Structural Competition in Grammar

by

Roni Katzir

Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2008

Author .....
Department of Linguistics and Philosophy
August 15, 2008

Certified by .....
Danny Fox
Professor of Linguistics
Thesis Supervisor

Accepted by .....
Irene Heim
Professor of Linguistics
Department Head

Structural Competition in Grammar

by

Roni Katzir

Submitted to the Department of Linguistics and Philosophy
on August 15, 2008, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis makes the following three claims:

(1) Competition exists in natural language: the grammaticality (and meaning) of using a linguistic object ϕ can be affected by the grammaticality (and meaning) of a different linguistic object ϕ'

(2) Structure plays a role in competition: ϕ' can only affect the grammaticality (and meaning) of ϕ if ϕ and ϕ' are structurally related (in particular, if ϕ' is no more complex than ϕ)

(3) Simpler is better: if ϕ is strictly more complex than ϕ', and if the two are equally good otherwise, ϕ will be blocked by ϕ'

The first claim is the most general and the least controversial. It adds little to what is commonly accepted in the domains of conversational implicature, focus alternatives, morphological blocking, and economy conditions in syntax and semantics. Chapter 1 presents background on some of the issues regarding this general claim. The second claim is more controversial. Most work on implicature has treated considerations of structural complexity as unimportant or downright orthogonal to conversational reasoning. In the domain of focus alternatives structure has been occasionally used (in particular, below the word level), but argued to be irrelevant otherwise. In Chapter 2 I will present a case study that shows that, at least sometimes, reference to structure (specifically to structural complexity) is necessary. Chapter 3, jointly written with Danny Fox, discusses a remaining question about the use of alternatives for implicature and provides arguments for a parallel treatment of implicature and focus, as well as for a constraint on the ability of contextual relevance to remove a formal alternative from the set of actual alternatives. In Chapter 4 I discuss certain cases of morphological blocking that cannot be based solely on structural pruning. For the patterns to be accounted for, a direct preference for simpler structures must be active in the grammar.

Thesis Supervisor: Danny Fox
Title: Professor of Linguistics

Acknowledgments

First and foremost, I wish to thank Danny Fox, my thesis supervisor, and Irene Heim, Sabine Iatridou, and David Pesetsky, my remaining committee members and unofficial advisors. In addition to teaching me more about linguistics than I ever thought existed, and in addition to being kinder, more supportive, and more generous with their time and ideas than I could ever hope to understand, this dream committee has offered me a personal example of how to be a linguist. From them I learned that one should strive to construct the most charitable reading of any proposal one wishes to argue against, while remaining critical towards one's own. They also taught me to treat conceptual arguments with suspicion and empirical ones with respect. If the reader finds anything of value in this work, Danny, David, Irene, and Sabine are the people who should be thanked for it. At least as significantly, they also deserve credit for filtering out an inordinate amount of bad ideas that I came up with over the years and that would have made their way into the final version otherwise.

From a technical point of view, Alec Marantz was not a member of my committee, but in every meaningful sense he was. Alec almost turned me into a morphologist, by showing me how much of morphology is actually syntax. Most directly, I thank him for triggering and accompanying what has become chapter 4 of this dissertation. But Alec also helped me see how much of morphology is not part of syntax, making it easier for me to remain a syntactician. In that, my debt of gratitude to him extends to just about every other project I have worked on.

And I thank Donca Steriade. When I started the program, I had no doubt that I would avoid phonology at all costs. That was before I met Donca. For reasons that remain unclear to me, Donca was willing to smile away my ignorance and attitude, and, from the beginning of the first year and up until the end, she has walked me through treacherous mazes of phonological marshlands in which I would have surely disappeared if I were to travel them on my own. Some of the results appear in chapter 4 of this dissertation, but what I owe her is much more than that.

While working on the topics that ended up in this dissertation I benefitted from the help

and feedback of an extraordinary number of extraordinarily smart people. From this very long list, particularly special thanks go to Klaus Abels, Márta Abrusán, Adam Albright, Artemis Alexiadou, Elena Anagnostopoulou, Jonathan Bobaljik, Emmanuel Chemla, Gennaro Chierchia, Noam Chomsky, Michel DeGraff, Paul Égré, David Embick, Kai von Fintel, Justin Fitzpatrick, Morris Halle, Peter Jenks, Jonah Katz, Michael Kenstowicz, Nathan Klinedinst, Idan Landau, Fred Landman, Winnie Lechner, Thomas Leu, Victor Manfredi, Luisa Martí, Andrew Nevins, Tim O'Donnell, Masha Polinsky, Milan Rezac, Dorian Roehrs, Erik Schoorlemmer, Ivy Sichel, Tal Siloni, Giorgos Spathas, Benjamin Spector, and Michael Wagner.

Polly Jacobson and two anonymous *Linguistics & Philosophy* reviewers have offered extremely helpful comments and suggestions on what has become chapter 2 of this dissertation.

Of all those who helped me with their morphological judgments, I am particularly indebted to María Ágústsdóttir, Artemis Alexiadou, Margrét Bjarnadóttir, Michael Fortescue, Jóel Friðriksson, Kim Grubert, Sabine Iatridou, Line Mikkelsen, Jan Rijkhoff, Giorgos Spathas, Lýður Thorgeirsson, and Torben Thrane.

I owe my conversion to linguistics to Shimon Edelman, MarcAriel Friedemann, Ido Grabinsky, Julia Horvath, Fred Landman, Aynat Rubinstein, Galit Sassoon, Aldo Sevi, Marit Shternau, and especially Tal Siloni.

Boris Katz made it possible for me to keep working on computation and pay the rent during the summers. Daniel Loreto made that work much more interesting. Federico Mora made it fun.

For helping me stay alive and sane during the past few years I thank Hiba Bou Akar, Jimmy Carlson, Nat Kuhn, and Dana Price. And also Aliya from Baraka Café and Salaam from Andala.

I am grateful to Chrissy Graham and Jen Purdy for being so helpful and so friendly, and for shielding me from some scary parts of the real world.

I thank Asaf Bachrach for collaboration, ideas, food, and company, as well as for a demonstration that life, or something very much like it, is possible. It is Asaf, more than anyone else, who turned me from a noisy and confused outsider who thinks he knows all

the answers to a noisy and confused insider who knows that he doesn't.

Over the past five years, I spent more time talking to Raj Singh than doing any other single activity. If I were offered the chance to go back, I would increase that time even more. Without Raj, my first four years in the program would have been so much duller. My fifth year would have been impossible. There is more to be said, but I cannot say it here without offending some people and embarrassing the rest. He knows, and that is enough.

I thank my close family: Nurit, Gadi, Nim, Yair, Karin, Noa, Guy, and Dror. Knowing they were far away has been difficult, but knowing they were somewhere helped.

I will not even try to express my gratitude to Ivona Kučerová and Pnina Rechavi. I am dedicating this work to them instead.

Contents

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 1

Introduction

1.1 Competition in grammar

This thesis attempts to investigate the role of structure in determining the form of competition in grammar. My main claims will be that structure plays a role in restricting the alternatives that are used for competition, and that alternatives that are structurally more complex do not affect the outcome of the competition. More tentatively, I will try to show that, in certain cases, competition in grammar can involve a preference for simple structures over complex ones. These claims all presuppose that grammar involves competition, so before we start we should convince ourselves that this presupposition is satisfied, and that (1) receives a positive answer.

(1) To determine whether a structure ϕ belongs to a language L, is it ever necessary to check for some $\phi' \neq \phi$ whether $\phi' \in L$?

In the early stages of Generative Grammar, (1) would receive a negative answer. A grammar G, as in (2), provided a recipe for generating structures:

(2) G:

$$
\begin{aligned}
S &\rightarrow NP\ VP \\
NP &\rightarrow Det\ N \\
NP &\rightarrow Name \\
VP &\rightarrow V\ NP \\
VP &\rightarrow V\ PP \\
&\ldots
\end{aligned}
$$

A structure ϕ was in L_G if and only if one could start with the initial non-terminal S and proceed by a series of expansions using rules in G (followed by various transformations). Derivatively, a string w was considered to be in the language L_G generated by G if and only if w was the concatenation of the terminals of a structure ϕ generated by G. Crucially, checking whether some other structure ϕ (or some string w') is also generated by G is irrelevant. This does not change even if we look at derivations rather than structures and strings. For example, choosing to expand NP to $Det\ N$ depends only on whether this production is in G. The presence of $NP \rightarrow Name$ in G is irrelevant for this choice. Neither the elements of G nor the conventions of use can make reference to structures.

This has been the traditional convention with respect to grammars and derivations, and it is a very natural one. Using a grammar G will involve keeping track of the current derivation, as well as checking whether particular rules are listed; on the traditional approach, this summarizes the derivational process.[1] Of course, one can imagine alternative conventions. Here are some possible, though unmotivated conventions:

(3) a. Never use any expansion of X more than 2 more times than any other expansion of X in one derivation, or

[1]This is not entirely correct. Devices such as obligatory transformations and rule ordering, present already in Chomsky (1957, pp. 44-45), made it possible to rule out an otherwise grammatical string by *adding* a rule to the grammar. The phenomena for which devices of this kind were proposed have later been taken to support a competition-based approach (as in Kiparsky, 1973, which we will discuss shortly). The devices themselves can be straightforwardly restated in terms of competition.

b. Choose only one expansion of X in a single derivation

c. ...

The convention we are interested in here is one of competition. It uses a grammar, along with the traditional convention to derivation, to which it adds a notion of *comparison* between structures, as well as the idea that an otherwise grammatical structure ϕ can be ruled out if there is a structure ϕ' that is strictly better than ϕ. Using $\precsim$ for the relation of `better-than` and $\prec$ for the irreflexive `strictly-better-than`, we can state the new convention as follows.[2]

(4) COMPETITION CONVENTION: if the traditional convention derives both ϕ and ϕ' from G, and if $\phi' \prec \phi$, then $\phi \notin L_G$

Within generative grammar, the introduction of competition as part of the architecture is largely due to Anderson (1969b), Kiparsky (1973) and Aronoff (1976), who provided evidence for competition in the domain of morpho-phonology. A closely related notion was formulated by Grice (1967/1975) in the domain of language use.

Let us look at an example of a distributional pattern that could be used to motivate (4).

1.1.1 The English past tense

To generate the past tense of English verbs, a mechanistic approach along the lines of (2) would perhaps divide verbs into regular verbs, such as *walk*, *talk*, *jump*, and so on, which take *-ed* as their past tense suffix, and irregular verbs, such as *put*, *go*, *sing*, and so on, which do not:[3]

(5) A grammar for the complementary distribution of regular and irregular forms of

[2] $\precsim$ is a transitive, reflexive relation (or *pre-order*) over structures, and $\prec := \precsim \setminus \precsim^{-1}$. In the examples below, it will often be convenient to define a comparison in terms of $\precsim$ but to state the competition in terms of $\prec$.

