experiencers and cascades. MIT Press, Cambridge.

Pfau, R. 2000. Features and Categories in Language Production. Ph.D dissertation, University of Frankfurt

Pinker, S. 1989. Learnability and cognition: the acquisition of argument structure. Cambridge, Mass. : MIT Press

Prince, A. and P. Smolensky 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Rutgers University Center for Cognitive Science Technical Report

Pollard, C. and I. Sag. 1994. Head-driven phrase structure grammar. Chicago: University of Chicago Press

Ramchand, G. 1997. Aspect and Predication Oxford Press

Ramchand, G. 1998. Deconstructing the Lexicon The Projection of Arguments: Lexical and Compositional Factors ed. by M. Butt, and W. Geuder. Stanford, CA: Center for the Study of Language and Information viii

Ramchand, G. 2003. First Phase Syntax. Ms. University of Oxford.

Rizzi, L. 1990. Speculations on verb second. In Mascar'o, J. and Nespor, M., editors, Grammar in Progress, pages 375--386. Foris, Dordrecht

Ritter, E. and S. Rosen. 1998. Delimiting events in syntax. The Projection of Arguments: Lexical and Compositional Factors ed. by M. Butt, and W. Geuder. Stanford, CA: Center for the Study of Language and Information viii

Sproat, R. 1985. On Deriving the Lexicon. Ph.D Dissertation MIT

Siddiqi, D. (2004). Obtaining lexical contrast in Optimality Theory. Ms, University of Arizona.

Travis, L. 1994. Event Phrase and a Theory of Functional Categories. Paper presented at the Canadian Linguistics Association, University of Calgary.

Urbanczyk, S. 1999. Reduplicative Size-segmentism Correlations as Root-affix Asymmetries. ms. Rutgers University, New Bruswick.

Woolford, E. 1984 “Dative Verbs with Unspecified Objects” Ling. Rev., 1984 June; 3 (4): 389-409