[3] Of course, regularities within the class of irregular verbs can be used to simplify the grammar by reducing the storage for listing rules for individual stems. And the use of features, as in Chomsky (1965), can reduce the duplication of statements about distinct categories. These important issues are orthogonal to the question of competition.

the past tense:

$$
\begin{aligned}
V_{past} &\rightarrow V1_{stem} - ed \\
V_{past} &\rightarrow V2_{stem} \\
V_{past} &\rightarrow went \\
&\ldots \\
V1_{stem} &\rightarrow walk \mid talk \mid \ldots \\
V2_{stem} &\rightarrow put \mid \ldots \\
&\ldots
\end{aligned}
$$

The possible affixes are in (near-)complementary distribution. A given stem usually has only one way to form the past tense. In particular, when an irregular past tense is possible, the regular ending is not:[4]

(6) *put (*put-ed), went (*go-ed), sing (*sing-ed)*

Competition is a state of affairs in which the availability of one form can make another form ungrammatical. For the past form in English, we can state a relation of `better-than` for the past tense, written as $\preceq_{pst}$ ($\prec_{pst}$ for *strictly-better-than*) that is based on lexical listedness: an irregular form is better than a regular one. So $put \prec_{pst} put\text{-}ed$, for example, and $went \prec_{pst} go\text{-}ed$. If competition holds for past formation in English, the complementarity of the regular and the irregular forms is no accident: it is the very possibility of generating an irregular form like *went* that makes its regular counterpart **go-ed* impossible.

A phrase-structure grammar can describe a pattern that involves complementarity, of course, as (5) above shows, but it could just as easily describe a pattern in which the irregular form supplements the regular form, rather than supplant it:

[4]Some exceptions exist. For example, *dive* has both *dived* and *dove* as possible past-tense forms, though there might be dialectal variation in the choice, and *hang* has both *hung* and *hanged*, though the two forms differ in meaning.

(7) A grammar for a hypothetical, non-complementary past-tense pattern:

$$
\begin{aligned}
V_{past} &\rightarrow V_{stem} - ed \\
V_{past} &\rightarrow V2_{stem} \\
V_{past} &\rightarrow went \\
&\ldots \\
V_{stem} &\rightarrow V1_{stem} \mid V2_{stem} \mid go \mid \ldots \\
V1_{stem} &\rightarrow walk \mid talk \mid \ldots \\
V2_{stem} &\rightarrow put \mid \ldots \\
&\ldots
\end{aligned}
$$

(7) generates a language where all stems can form the past tense using *-ed*, and some stems, like *put*, *go*, and *sing* have an additional, irregular past form. In this sense, the competition-free approach treats the complementarity of the irregular and the irregular forms in English as an accident. In practice, however, the complementary pattern in (5) seems to be quite widespread. Here are some familiar examples:

(8) Complementarity with other affixes:

a. *boy – boy-s, girl – girl-s, ...*

b. *child – child-ren (*child-s), man – men (*man-s), ...*

(9) Complementarity with periphrastic forms:

a. *intelligent – more intelligent, upset – more upset, ...*

b. *good – better (*more good)*

Significantly, many of the phenomena that involve complementarity exhibit an asymmetry between the possible outcomes. If there are two possible outcomes, the distribution of one would be statable on its own, while the distribution of the other will be more naturally defined as the complement of the first. This asymmetry was argued by Kiparsky (1973) to be fundamental to the application of phonological rules, justifying a principled treatment. The principle Kiparsky proposes for capturing the asymmetry is his *Elsewhere*

Condition, which he formulates as follows, noting the long history of similar ideas:

(10) The Elsewhere Condition (Kiparsky, 1973, (4), p. 94)

Two adjacent rules of the form

$A \rightarrow B \,/\, P\_Q$

$C \rightarrow D \,/\, R\_S$

are disjunctively ordered if and only if:

a. the set of strings that fit PAQ is a subset of the strings that fit RCS, and

b. the structural changes of the two rules are either identical or incompatible

We have already seen examples of this *elsewhere* asymmetry in the domain of affixation, where the different status of the possible outcomes is suggested by the terms *regular* and *irregular*. The distribution of irregular forms is often stated either extensionally, by listing the relevant enviroments, or intensionally, in terms of morpho-phonological properties of those environments. The distribution of regular forms, on the other hand, is more naturally described as those environments that do not belong to the environments of the irregular form.

A slightly different example of an elsewhere asymmetry is provided by the paradigm of the present tense forms in English. Of the two possible endings, *-s* and ∅, the former is most straightforwardly specified as the form for 3rd person singular, while the latter is more naturally specified as appearing everywhere else. The asymmetry in these cases is not simply an artifact of which of the forms we chose to start with. For regular and irregular past tense or plural forms, the usual choice results in a considerably more compact description than the opposite direction, in which all the stems that take a regular ending are listed. In the case of the present tense paradigm, the entries for ∅ cross the lines of person and gender, so characterizing them directly makes less sense than treating them as the complement of the environment for *-s*.

The *elsewhere* signature is not limited to morpho-phonology. A remarkably similar pattern was observed and discussed by Hawkins (1991) and Heim (1991) for the use and meaning of the definite and indefinite articles in a language like English. Simplifying somewhat, using the definite article in the singular is subject to a uniqueness requirement:

the big dog is in the corner can only be used felicitously if there is exactly one big dog that is salient in the context. At first sight, using the indefinite article seems to be subject to a *non-uniqueness* requirement: when the referent of the noun phrase is clearly a singleton, using the indefinite article is infelicitous, as the following example from Heim (1991) shows.

(11) The/#A weight of our tent is 4lb. (=Heim's (117))

As Heim shows, however, there is an asymmetry between the uniqueness and the non-uniqueness requirements. The uniqueness requirement associated with the definite article seems to be a real part of its meaning. The indefinite article, on the other hand, seems to be an *elsewhere* condition, used when the definite article is impossible, including in certain cases where non-uniqueness is not implied:

(12) a. Robert caught a 20ft. catfish (=Heim's (121))

b. A pathologically noisy neighbor of mine broke into the attic (=Heim's (122))

(12a) can be used without suggesting that there was more than one 20ft. catfish, and (12b) can be used without suggesting the existence of other pathologically noisy neighbors. Rather, what makes the indefinite article possible in these sentences is the inappropriateness of the definite article. As Heim (1991) proposes, this would follow if the indefinite article were not specified with respect to uniqueness, and if an ordering principle required the definite article to be used when possible:

(13) In utterance situations where the presupposition for *[the ζ] ξ* is already known to be satisfied, it is not permitted to utter *[a ζ] ξ*

Heim notes that (13) (= her (123)) bears resemblance to the Gricean notion of competition, but that unifying the two phenomena faces certain challenges.[5] Instead, she points out a more conservative generalization of (13) to a principle, referred to as *Maximize Presupposition!* in later literature, requiring a presuppositional entry to be used when possible. This principle has been argued to derive *elsewhere* asymmetries in the domain of tense semantics (Sauerland, 2002), number (Sauerland, 2003b), and more generally for binary feature distinctions (Sauerland, 2003a).

[5]The main difficulty has to do with the status of world knowledge within a conversational setting. See Magri (2006) for further discussion.

In various forms, complementarity and the *elsewhere* signature have been argued to be a fundamental aspect of many other phenomena in grammar. As mentioned, these patterns could be treated as accidents, resulting perhaps from diachronic pressure or other considerations that are not directly related to grammar. Most frameworks, however, tend to agree that the *elsewhere* signature is not entirely accidental, and that synchronic grammar involves some kind of competition. In other words, the consensus seems to be that our earlier question in (1) should be answered to the affirmative:

(14) At least in some cases, the question of whether $\phi \in L$ depends on whether for some $\phi' \neq \phi, \phi' \in L$

The disagreement is about the scope of this competition. On some accounts, the comparison of whole structures can be reduced to local comparisons within smaller domains. In our example of the past tense in English we already assumed such a factorization, if also implicitly. Grammaticality is a property of complete outputs of the grammar, but we stated $\precsim_{pst}$, the comparison between regular and irregular past tense forms, in terms of single words. The ordering $\precsim$ among whole structures can be made to be dependent on the local $\precsim_{pst}$: if $\phi[\pi_1, \ldots, \pi_n]$ is a structure in which $\pi_1, \ldots, \pi_n$ are all the occurrences of past tense forms, and if $\pi'_1 \precsim \pi_1, \ldots, \pi'_n \precsim \pi_n$, then $\phi[\pi'_1, \ldots, \pi'_n] \precsim \phi[\pi_1, \ldots, \pi_n]$. For cases like the English past tense, where comparison of the whole structures can be reduced to a conjunction of local comparisons, we will say that the competition is *local*. When such a factorization is not available we will say that the competition is *global*. In section (1.2) I will review some of the discussion in the literature on the scope of competition in grammar. Locality *per se* will not be my primary focus in the coming chapters. However, issues related to locality will come up throughout the dissertation in various contexts, and it will be convenient to cover this background first.

1.2 The scope of competition

1.2.1 Global competition

An early, globalist, and highly influential competition account was presented by Grice (1967/1975) and developed in subsequent work by Horn (1972), Gazdar (1979), Hirschberg (1985/1991) and others.[6] On Grice's account, participants in a conversation can use general reasoning to convey and infer meaning components of an utterance that go beyond the semantic content of that utterance. For example, (15) entails, as part of its semantic content, that John talked to either Mary or Sue:

(15) John talked to Mary or Sue

As far as the semantic content is concerned, the sentence can be true if John talked to both Mary and Sue. However, under normal circumstances, a speaker who utters (15) suggests that John did not talk to both Mary and Sue. The speaker also suggests that they do not know which of Mary and Sue John talked to. Informally, these suggestions are arrived at via competition with other possible utterances, each of which is in some sense strictly better than ϕ =(15).[7]

(16) a. ϕ'=John talked to Mary

b. ϕ''=John talked to Sue

c. ψ=John talked to Mary and Sue

The sense in which the sentences in (16) are strictly better than (15) involves considerations of semantic entailment and assertability in a given context. Each of the sentences in (16) asymmetrically entails (15). Consequently, each of them, if true and assertable otherwise in the current context C, would make a more informative utterance than (15). The relations $\precsim$ and $\prec$ here, then, should be stated between objects that include a representation of assertability in context C. Let us use ! to denote assertability, writing $!(\phi)(C)$ to represent

[6]Grice's proposal and the later developments will occupy us in the coming chapters. The current section provides only a very rough outline.

[7]See Sauerland (2004) for discussion. We will come back to this topic in chapter 2.

the assertability of ϕ in the context C. We can now state the comparison that is relevant for evaluating ϕ as follows:

(17) a. $!(\phi')(C) \prec !(\phi)(C)$

b. $!(\phi'')(C) \prec !(\phi)(C)$

c. $!(\psi)(C) \prec !(\phi)(C)$

On the assumption that competition results in blocking, the ordering in (17) means that the assertability of ϕ in the context of C, $!(\phi)(C)$, would be blocked by the assertability of any of the alternatives in (16). Since (17) was asserted, the hearer can conclude that none of the alternatives was assertable:

(18) a. $\neg!(\phi')(C)$

b. $\neg!(\phi'')(C)$

c. $\neg!(\psi)(C)$

The conclusions in (17) amount to the informal, extra-semantic suggestions we listed above: the speaker does not hold the belief that John spoke to Mary, the speaker does not hold the belief that John spoke to Sue, and obviously the speaker does not hold the belief that John talked to both Mary and Sue.[8]

What makes Grice's account an instance of global competition is that the comparison we used depended on entailment and assertability at the level of whole sentences. If we tried to represent the global comparison as the conjunction of local ones, we would encounter difficulties right away. Consider, for example, the idea of factorizing each assertion into finite clauses. Our earlier examples involved just one such clause, so there would be no difference between the global comparison and the local one. Embedding (15) allows us to see how the two comparisons differ.

(19) If John talked to Mary or Sue, he must already know about the party

(20) Every boy who suspects that John talked to Mary or Sue should alert the headmaster

[8]Accounting for the suggestion that the speaker believes that John did not speak to both Mary and Sue, equivalent to $!(\neg\psi)(C)$ in our notation, is more involved. See van Rooij and Schulz (2004), Sauerland (2004), Fox (2006), and Spector (2006b) for important discussion.

In (19) and (20), our earlier (15) is embedded in a downward entailing (henceforth D.E.) context, where entailment relations are reversed. This means that, from the perspective of global comparison, the sentences in (16), which used to be better alternatives than (15) as assertions at the root level, will give rise to weaker, and hence worse assertions in the embedded context. Using $O_{DE}(\cdot)$ to represent one of the embedding contexts, we have the following:

(21) Global comparison and embedding in a D.E. context: reversal of $\prec$

a. $!(O_{DE}(\phi))(C) \prec !(O_{DE}(\phi'))(C)$

b. $!(O_{DE}(\phi))(C) \prec !(O_{DE}(\phi''))(C)$

c. $!(O_{DE}(\phi))(C) \prec !(O_{DE}(\psi))(C)$

Since competition looks only for better candidates, $O_{DE}(\phi')$, $O_{DE}(\phi'')$, and $O_{DE}(\psi)$ will be irrelevant for the evaluation of $O_{DE}(\phi)$.[9]

In the local approach that we mentioned, whole assertions are broken into individual finite clauses, so $O_{DE}(\phi)$ will be broken into O_{DE} =*If x, he must already know about the party* and ϕ =*John talked to Sue or Mary*. By comparing each clause on its own, we will obtain the relation of *better-than* for a whole structure as a conjunction of the relations of *better-than* for the individual clauses of which that structure is built. For the embedding context, we have $!O_{DE} \precsim !O_{DE}$ (pre-orders are reflexive, so every structure is as good as itself). For the embedded clauses, the original `better-than` relations in (17) above still hold:

(22) a. $!(\phi')(C) \prec !(\phi)(C)$

b. $!(\phi'')(C) \prec !(\phi)(C)$

[9]On the other hand, in the context of $O_{DE}(\cdot)$, what used to be stronger alternatives should now suggest that what used to be the weaker alternative could not have been used. This is true for $O_{DE}(\psi)$: *If John talked to (both) Mary and Sue, he must already know about the party* suggests that *If John talked to Mary or Sue, he must already know about the party* could not have been used. For the single disjuncts, on the other hand, no such inference is made: under normal circumstances, *If John talked to Mary, he must already know about the party* suggests nothing about whether talking to Sue would have had the same result. The reason, I believe, is that $\precsim$ incorporates more than just semantic entailment, and that structural complexity also plays a role. I will discuss this idea in detail in the coming chapters.

c. $!(\psi)(C) \prec !(\phi)(C)$

Conjoined, we obtain an ordering of the complete structures that follows (22) and is the opposite of (21)

(23) Local comparison and embedding in a D.E. context: no reversal of $\prec$

a. $!(O_{DE}(\phi'))(C) \prec !(O_{DE}(\phi))(C)$

b. $!(O_{DE}(\phi''))(C) \prec !(O_{DE}(\phi))(C)$

c. $!(O_{DE}(\psi))(C) \prec !(O_{DE}(\phi))(C)$

In this case, then, the local approach we have considered seems to do worse than the global one. This does not show, of course, that other decompositions would also fail.[10] A more general argument in favor of the global account comes from the fact that, while a local comparison of the kind we saw above seems like an arbitrary stipulation, the global comparison seems to fall naturally out of assumptions about cooperative behavior in conversation. Grice offered a preliminary set of such assumptions, presented as *Maxims of Conversation* which specify the behavior of cooperative speakers in an unmarked state of affairs:

(24) Grice's Maxims of Conversation:

a. Quantity

i. Make your contribution as informative as is required

ii. Do not make your contribution more informative than is required

b. Quality

i. Do not say what you believe to be false

ii. Do not say that for which you lack adequate evidence

c. Relation

i. Be relevant

d. Manner

[10]In fact, problems for the global account have been pointed out as early as Cohen (1971). Recently, arguments in favor of a local approach have been provided by Landman (2000), Chierchia (2004), Fox (2006), Chierchia et al. (2008), and others. We will discuss some of the relevant issues in the following chapters.

i. Avoid obscurity of expression

ii. Avoid ambiguity

iii. Be brief

iv. Be orderly

In our discussion above, we stated $\preceq$ in terms of semantic entailment, reflecting one way in which (24a-i) has been interpreted (see Gamut, 1991), and we stated the notion of assertability in terms of speaker's belief, reflecting (24b-i).

A maxim that we will pay special attention to is the brevity condition in (24d-iii). Brevity has sometimes been proposed as an explanation for certain *blocking* phenomena, where competition arises between elements that mean the same thing but are not equally good in terms of their form.[11] In our earlier examples from English morphology, we may note that the irregular forms are often shorter than the regular ones. In the past tense, for example, *put* is shorter than **put-ed*, and *sang* is shorter than **sing-ed*. In plural formation, *men* is shorter than **man-s*. And in the comparative, *smarter* is shorter than **more smart*. Above, when we discussed these examples, we considered the idea of ordering irregular (or listed) forms as better than regular ones. Could we replace this seemingly arbitrary ordering with the more natural brevity condition in (24d-iii)?

As Poser (1992) shows, there are reasons to doubt whether reducing blocking to the Gricean notion of brevity could work. First, while some irregular forms are indeed shorter than their (blocked) regular counterparts, say when measured in terms of phonetic effort, in other cases the regular form is as short as the irregular one, as in the pair *oxen, *oxes*. In some cases, such as *child-ren, *childs*, the regular form is arguably shorter than the irregular one.

Poser notes two further problems for what he refers to as the pragmatic approach. First, the judgments for supposed violations of Grice's Maxim of Manner seem weaker than the intuitively much clearer grammaticality contrasts in blocking effects: the ill-formedness of *pale red* (presumably because of the availability of *pink*; cf. Householder, 1971) is nothing in comparison with the ungrammaticality of **man-s* (supposedly because of the availability

[11] I am using the term *blocking* in a somewhat different sense from that of Aronoff (1976), who introduced the term. I will say more about Aronoff's proposal in section 1.2.2.

of *men*). Second, Poser points out that, contrary to what a pragmatic account might predict, the evaluation for blocking seems to be subject to both syntactic and semantic restrictions: the domains in which the candidates for blocking differ are small, and the semantic computations that determine identity are simple. For example, *the red book* does not block *the book which is red*, and *John is smarter than Tom* may block **John is more smart than Tom* but it does not affect the acceptability of *John's intelligence exceeds Tom's* or *John has greater intelligence than Tom*. A related issue, raised and discussed by Horn (1978, 1984), is that allowing free competition of forms, as is suggested by pragmatic accounts such as McCawley (1978)'s, leads us to expect a much broader range of blocking effects than we actually find. For example, Horn notes that if, as has been suggested, *John moved the plate* blocks the unmarked reading of *John made the plate move*, we would expect *John dropped the plate* to block the unmarked reading of *John made the plate fall*, contrary to fact.

Further objections to the global perspective are raised by Embick and Marantz (2008), who discuss Bresnan (2001)'s account of contraction in English. One of the characteristics of competition, as we saw earlier, is that a candidate that would be ungrammatical in the presence of a better competitor can become grammatical if a better competitor is unavailable. In the case of regular and irregular past tense forms, for example, we mentioned a competition account in which the irregular form is in some sense better than the regular form, allowing a form like *sang* to block the regular form *sing-ed*. When an irregular form is unavailable, as in the case of *jump*, no better candidate than *jump-ed* is available, so the regular form becomes grammatical. Bresnan uses the absence of **amn't* from the lexicon of Standard English to derive the possibility of certain forms that would otherwise be ungrammatical or dispreferred. She notes, for example, that while the form *Is he not working?* is dispreferred to the contracted *Isn't he working?*, as in (25), the first person variant *Am I not working?* is possible, correlating with the impossibility of **Amn't I working?*, as shown in (26). Similarly, the less faithful *Aren't I working?*, in which the form of the auxiliary fails to express the features of the subject, is limited to the first person singular, and it is not available when a more faithful form exists, hence **Aren't he working?*. [12]

[12]Bresnan (2001) works within the framework of Optimality Theory (OT, Prince and Smolensky, 1993). In (25) and (26), LEX is a constraint that penalizes unfaithfulness to the lexicon, as in the use of the form

(25)

	LEX	*NEG-C	*NEG-VP
☞ *Isn't he working?*			
Is not he working?		*!	
Is he [not working]?			*!

(26)

	LEX	*NEG-C	*NEG-VP
Amn't I working?	*!		
Am not I working?		*!	
☞ *Am I [not working]?*			*

As Embick and Marantz (2008) note, Bresnan's account makes problematic predictions even for the forms that she discusses. For example, they point out that speakers do not seem to find a difference between the acceptability of *Is he not working?* and that of *Am I not working?*, surprisingly for Bresnan's account in which the former loses in a competition and the latter wins. Furthermore, they observe that lexical gaps in general do not license the emergence of an otherwise dispreferred analytical form: the absence of a past form for *forgo* (**forgoed, *forwent*) does not license **did forgo*; and the absence of a participial form for *stride* (**stridden*) does not license **done stride*.

The general problem with global accounts such as Bresnan's is that they allow too many candidates to enter the competition. When difficulties arise in one place in the structure,

---

amn't*; *NEG-C penalizes analytic negation adjoining to the complementizer, which Bresnan assumes takes place in **Is not he working?* but not in *Isn't he working?*; and *NEG-VP penalizes analytic negation adjoining to *VP*, as in *Am I [not working]?* For purposes of presentation, I have simplified Bresnan's discussion, omitting certain additional constraints and candidates that Bresnan considers. It is also worth noting that OT offers a perspective on competition that is different from the one that we have been assuming. In particular, we have assumed that the various pre-orders $\{\preceq_\alpha\}_{\alpha \in I}$ that combine to provide the pre-order $\preceq$ over whole structures that is relevant for competition are themselves unordered. The pre-order $\preceq$ is the intersection of all the local pre-orders: $\preceq = \bigcap_{\alpha \in I} \preceq_\alpha$. On our assumptions, then, a form like ϕ_1 =John sing-ed a better song than Bill*, which contains a losing past-tense candidate and a winning comparative, is neither better nor worse than ϕ_2 =**John sang a more good song than Bill*, which has a winning past-tense candidate and a losing comparative. In OT, on the other hand, different constraints are assigned different importance, allowing ties such as the one between our ϕ_1 and ϕ_2 to be broken. For our current discussion, I believe that these differences can be ignored.

as is the case for lexical gaps, there will usually be candidates that are different enough to avoid the difficulty but are otherwise suboptimal. For global comparison, these candidates will often be considered, leading to the prediction that one of them would win. As the objections of Horn (1978, 1984), Poser (1992), and Embick and Marantz (2008) make clear, this prediction fails to match observed patterns of blocking effects. While there is a sense in which one form can be considered better than another, as in the English past, where the availability of an irregular form rules out the regular form, and as in the English comparative, where a synthetic form like *smarter* rules out the analytic **more smart*, the set of candidates that enter the competition must be highly restricted. For the past tense, even the intuitively local *do*-support variant seems to be excluded from the set of alternatives (as the impossibility of **did forgo* shows). And for the comparative, *smarter* might be relevant for **more smart* but it does not seem to play a role for some of the more complex paraphrases: *John's intelligence is higher than Bill's* is odd, perhaps, but not ungrammatical.

Let us now move on to two kinds of competition accounts where the set of alternatives is restricted enough to avoid the problems of global comparison.

1.2.2 Intermediate competition and the lexicon

For Aronoff (1976), the locus of competition is the lexicon: a listed item occupies a 'slot' that prevents other items from being listed. Regularly formed, unlisted items, are not subject to this restriction. Thus, the listed form *glory* prevents **gloriosity*, which would need listing, from being available; the regularly-formed *gloriousness*, on the other hand, does not need to be listed and is consequently exempt from blocking. Similar considerations apply to *fury* and **furiosity*. For adjectives like *various* and *curious*, for which there is no counterpart to *glory* and *fury*, blocking does not occur: the relevant lexical slot is free to list the *-ity* form, which now appears alongside the regular *-ness* form.

(27) Listedness blocking Aronoff, 1976

XOUS	∅	-ITY	-NESS
glorious	glory	*gloriosity	gloriousness
furious	fury	*furiosity	furiousness
various	*vary	variety	variousness
curious	*cury	curiosity	curiousness

Aronoff's proposal has been influential in the literature on competition, but in a sense it does not really involve competition in grammar at all. The generalization that Aronoff wants to account for is the near-complementary distribution of two *listed* forms, and his proposal makes crucial use of the idea of lexical slots. Regularly-formed items, such as nominals derived by affixation of *-ness*, are irrelevant to blocking. Competition, as we defined it earlier, involves the generation of the losing form to a better one. In Aronoff's account, the loser is not generated. If **gloriosity* were generated, for example, it would not have to be listed in the lexicon and would consequently be exempt from blocking. It seems more natural, then, to think of Aronoff's blocking as a meta-constraint on the form of grammar (where the grammar contains also the lexicon) or on learning. The main reason I mention this account in the current context is because of its relevance for Poser (1992)'s suggestion, which does involve competition in the sense defined above, and to which we now turn.

Poser (1992) suggests that competition can take place at an intermediate level (his *small categories*, such as *A* and *V*), locally enough to avoid the problems for Gricean blocking accounts, but not so local as to exclude structures that are bigger than lexical items. Aiming at deriving complementarity patterns such as that of the English comparative, Poser proposes that if the lexicon contains a word that expresses the same features as a phrasal small category, the word must be used. For example, since the lexical form *smarter* is available, the phrasal *more smart* is blocked. Phrasal forms are possible only when there is no lexical alternative to block them. For example, the absence of a lexical form such as **intelligenter* makes possible the phrasal form *more intelligent*.

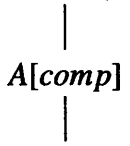
We can define a pre-order $\prec_{\pi}$ that corresponds to Poser's proposal.[13]

(28) Lexical preference: if a lexical item λ and a phrase ϕ belong to the same small category and express the same features, then $\lambda \prec_{\pi} \phi$

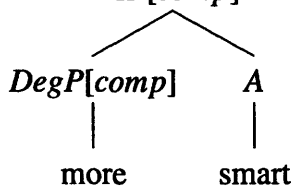
For the English comparative example of *smart-er* vs. **more smart*, the structures would be as follows:

(29) Structures for the English comparative

a. λ =*AP*[*comp*]



b. ϕ =



Poser's account restricts the alternatives to a set that is constrained enough to avoid some of the problems of the global approaches that we saw earlier. For example, on the assumption that *smarter* and **more smart* are as schematized in (29) and that the syntactic features of the two forms are identical, *John is smarter than Bill* will be a competitor for **John is more smart than Bill* but not for a paraphrase like *John's intelligence is higher than Bill's*, which may mean the same as *John is smarter than Bill* but has no *AP* with the same features as *smarter*.

By relying on syntactic features within particular categories rather than on semantic or pragmatic equivalence, Poser's comparison can derive the blocking of medium-sized constituents by words without being committed to the ungrammaticality of all other structures that means the same. This is clearly an advantage over the global approach. However, as pointed out by Embick and Marantz (2008), the distinction between words and phrases, on which Poser's account is based, makes wrong predictions with respect to the possibility of

[13]Only the irreflexive version is defined here, and only in its local form. The reflexive version and the extension to full structures are defined as usual.

preventing competition from arising. Abstractly, using the schematic representations below, the prediction is that a word like w_1 in (30) can block the combination of w_2 and w_2 when they form a constituent, as in (31b), but not in a configuration like (32), where any constituent that contains w_2 and w_3 also contains additional material.

(30) Lexicon:

a. $w_1[\alpha,\beta]$

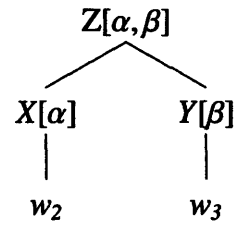
b. $w_2[\alpha]$

c. $w_3[\beta]$

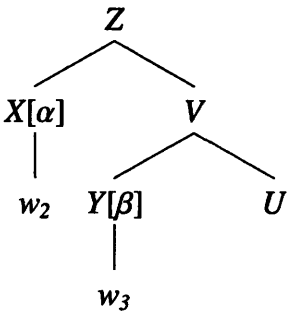
(31) a. $Z[\alpha,\beta]$

w_1

b.



(32)



For comparatives, this prediction can be tested by adding a *PP* complement, corresponding to *U* in (32). A synthetic comparative like *quicker* or *easier* should be able to block the analytic forms **more quick* and **more easy*, but only if the analytic form is a minimal constituent. As soon as the adjective takes a complement, as in *proud of his accomplishments* or *easy to understand*, the analytic form will have no constituent that competes with the synthetic form, and blocking is predicted not to arise. In fact, however, as Embick and Marantz observe, *PP* complements do not seem to interfere with blocking at all:

(33) John is ...

a. proud-er/*more proud (of his accomplishments)

b. easi-er/*more easy (to understand)

...than Bill

The problem that Embick and Marantz observe is in some sense the opposite of the problem that global approaches faced. For global competition, the problem was caused by the availability of too many alternatives, giving rise to predictions about blocking effects that are never attested, as with *John's intelligence is greater than Bill's*, and to imperfect winners that are never found, as in **did forgo*. For Poser's account, the problem is caused by the availability of too few alternatives, giving rise to prediction that adding material to the phrase should obviate blocking, contrary to fact.

Embick and Marantz's conclusion is that competition is restricted, but that a distinction between words and phrases is irrelevant. Rather, they offer a highly localized notion of competition in which the only candidates involved are the possible phonological forms that could be inserted as the phonological exponents of a subset of the terminal elements in the structure. Specifically, in the terms of Distributed Morphology (DM; Halle and Marantz, 1993), competition takes place only at the point of Vocabulary Insertion, where phonological exponents compete according to specificity, following the *Subset Principle* (34), the DM reformulation of Kiparsky's Elsewhere Condition.

(34) Subset Principle (Halle, 1997, p. 128):

> "The phonological exponent of a Vocabulary item is inserted into a morpheme in the terminal string if the item matches all or a subset of the grammatical features specified in the terminal morpheme. 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."

1.3 Outline

The main goal of this dissertation is to point out ways in which structure affects competition. In particular, I will try to show that the status of ϕ will only depend on structures that are at most as complex as ϕ, under a certain notion of complexity. Chapter 2 argues for this restriction in the domain of scalar implicature, and chapter 3, joint work with Danny Fox, extends the argument to focus semantics. More tentatively, I will suggest that the relation `at-most-as-complex-as` is one of the pre-orders that enter the comparison of candidates, giving rise to a direct preference for simpler structures. Chapter 4 presents evidence for a preference for simplicity in the domain of morpho-semantic mismatches. The two notions of simplicity that we will arrive at, the one for implicature and focus and the one for morphology, are similar but not identical. For implicature and focus, the notion of simplicity that we will use will be one of structural embedding (closely related to the graph-theoretic notion of *graph minor*): we will say that ϕ' is at most as complex as ϕ if ϕ can be mapped onto ϕ' by deleting material in ϕ and allowing certain substitutions. Crucially, important structural relations will remain intact by this mapping. For morphology, on the other hand, the relevant notion of simplicity will be one of *counting*: ϕ' will be considered at most as complex as ϕ with respect to a certain element α if ϕ' has at most as many occurrences of α as ϕ does. Differently from embedding, counting does not preserve various hierarchical relations. I do not know whether the two kinds of competition are instances of a more general mechanism; my advice to the reader is to treat the two sections, chapters 2 and 3 on the one hand and chapter 4 on the other, as unrelated.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 2

Implicatures

The general claim of this dissertation is that competition exists throughout the grammar, and that structure plays a central role in determining which candidates can compete. In the coming chapters I will try to show that the way in which structure affects the competition is through a preference for simpler structures over complex ones. While the form of comparison is roughly the same across different parts of the grammar, the fate of the loser will vary depending on several factors, among which meaning is perhaps the most significant. When two candidates have the same meaning and one of them is strictly simpler than the other, the more complex one will tend to be ungrammatical. This will be the case, for example, in so-called morphological blocking. When the two candidates have different meanings, on the other hand, a candidate ϕ that would otherwise lose to a better candidate ϕ' might be rescued by the beliefs and attitudes of the speaker: if the speaker can truthfully assert ϕ but lacks the evidence to support the assertion of ϕ', the speaker may have to stick to ϕ. This is what happens in the domain of conversational implicature,[1] and more specifically in what has come to be known as *scalar implicature*[2] Scalar implicatures (or simply *implicatures* in what follows) are at the center of the current chapter, and I will use them to establish the relevance of structural complexity to constraining the candidate set for competition.

[1]The term is due to Grice (1989), who also offered the first systematic discussion of the phenomenon. The phenomenon itself has a much longer history. See Horn (1989) for references and discussion.

[2]Horn (1972) introduced the notion of scales (already present in Horn, 1969). See Gazdar (1979), Atlas and Levinson (1981), and Hirschberg (1985/1991), among others, for further developments. We will discuss scales in more detail below.

The proper treatment of implicatures has been an issue of some debate in recent literature, and the question of whether the relevant computations are global (at the level of whole sentences) or local (at the level of each embedded proposition) has turned out to be particularly divisive. In presenting the argument for structural complexity in this chapter I will attempt to abstract away from as many of the controversial issues as possible. For the most part, the same treatment of alternatives is shared by all the major approaches. I will also resist the temptation to argue here for a direct preference for simpler structures (though I will provide a brief sketch of how this could be done), and I will avoid relating implicatures to some obviously related phenomena (focus, questions, and others). By the end of this chapter, the reader is only expected to accept that structural complexity determines the definition of the alternatives. Getting there will take some preparation, especially since the literature has largely agreed that structure is irrelevant to the choice of alternatives, and I will have to explain why I think otherwise, but in terms of the main goals of this dissertation, the current chapter makes only the rather general point that complexity matters. The more interesting questions of locality, simplicity, and the relation to alternatives elsewhere will be the topic of the chapters that follow.

2.1 Outline

Computing the implicatures of a linguistic object involves reference to alternative objects that were not used. This chapter argues for a structure-sensitive characterization of these alternative objects. For any structure ϕ, the alternatives will be all those structures that are at most as complex as ϕ, under a particular notion of complexity:

(35) $A_{str}(\phi) = \{\phi' | \phi' \lesssim \phi\}$

Complexity is not a new notion in the domain of conversational reasoning. It is present in Grice's Maxim of Manner and elsewhere. However, there has also been skepticism about the relevance of complexity for implicatures. One reason for skepticism is the scarcity of instances where an inference about ϕ seems to require reasoning about a strictly simpler ϕ'. In most cases, ϕ and ϕ' are of roughly the same complexity. This is surprising under (35). On the other hand, under the widely assumed characterization of alternatives in terms of

scales (Horn, 1972; Gazdar, 1979), the scarcity of complexity-related implicatures makes much more sense. And there have been other objections. Perhaps the most serious objection to the use of complexity is an argument by Matsumoto (1995), who observes that in certain cases an inference about an utterance ϕ requires reference to an alternative that appears to be strictly more complex than ϕ. Matsumoto's conclusion is that complexity does not play a role in the computation of implicatures. Instead, he follows Horn (1989) in adopting a semantic constraint on scale-mates that makes use of the notion of *monotonicity*. We will review these arguments in section 2.2.

Section 2.3 presents an empirical observation that will suggest that complexity might be the right way to go after all. The observation is this: in certain contexts, a complex, non-monotonic expression can give rise to an implicature that requires a comparison with a simpler alternative. This goes against the monotonicity condition, and indirectly also against the notion of scale, while supporting the complexity approach.

To turn the complexity approach into an actual proposal, we will need an explicit definition of structural complexity. We will also have to address Matsumoto's objection as well as the concern about the scarcity of comparisons with strictly simpler alternatives. Section 2.4 offers a definition of complexity, based on the idea that we can transform ϕ into a structure that is no more complex if we restrict ourselves to (a) *deleting* elements in ϕ, and (b) *substituting* elements in ϕ with other elements from an appropriately defined source. In section 2.5 we will see that a large family of conversational inferences is correctly predicted, including the disjunction puzzles analyzed by Sauerland (2004), as well as some new cases. Along the way we will see why inferences that refer to simpler alternatives are rare: usually, simpler structures result in weaker assertions, so no inference arises. Once this is observed, the problem can be avoided by embedding complex structures in downward-entailing contexts, where entailment relations are reversed. As predicted, we will find that in such contexts the relevant inferences arise. In section 2.6 we turn to Matsumoto's argument against complexity. Addressing the argument will force us to be more careful about our notion of substitution source. This, in turn, will lead to certain new predictions.

The primary goal of this chapter is to argue for a structure-sensitive characterization of the alternatives. The secondary goals are to argue against a monotonicity constraint on alternatives, as well as against the intermediate step of scales. Before we start, I would like to mention what this chapter is *not* about. I will have very little to say about the question of whether implicatures are computed locally (as in Landman, 2000, Chierchia, 2004, Fox, 2006, and others) or globally (as in the traditional accounts, as well as in recent proposals by Sauerland, 2004 and Russell, 2006). Much of the discussion will be framed in terms of global, neo-Gricean reasoning, but this is done for ease of presentation only, and in the following chapters we will switch to a local perspective. As far as I can tell very little of the current discussion depends on this choice. Secondly, I will avoid talking about the relation of implicatures to exhaustivity in answers, as in the proposals of van Rooij and Schulz (2004), Sevi (2005), Spector (2007), and others. Finally, I will have nothing to say about implicatures that depend on context and world knowledge, of the kind discussed by Hirschberg (1985/1991).

2.2 Background

2.2.1 Conversational reasoning

Informally speaking, a cooperative speaker can be assumed to make maximally helpful contributions to the conversation. If ϕ and ϕ' are potential contributions, and if ϕ' is strictly better than ϕ (i.e., ϕ' is at least as good as ϕ in every relevant respect, and strictly better than it in at least one relevant respect), then a cooperative speaker will prefer using ϕ' when possible. Consider, for example, a situation in which (36) and (37) are both potential contributions to the conversation.

(36) John ate some of the cake

(37) John ate all of the cake

(37) is more informative than (36), but in all other respects the two sentences seem equally good. A cooperative speaker, then, will prefer using (37) to (36). If despite this preference the speaker has uttered (36), the hearer can conclude that there were considerations other

than those mentioned above that prevented the speaker from using (37). Depending on what else can be assumed, the hearer may take this inference to license further conclusions. For example, they can conclude that the speaker does not have an opinion about (37), or that the speaker believes (37) to be false. Still in loose terms, the preference can be stated like this:

(38) Conversational principle (informal version): do not use a linguistic object ϕ if there is another object ϕ' such that both

a. ϕ' is better than ϕ, and

b. ϕ' could have been used

Several aspects of (38) must be specified if we want to be able to apply it at all. First, one has to decide what kinds of objects ϕ and its alternatives are. Then one has to provide concrete content to the term *use* (in fact, we will soon see that it is common to provide different content to each of the two occurrences of *use*), as well as to the term *better*. Once this is done, (38) can license the following kind of inference: if a cooperative speaker has used ϕ, then for all better alternatives ϕ' it is not the case that the speaker could have used ϕ' instead of ϕ. In our example, since (37) is better than (36), the use of (36) licenses the inference that it is not the case that the speaker could have used (37).

Depending on what else is assumed, these inferences can sometimes license further inferences about why it is that the better ϕ' could not be used. For example, if the speaker can be assumed to be in a position to use either ϕ' or $\neg\phi'$, the weak inference *it is not the case that the speaker could have used* ϕ' can be strengthened into *it is the case that the speaker could have used* $\neg\phi'$. If such an assumption can be made about the speaker of (36), and if using an alternative implies believing that it is true, the hearer will conclude that the speaker believes that (37) is false, and that John ate some but not all of the cake.

The general 2-step architecture we just went through follows the proposals of Soames (1982), Hirschberg (1985/1991), Horn (1989), and more recently Sauerland (2004), and I will make extensive use of it in discussing possible choices for the definition of alternatives. The first step derives weak, or, using Sauerland's terminology, *primary* implicatures using a conversational principle. The second step strengthens these to *secondary* implicatures

under certain conditions (see Fox, 2006 for important modifications of this part). In terms of this architecture, the present chapter focuses entirely on the first step.

2.2.2 The Symmetry Problem

As an attempt to turn (38) into a working definition, we can first try the following idea, based on an over-simplification of Grice, and more directly on the formulation in Gamut (1991). The linguistic objects, ϕ and its alternatives, are full sentences. Using them, in the sense of the first occurrence of *use*, means asserting them in discourse. *Better* means more informative, understood in terms of semantic entailment. And using an alternative ϕ', in the sense of the second occurrence of *use*, means that it is believed to be true, relevant, and supported by evidence. This last condition means, roughly, that ϕ' would be assertable in the absence of any other alternative. It will be convenient to have a shorter way to say this. I will do this by defining a notion of *weak assertability*:[3]

(39) WEAK ASSERTABILITY:

A structure ϕ will be said to be *weakly assertable* by a speaker S if S believes that ϕ is true, relevant, and supported by the evidence.

Our version of the conversational principle will now look like this:

(40) Conversational principle (naïve version): do not assert ϕ if there is another sentence ϕ' such that both

a. $[\![\phi']\!] \subset [\![\phi]\!]$, and

b. ϕ' is weakly assertable

We now have what looks like a concrete version of the conversational principle in (38), which we could use to reason about examples like (36) above. As noted by Kroch (1972), however, (40) does not work. The problem, dubbed the *symmetry problem* in class notes by Kai von Fintel and Irene Heim (see also Horn, 2000 and von Fintel and Fox, 2002), has to do with the fact that (40) allows us to reason about too many alternatives: for any ϕ'

[3]The definition of weak assertability makes reference to a speaker S. Most of the time, however, no confusion is likely to arise, and I will usually suppress this parameter.

that is stronger than ϕ, and that we would like to reason about, there is another alternative, $\phi'' = \phi \wedge \neg\phi'$, which is also stronger than ϕ, and which would license an inference in the opposite direction. Combined, ϕ' and ϕ'' license only ignorance inferences, contrary to fact.

To see how the symmetry problem arises in a simple example, let us return to (36) above, repeated here:

(41) John ate some of the cake

Assuming that the speaker obeys (40), the hearer can conclude that there was no sentence that is strictly more informative than (41) that the speaker could have (weakly) asserted instead. In particular, the speaker could not have asserted this (repeated from (37)):

(42) John ate all of the cake

(42) is more informative than (41). Assuming that (42) is relevant, the hearer may conclude that either (42) is not supported by the evidence that the speaker had or that the speaker believes it to be false. If (42) were the only sentence to consider in the context of (41), and if it is plausible that the speaker has access to the relevant evidence, the hearer could conclude that the speaker believes (42) to be false. However, (40) does not restrict us to (42) alone. Consider (43), for example:

(43) John ate some but not all of the cake

Like (42), (43) is also strictly more informative than (41), and the same reasoning as above would lead the hearer to conclude that either (43) is not supported by the evidence that the speaker has, or that the speaker believes it to be false. Again, if (43) were the only sentence to consider, and if it is plausible that the speaker is well-informed, the hearer could conclude that the speaker believes that (43) is false, and that John ate all of the cake. And here is the problem: given that (41) is true, the two alternatives (42) and (43) cannot be simultaneously false. Either John ate all of the cake or he didn't.[4] The speaker cannot believe that (42) and (43) are both false (unless the speaker is entertaining contradictory beliefs), and the hearer can reason that the speaker is not opinionated with respect to at

[4]Assuming that all presuppositions are satisfied.

least one of the two alternatives. But for a speaker who believes that (41) is true, being opinionated about one of the two alternatives means being opinionated about the other, and the hearer can conclude that the speaker is not opinionated about either. In other words, the speaker did not have access to the relevant evidence after all, and so was not in a position to believe either (42) or (43). The speaker is simply ignorant of the relevant facts.

Ignorance inferences of this kind often conflict with other beliefs of the hearer. For example, the hearer might be quite confident that the speaker was watching John throughout the cake-eating event. And yet there is nothing deviant in (41) even in this context. The assertion implies neither that the speaker has lost her recollection of whether or not John ate all of the cake nor that the speaker is non-cooperative. In fact, the standard inference in this situation is that the speaker believes that John did not eat all of the cake. That is, actual hearers reason as if only (42) were an alternative to (41) for purposes of conversational reasoning. Fixing (40), then, will require excluding (43) from the reasoning process.

2.2.3 Restricting the alternatives

The neo-Gricean approach

(40) said that when ϕ is uttered, every alternative ϕ' should be considered. The symmetry problem arose from the fact that for each such ϕ', a symmetric ϕ'' was available. A widely accepted solution involves being more careful about the alternatives that are being referred to. Each sentence ϕ can be thought of as being associated with a set of alternatives, $A(\phi)$, and it is only those alternatives that are considered for purposes of deriving implicatures. If $A(\phi)$ is chosen appropriately, it will consist of those alternatives that correspond to the actual inferences that are made. Symmetric alternatives of the kind that gives rise to unattested ignorance inferences will be excluded. Here is a minimally revised version of (40) that can refer to alternatives.

(44) Conversational principle (alternative-sensitive): do not assert ϕ if there is another sentence $\phi' \in A(\phi)$ such that both

a. $[\![\phi']\!] \subset [\![\phi]\!]$, and

b. ϕ' is weakly assertable

For (41) above, for example, the goal would now be to find a set of alternatives that will include (42) but not (43). Of course, one would like to have a principled way to arrive at $A(\phi)$ for any given ϕ. If we had to associate each sentence with an arbitrary set of alternatives, the theory would hardly be predictive.

Scales

Scales, introduced by Horn (1972) and developed further by Gazdar (1979), Atlas and Levinson (1981), and others, offer a partial solution to the problem of deriving $A(\phi)$. The intuition is this: to determine that (42) is a good alternative for (41) while (43) is not, all we have to know is that *all* is an alternative expression to *some* while *some but not all* is not. Determining whether sentences containing these expressions are alternatives to each other is done mechanically, by substitution. The idea, then, is to break down the task of deriving A, the relation between full sentences and their alternatives, into two components. The first component is a new relation, S (for `scale-mate-of`), defined over a proper subset of linguistic objects, sometimes referred to as *scalar items* (the set of scalar items is typically finite, and is often taken to be a subset of the lexicon). The second component uses S to derive A by substitution: if $< \alpha, \alpha' > \in S$, and if ϕ is a sentence that has an occurrence of α, then $< \phi, \phi[\alpha'/\alpha] > \in A$, where $\phi[\alpha'/\alpha]$ is obtained from ϕ by replacing that occurrence of α with α'.[5] In our example, it would suffice to specify that <*some, all*>$\in S$ while <*some, some but not all*>$\notin S$.[6]

The scalar approach offers a way to avoid stating separately for every sentence what its alternatives are. However, if we don't know how scales are derived, their usefulness will be limited. For any two expressions we would need to know in advance whether they are scale-mates or not; otherwise, no predictions can be made. Perhaps we can do no

[5]As Sauerland demonstrates, the replacement of elements by their scale-mates should not be restricted to one-step substitutions. Rather, any sequence of replacements of elements with their scale-mates that can take us from ϕ to ϕ' will license using ϕ' as an alternative.

[6]A further point that Sauerland makes is that `scale-mate-of` must be a symmetric relation: what matters is the entailment relation between the whole sentences ϕ and $\phi[\alpha'/\alpha]$ and not between the scalar items α and α' within them. This is important for dealing with the phenomenon of scale reversal in downward-entailing contexts. The order in which the items are listed on the scales here, then, should not be taken as significant.

better than that. Gazdar (1979, p. 58), for example, concluded that scales are "... in some sense, 'given to us' ". I will try to show that we can do better than that, and that structural complexity provides an adequate characterization of alternatives. This direction, though, will eventually lead us to abandon the notion of scales.

Complexity and its problems

The role of complexity in communication was explored by Zipf (1949), Grice (1989), McCawley (1978), Atlas and Levinson (1981), and Horn (1984), among others (for more recent work, see Blutner, 2000, Parikh, 2000, and van Rooij, 2004). For our example (41) we can state the reasoning as follows. (41) and (42) are of roughly the same complexity, in some sense, and so they can be compared for purposes of computing implicatures. (43), on the other hand, is more complex than (41), and is consequently ignored. I stated the informal reasoning process in terms of comparing whole utterances. It is easy to do the same in terms of scales: *some* and *all* are of the same complexity, and can therefore be on the same scale, while *some but not all* is more complex than *some* and cannot be on the same scale with it.[7]

Within the Gricean framework, complexity has often been related to the Maxim of Manner. The status of this maxim, however, has remained unclear. For example, one wonders whether a cooperative speaker should avoid providing relevant information, as in the excluded *some but not all*, just in order to maintain simplicity. It is also somewhat surprising that the supposed effects of Manner are manifest almost exclusively in licensing alternatives of the same complexity, as with (41) and (42). Shouldn't we also expect to find a complex sentence having implicatures that are based on negating simpler alternatives?

[7]In the literature, one often finds attempts to characterize alternatives in terms of *degree of lexicalization*, as distinct from brevity. Thus, (Atlas and Levinson, 1981, p. 44) write that "... to constitute a genuine scale for the production of scalar implicatures, each item must be lexicalized to the same degreee." As discussed by Matsumoto (1995), it is not always clear how one should distinguish between brevity and lexicalization (Matsumoto himself argues that neither of these is relevant for implicature). See Horn (2000) for further discussion on this matter, as well as for reasons to reject Atlas and Levinson (1981)'s symmetric condition in favor of a formulation that allows using alternatives that are strictly more lexicalized than what was uttered, much in the spirit of the current proposal.

As mentioned in the introduction, however, candidates for such implicatures appear to be few and far between.

More serious concerns about the Maxim of Manner have to do with the precise notion of complexity that is used. Certain definitions, such as syllable count and phonetic effort, seem to yield incorrect results (see Poser, 1992 for discussion). Worse, as argued by Matsumoto (1995), some implicatures require alternatives that are strictly more complex than the actual utterance under any reasonable definition of complexity. If this is correct, complexity cannot be the way to go.

The argument is based on the results of asserting (45a):

(45) a. It was warm yesterday, and it is a little bit more than warm today (Matsumoto, 1995, ex. 39 p. 44)

b. It was a little bit more than warm yesterday, and it is a little bit more than warm today

As Matsumoto notes, uttering (45a) gives rise to the inference that it was only warm yesterday, and not a little bit more than warm. Intuitively, this inference seems to rely on considering the stronger alternative (45b) and concluding that it was false. The problem, of course, is that (45b) is more complex than (45a).

Monotonicity

Complexity, then, appears to be on the wrong track. A more promising approach, Matsumoto suggests, is the one that was outlined in Horn (1989), and in which the semantic notion of *monotonicity* (or *scalarity*) plays a role.

(46) "Positive and negative quantifiers, modals, and related operators must be represented on distinct, though related scales. There can be no single scale on which operators like *some* and *not all*, or *possible* and *unlikely*, can be plotted. Rather, there is one scale defined by the positive operators and one by their negative counterparts." (Horn, 1989, p. 235)

(47) "Scalarity Condition: Expressions that form a Horn scale must be either 1) all positively scalar (e.g., *<all, some>*) or 2) all negatively scalar (e.g., *<no, few>*)."

(Matsumoto, 1995, p. 46)

Condition (47) licenses the alternatives that we have seen so far.[8] It also prevents expressions such as *some but not all* from being on the same scale as *some*, or indeed on any scale.

In fact, one could try to further strengthen the argument for monotonicity and against complexity, as was pointed out to me by Danny Fox (p.c.). Imagine a complexity-based theory that would somehow allow (45b) to be used as an alternative for (45a). Such a theory, it seems, will also allow (48b) to be used as an alternative to (48a):

(48) a. John talked to some of the girls yesterday, and he talked to some but not all of the girls today

b. John talked to some but not all of the girls yesterday, and he talked to some but not all of the girls today

However, uttering (48a) does not give rise to the inference that (48b) is false. In fact, (48a) does not seem to give rise to any inference at all. It is just an odd sentence. For a monotonicity-based account, this is not a problem. The non-scalar *some but not all* is ruled out as an alternative on semantic grounds. It seems, then, that a complexity-based approach would have to include both a special mechanism that would license *a little bit more than warm* as an alternative for *warm* in sentences such as (45a), and a monotonicity condition to rule out *some but not all* as an alternative to *some* in the otherwise analogous (48a). This looks like a fairly direct argument against the complexity approach.

2.3 A puzzle

Having convinced ourselves that monotonicity works and that complexity does not, let us now look at the following sentences. In each example, uttering the (a) sentence seems to imply that the (b) sentence is unassertable.

(49) a. I doubt that exactly three semanticists will sit in the audience

[8]Though see Sevi, 2005 for complications with applying the notion correctly.

b. I doubt that three semanticists will sit in the audience

(50) a. If we meet John but not Mary it will be strange

b. If we meet John it will be strange

(51) (Danny Fox, p.c.)

a. Everyone who loves John but not Mary is an idiot

b. Everyone who loves John is an idiot

By uttering (49a), for example, the speaker asserts that they find it unlikely that the number of semanticists in the audience will be exactly three, but they also suggest that they do not find it unlikely that the number would be at least three. Similarly, (50a) asserts that all situations in which we meet John but not Mary will be strange, while implying that, for all the speaker knows, there can be situations in which we meet John and in which there is no strangeness (in those situations we would necessarily also meet Mary, though it is conceivable that there would also be some strange situations in which we meet both John and Mary). And (51a) suggests that, as far as the speaker is concerned, it is possible to love John without being an idiot (again, as long as one also loves Mary). In each example, the (b) sentence is stronger than the (a) one, so if we could use the (b) sentences as alternatives, we would be able to derive the implication that the (b) sentence is not assertable as an implicature of the (a) sentence.[9] The problem is that both *exactly three semanticists* and *John but not Mary* are non-scalar, and so the monotonicity approach predicts that they would not be used. On the other hand, the complexity approach makes the correct predictions here.

I will try to argue that the situation in (49)-(51) is not an accident, and that complexity, under the right formulation, provides a correct characterization of the alternatives, while monotonicity is wrong. To be able to make this claim I will have to provide a concrete definition of complexity-based alternatives, and I will also have to explain why it is that in some cases complexity appears to be wrong and monotonicity appears to be right.

[9]That these are indeed implicatures can be seen by their cancelability. For example, the suggestion that it is possible to love John without being an idiot disappears if we continue (51a) with "In fact, everyone who loves John [period] is an idiot." I thank an anonymous L&P reviewer for pointing this out to me.

2.4 Proposal

2.4.1 Structural complexity

Certain operations simplify a structure. Viewing syntactic structures as trees (or directed graphs, more generally, with edges representing motherhood), simplifying operations may involve *deletion* (removing edges and nodes) and *contraction* (removing an edge and identifying its end nodes). Other operations do not simplify but also do not add complexity. *Substitution* of one terminal element for another terminal element of the same category is an example, and more generally, substitutions of structures for other structures given an appropriately defined substitution source. For the moment we can think of the substitution source as the lexicon of the language, though we will revise this once we get back to Matsumoto's sentences in section 2.6.[10]

(52) SUBSTITUTION SOURCE (first version, to be revised in 75):
Let ϕ be a parse tree. The *substitution source* for ϕ, written as $L(\phi)$, is the lexicon of the language.

We can now say that a structure ψ is no more complex than ϕ if ψ can be obtained from ϕ by a finite number of operations of the kind discussed above. Here are the definitions.

(53) STRUCTURAL COMPLEXITY:
Let ϕ, ψ be parse trees. If we can transform ϕ into ψ by a finite series of deletions, contractions, and substitutions of constituents in ϕ with constituents of the same category taken from $L(\phi)$, we will write $\psi \lesssim \phi$. If $\psi \lesssim \phi$ and $\phi \lesssim \psi$ we will write $\phi \sim \psi$. If $\psi \lesssim \phi$ but not $\phi \lesssim \psi$ we will write $\psi < \phi$.

[10]The existence of a lexicon is a contentious issue (see Pesetsky, 1985 and Marantz, 1997 for discussion), and in non-lexicalist approaches such as Distributed Morphology (DM, Halle and Marantz, 1993) there is no place where words are listed to the exclusion of non-words. The use of the lexicon under the current proposal is done only for obtaining a set of easily accessible constituents for substitution. These do not have to be words, and the revision that will be needed to deal with Matsumoto's examples will bring in things that are strictly non-words. It might therefore be better to use something like the DM notion of *vocabulary* in the definition of $L(\phi)$, though for presentation purposes I will stick to the term *lexicon*.

Using the structural relation $\lesssim$ we can now provide a complexity-based definition of the set of alternatives.

(54) STRUCTURAL ALTERNATIVES:

Let ϕ be a parse tree. The set of *structural alternatives* for ϕ, written as $A_{str}(\phi)$ is defined as $A_{str}(\phi) := \{\phi' : \phi' \lesssim \phi\}$.

2.4.2 Using the alternatives

Having a set of alternatives is one thing. Using them is another. In the informal discussion above I have assumed a global comparison of alternatives. This is the traditional (neo-)Gricean approach (see Horn, 1972; Gazdar, 1979; Atlas and Levinson, 1981; as well as Sauerland, 2004; Russell, 2006; Spector, 2007; Geurts, 2007 for a recent defense of the global approach). Over the past few years it has been recognized that the global approach faces non-trivial difficulties, and a localist approach to implicatures has been developed by Landman (2000), Chierchia (2004), and Fox (2006). It is not the goal of this chapter to decide between these two approaches. Both make crucial use of alternatives, and both are compatible with the definition of structural alternatives in (54). I will keep discussing implicatures using the globalist perspective of Sauerland's 2-step architecture, outlined above, but this is done for ease of exposition only.[11]

Obtaining the primary implicatures is done through the neo-Gricean principle (44) above, where $A(\phi)$ is now specified to be $A_{str}(\phi)$, the structural alternatives of ϕ defined

[11]Here is a brief sketch of how the current discussion could be recast in a localist system, along the lines of Fox (2006). Implicatures, in Fox's system, are derived using an exhaustive operator, *Exh*, roughly a non-presuppositional counterpart of *only*. When *Exh* attaches to a clause ϕ, the result is an assertion that $[\![\phi]\!]$, the proposition denoted by ϕ, is true, and that all the alternatives to p in a set of alternatives A that can be safely negated (Fox's *innocently excludable* alternatives, here written as $IE(A, p)$) are false:

$$[\![[\text{Exh } \phi]]\!] = \lambda w_s.[\![\phi]\!](w) \& \forall q \in IE(A, p).\neg q(w)$$

On an account that derives implicatures using such an operator, the current proposal would translate into restricting the set A to be a subset of the set of denotations of the structural alternatives to ϕ. That is, $A = \{[\![\phi']\!] \mid \phi' \in A_{str}(\phi)\}$.

in (54).

(55) Conversational principle (with structural alternatives): do not use ϕ if there is another sentence $\phi' \in A_{str}(\phi)$ such that both

a. $[\![\phi']\!] \subset [\![\phi]\!]$, and

b. ϕ' is weakly assertable

2.4.3 Aside: Structural alternatives and the Maxim of Manner

Notice that (55) expresses an indirect preference for simple sentences: if ϕ' is strictly simpler than ϕ then ϕ' will be taken into account in computing the implicatures for ϕ, but ϕ will not be relevant for the implicatures of ϕ'. On the other hand, there is no direct pressure to minimize structure: the fact that ϕ' is strictly simpler than ϕ does not on its own mean that ϕ' is preferable to ϕ if the two are truth-conditionally equivalent. In other words, we are using complexity as a neo-Gricean filter on structures and not as a preference that might correspond to the Gricean Maxim of Manner. Let us look at what a Gricean formulation would look like. Here again is the conversational principle we started with:

(56) Conversational principle (informal version, repeated from (38)): do not use a linguistic object ϕ if there is another object ϕ' such that both

a. ϕ' is better than ϕ, and

b. ϕ' could have been used

The choice of interpreting *better* in terms of semantic entailment in (56) was a simplification, based on one part of Grice's Maxim of Quantity. To incorporate Manner, we could combine semantic entailment $\subseteq$ with our notion of structural complexity $\lesssim$. Using $\precsim$ for `at-least-as-good-as`, we could write this as follows:

(57) $\precsim := \{(\phi, \psi) \mid \phi \lesssim \psi \wedge [\![\phi]\!] \subseteq [\![\psi]\!]\}$

We could then take *better* to mean the irreflexive $\prec$ (defined as $\precsim \setminus \precsim^{-1}$)), changing our neo-Gricean (55) into a Gricean principle:[12]

[12]More generally, we could start with any set of pre-orders over the domain of structures Φ (that is, reflexive, transitive relations over $\Phi \times \Phi$), $\{\leq_\alpha\}_{\alpha \in I}$. In addition to semantic entailment and structural complexity,

(58) Conversational principle (Gricean): do not assert ϕ if there is an alternative ϕ' such that

a. $\phi' \prec \phi$

b. ϕ' is weakly assertable

Treating complexity as part of the comparison, as in (58), is perhaps more natural from a globalist, pragmatic perspective than using it as part of the filtering mechanism, as in (55).[13] More significantly, the two approaches make distinct predictions. If $[\![\phi]\!]=[\![\phi']\!]$ then under (55) the weak assertability of ϕ' will never affect the assertability of ϕ. Under (58), if ϕ' is strictly less complex than ϕ, the weak assertability of ϕ' will prevent ϕ from being assertable. For the examples that we will discuss in the remainder of this chapter, the differences between the different formulations will not matter. I will keep using the neo-Gricean (55), mostly in order to facilitate the comparison with other proposals.

2.5 Examples

2.5.1 Some, all, some but not all

Let us start by verifying that our definitions can handle the simple (41) with which we started, repeated here.

(59) a. ϕ = John ate some of the cake

b. ϕ' = John ate all of the cake

c. ϕ'' = John ate some but not all of the cake

Assuming that *all* and *some* are of the same syntactic category, and assuming that both are in the lexicon, we can substitute one for the other and get from ϕ to ϕ' (or from ϕ' to ϕ) without adding complexity. That is, $\phi' \sim \phi$, and by definition, $\phi' \in A_{str}(\phi)$. Consequently,

one can use pre-orders that correspond to obscurity, appropriateness, politeness, and so on. The relation of `at-least-as-good-as` can then be defined as the intersection of these pre-orders: $\preceq := \bigcap_{\alpha} \preceq_{\alpha}$. The rest proceeds as above.

[13] See Fox (2006) and Russell (2006) for considerations regarding the role of formal alternatives in conversational reasoning.

when ϕ is uttered, ϕ' is evaluated by (55), which tells us that ϕ should not be uttered if (i) $[\![\phi']\!] \subset [\![\phi]\!]$, and (ii) ϕ' is weakly assertable. Assuming that the speaker obeys (55), then, the utterance of ϕ means that at least one of (i) and (ii) does not hold. (i) is true, so (ii) must be false. That is, ϕ' is believed to be not true, not relevant, or not supported by the evidence.

ϕ'' is also stronger than ϕ, and we saw that if it is not excluded from being an alternative it gives rise to the symmetry problem. Under the current definitions, ϕ'' is successfully excluded. The alternatives of ϕ are only structures that are at most as complex as ϕ, and ϕ'' is not such a structure: there is no way in which we can transform ϕ into ϕ'' by the operations relevant for structural complexity listed in (53). In fact, ϕ'' is strictly more complex than ϕ. (55) does not consider ϕ'', and so no inference is made.

Summing up, uttering ϕ gives rise to the inference that it is not the case that ϕ' is weakly assertable, and it does not give rise to any inference with respect to ϕ''. Assuming that no other alternatives are available, the only primary implicature of ϕ is that ϕ' is not weakly assertable. This now feeds the second process in which the primary implicatures are strengthened. Under the appropriate assumptions, the secondary implicature is that ϕ' is false. In total, we obtain the desired inference that John ate some but not all of the cake.

2.5.2 'or', 'and', L, R

Disjunction gives rise to two main inferences: that only one of the disjuncts is true (*exclusiveness*), and that the speaker does not know which one it is (*ignorance*). We start with exclusiveness. Uttering ϕ, for example, usually implies that it is not the case that John ate both the apple and the pear. That is, ϕ implies that ϕ' is false.

(60) a. ϕ = John ate the apple or the pear

b. ϕ' = John ate the apple and the pear

c. ϕ'' = John ate the apple or the pear but not both

To obtain the exclusive reading of ϕ we must make sure that ϕ' is an alternative, but the symmetric ϕ'' is not. Under the current approach, this is the result of the same structural consideration that prevented the symmetry problem elsewhere: ϕ' is at most as complex as

ϕ, while ϕ'' is strictly more complex. So far, nothing special.

Turning to the ignorance inference, it was observed by Grice (1989) (see also Gazdar 1979, among others) that disjunction usually implies that the speaker does not know which of the disjuncts holds. In ϕ above, the speaker can be taken not to be in a position to say which of the two fruits John ate. A proposal for deriving the ignorance inference has been developed by Sauerland (2004), and I will adopt the proposal in its essentials. I will show, however, that one particular aspect of the proposal can be simplified: the effects of a seemingly stipulative definition of scales are directly predicted from our complexity-sensitive alternatives.

As Sauerland notes, the ignorance inference for disjunction can be obtained if ϕ has ψ' and ψ'' as alternatives.

(61) a. ψ' = John ate the apple

b. ψ'' = John ate the pear

Semantically, ψ' and ψ'' are both strictly stronger than ϕ. If ψ' and ψ'' are alternatives to ϕ, we may conclude that neither of them was weakly assertable. Since both seem to be relevant in situations in which ϕ is, as was ϕ' above, we obtain the following primary inferences:

(62) a. It is not the case that the speaker believes that ϕ'

b. It is not the case that the speaker believes that ψ'

c. It is not the case that the speaker believes that ψ''

As before, these primary inferences now feed the process of strengthening, resulting in secondary implicatures. For (62a) strengthening is straightforward, yielding the inference that the speaker believes that ϕ' is false. For (62b) and (62c), on the other hand, no strengthening takes place, since it is impossible to strengthen both inferences simultaneously without contradicting the original assertion ϕ. See Sauerland (2004) and Fox (2006) for details and discussion.

The key to obtaining the correct inferences for ϕ is the alternatives. Specifically, we need ϕ', ψ', and ψ'' to be alternatives, while ϕ'' must be excluded. Structurally defined alternatives, using $\lesssim$, derives precisely these alternatives. Notice, however, that there was

something new in this example. So far, the good alternatives were always of the same complexity as the original utterances. The excluded structures were strictly more complex. Here, for the first time, we needed to use alternatives that were strictly less complex than the original structure. Relatedly, this is also the first example in which our proposal behaves differently from a scale-based approach. For a scale-based approach, including ϕ' and excluding ϕ'' is fairly straightforward: *or* and *and* are standardly assumed to be scale-mates, and there is no scale which allows us to get to ϕ''. Deriving ψ' and ψ'', on the other hand, is less simple. The problem, discussed by Sauerland (2004), is this: for any disjunction $p \vee q$ we would like to have both p and q as alternatives. The relation `scale-mate-of`, however, is really an equivalence relation. This would mean that for any p and any q we would have $p \vee q$ as an alternative, which in turn would mean that p and q are scale-mates of each other, regardless of what they are. But if every sentence is an alternative to any sentence, all we would get is ignorance inferences, and no real (secondary) implicatures would ever arise.

Sauerland's solution is to posit two binary connectives, L(eft) and R(ight), which are scale-mates to $\vee$ and $\wedge$. L returns its left argument, and R returns its right argument: $[\![\phi L \psi]\!] = [\![\phi]\!] = [\![\psi R \phi]\!]$. With these new scale-mates, $\phi \vee \psi$ will always have both $\phi L \psi$ and $\phi R \psi$ as alternatives, simulating the effect of using each disjunct as an alternative. Importantly, this does not actually turn the disjuncts into alternatives: ϕ is not the same as $\phi L \psi$ (or $\psi R \phi$); consequently, the problem of arbitrary sentences becoming alternatives of each other disappears.[14]

The introduction of L and R and the conditions on them are somewhat stipulative. Our structural alternatives derive the effect of these connectives using nothing more than what was needed to obviate the symmetry problem with *some*, *all*, and *some but not all*.

2.5.3 Strictly simpler alternatives

Under the current proposal, we should expect to find effects similar to those of L and R but in other domains. That is, whenever we have a complex structure ϕ, we predict that any

[14]Or, at least, it disappears if we assume that uttering ϕ can never be mis-parsed as $\phi L \psi$ or as $\psi R \phi$. Sauerland addresses this concern by a prohibition against using L and R in speech, an effect that he attributes to the Maxim of Manner.

simplification of ϕ will be an alternative to it. For example, a structure that contains a modified noun phrase should have as an alternative the same structure but without modification. Normally, however, we see no such effects, which is one of the reasons for skepticism towards complexity-based theories. But there is a reason why such inferences are rare. In most cases, a sentence with a modified noun phrase will (asymmetrically) entail the variant with no modification:

(63) a. ϕ=A tall man came to the party

b. ϕ'=A man came to the party

Since ϕ is stronger than ϕ', we cannot negate it without contradicting the assertion. To test for complexity effects in implicatures we need to reverse the entailment relations. Under downward-entailing operators, then, we will expect to find implicatures that are based on complexity. In fact, we already saw some such cases in (49) above. Here are some further examples. In each case, the (a) sentence implies that the (b) sentence is not assertable.

(64) a. ϕ = If any tall man comes to the party, he will be disappointed

b. ϕ' = If any man comes to the party, he will be disappointed

(65) a. ψ = Every candidate who sang was elected

b. ψ' = Every candidate was elected

(66) a. ξ = John doubts that many dogs with long tails will be sold

b. ξ' = John doubts that many dogs will be sold

As with disjunction, these inferences pose a challenge to scale-based approaches. If we wanted to capture the implicatures above in terms of scales, we would probably say that *tall man* is a scale-mate of *man*, *candidate who sang* is a scale-mate of *candidate*, and *dogs with long tails* is a scale-mate of *dogs* (and perhaps more generally that any noun phrase that contains a modifier is a scale-mate of the same noun phrase but without the modifier). But this would predict that in upward-entailing contexts we would get implicatures in the opposite direction, from the simple structure to the complex one. This prediction is not borne out: in none of the examples below does the (a) sentence imply that the (b) sentence

is false.[15]

(67) a. ϕ = A man came to every party

b. ϕ' = A tall man came to every party

(68) a. ψ = Each reporter talked to a candidate

b. ψ' = Each reporter talked to a candidate who sang

(69) a. ξ = John is sure that many dogs will be sold

b. ξ' = John is sure that many dogs with long tails will be sold

The (a) sentences suggest nothing about their (b) counterparts. Under the direct approach, using structural alternatives, the absence of inferences about more complex structures, as in (67), is predicted.

As we saw in section 2.3 above, the same pattern holds also for non-monotonic noun phrases. Here are the examples again.

(70) a. I doubt that exactly three semanticists will sit in the audience

b. I doubt that three semanticists will sit in the audience

(71) a. If we meet John but not Mary it will be strange

b. If we meet John it will be strange

(72) (Danny Fox, p.c.)

a. Everyone who loves John but not Mary is an idiot

b. Everyone who loves John is an idiot

In (70) the triggering elements, *exactly three semanticists* and *John but not Mary*, were not monotonic, though their alternatives were. It is just as easy for the alternatives themselves to be non-monotonic:

(73) a. I doubt that exactly three semanticists and exactly two syntacticians will sit in the audience

b. I doubt that exactly three semanticists will sit in the audience

[15]The use of universal operators in these examples guards against the possibility that inferences would be blocked because of symmetry. See section 2.6.2 for discussion.

In total, then, we have a system that dispenses with the notion of scale, derives the effect of L and R for disjunction, and makes new (and correct) predictions about a family of inferences that arise in downward-entailing contexts, inferences that pose a challenge for the monotonicity-based approach and for the very notion of scale.

2.6 Complex alternatives: redefining the substitution source

It is now time to go back to Matsumoto's argument against the use of complexity in implicatures. The argument was based on example (45a) above, repeated here:

(74) a. ϕ = It was warm yesterday, and it is a little bit more than warm today

b. ϕ' = It was a little bit more than warm yesterday, and it is a little bit more than warm today

Our task is to allow ϕ' to be an alternative for ϕ even though it is strictly more complex. I would like to suggest that what licenses ϕ' as an alternative is the fact that, in some sense, its building blocks are already present in ϕ. To obtain ϕ' from ϕ what we need is to substitute *a little bit more than warm* for *warm* in ϕ. The structure for *a little bit more than warm* is not in the lexicon, which prevented us from performing the relevant substitution so far, but it is there as part of the right conjunct of ϕ. If we could enrich the substitution source with the subtrees of the current utterance, we would get what we want. Here is the definition:[16]

(75) SUBSTITUTION SOURCE (final version):

Let ϕ be a parse tree. The substitution source for ϕ, written as $L(\phi)$ is the union of the lexicon of the language with the set of all sub-trees of ϕ

[16]Not all speakers share Matsumoto's intuitions about (74a). Among those who do, some seem to require a special prosodic marking of *warm* and *a little bit more than warm*, similar perhaps to the marking discussed by Geurts (2007) in the context of Levinson (2000)'s examples. I have nothing interesting to say about the source of this inter-speaker variation. For now I will assume that speakers differ in whether they use only the lexicon as their substitution source, as in our original definition, or whether they allow also substitutions outside the lexicon, as in (75), possibly subject to the requirement that substitutions outside the lexicon be prosodically marked in the relevant way.

With our revised definition for $L(\phi)$ in place we can now derive the desired implicature. Since *a little bit more than warm* is a subtree of ϕ, we can use it for substitution and obtain ϕ', which means that $\phi' \lesssim \phi$, appearances to the contrary notwithstanding. Semantically, we may assume that [[a little bit more than warm]]⊂[[warm]]. A speaker who obeys (55) will therefore not assert ϕ if ϕ' is weakly assertable. Assuming that there are no symmetric alternatives to consider, strengthening may now proceed, giving rise to the inference that it was warm but not a little bit more than warm yesterday, and it is a little bit more than warm today.

2.6.1 Concern I: arbitrariness

We have thus derived the inferences which Matsumoto used against complexity-based approaches. Though the change to our original system was not big, it does look like a hack, added to simulate Matsumoto's results without giving up on structural comparisons. We could, perhaps, justify our move by saying that the substitution source is the collection of objects that are available for further syntactic operations. Those include the lexicon, used as a source for the terminal elements in the derivation, but they also include any constituent of the current structure (and maybe nearby structures as well)[17] that could be used for movement, pronominalization, the satisfaction of anaphoricity requirements on focus and ellipsis, and so on. Perhaps. Fortunately, we do not have to rely on such considerations. As

[17]That nearby structures are also available for substitution can be seen from the possibility of replicating Matsumoto's results across discourse. An anonymous L&P reviewer, for example, offers the following exchange, to be imagined as part of a long-distance phone call:

A: I love late spring. It's warm here.

B: It's a little bit more than warm here. I'm inside with the air conditioning.

As in Matsumoto's example, B's utterance implies that it is warm, but not a little bit more than warm on A's side. A similar point can be made using the following exchange, from Gazdar (1979), where B's answer is taken to imply that B's mother is not well:

A: Is your mother well and back?

B: Well, she's back, yes.

A: She's not well then.

it happens, the two proposals are not equivalent. They make different predictions, and it is the predictions of the current one that are borne out.

Under the current proposal the inference in cases like (74a) depends on the presence of the more complex alternative somewhere within the structure. If it is not there, no implicature will be generated. Under a scale-based approach, on the other hand, once a scale-mate always a scale-mate. (74a) would teach us that *warm* and *a little bit more than warm* are scale-mates; after that we would expect to find similar inferences elsewhere, regardless of whether *a little bit more than warm* is present or not. The facts seem to support the complexity approach. When *a little bit more than warm* is absent, no inference about yesterday's temperature is made:

(76) It was warm yesterday, and it is cold today

(77) It was warm yesterday, and it is hot today

It is also worth noting that the current approach, differently from the scale-based one, predicts that more or less any other inference can be triggered if the relevant material is already part of the structure.[18] This, too, seems to be correct: uttering (78) implies that yesterday it was just warm, not warm and sunny with gusts of wind.

(78) It was warm yesterday and it is warm and sunny with gusts of wind today

2.6.2 Concern II: non-scalar items

Matsumoto's example was meant as an argument against using complexity in the computation of implicatures. Upon closer inspection, as we just saw, this example turned out to teach us about the exact notion of complexity that is involved. It also constitutes another argument against the notion of scale. But should the definition of alternatives ignore any non-structural considerations? The patterns that we discussed suggest two conflicting conclusions. On the one hand, we observed that non-monotonic elements can trigger implicatures (as in (49) above), and that the alternatives themselves can be non-monotonic (as in (73)). Those implicatures are often difficult to find, since the simpler alternatives

[18]Subject to considerations of symmetry, as we discuss immediately below.

are usually weaker, but in downward-entailing contexts the implicatures surfaced. On the other hand, we have encountered other data that seem to suggest the opposite: when Matsumoto's example is modified to involve non-monotonic elements instead of the original monotonic ones, no inference arises. The relevant example was (48a) above, repeated here.

(79) a. John talked to some of the girls yesterday, and he talked to some but not all of the girls today

b. John talked to some but not all of the girls yesterday, and he talked to some but not all of the girls today

Uttering the distinctly odd (79a) does not give rise to any inference about (79b), as we saw. Should we then find a way to incorporate monotonicity as a condition in cases like (79a)? I think not. The problem is one of symmetry. Notice that (79b) is not the only alternative to (79a). The following is also a stronger alternative:

(80) John talked to all of the girls yesterday, and he talked to some but not all of the girls today

Combining (79b) and (80) all we get for (79a) is an ignorance inference.[19] In this case, then, monotonicity and brevity make similar predictions. It is possible to tease these predictions apart, though. If monotonicity holds, it should hold in all cases. If the problem is one of symmetry, embedding the structure in a context where the two symmetric alternatives do not exhaust the space of possibilities will solve the problem, and an implicature is predicted to arise. What we should do, then, is modify (79a) under a universal operator, as in (81). The two symmetric alternatives, no longer exhausting the possibilities, are listed in (82)

(81) Yesterday, every boy talked to some of the girls, and today every boy talked to some but not all of the girls

(82) a. Yesterday, every boy talked to all of the girls, and today every boy talked to some but not all of the girls

b. Yesterday, every boy talked to some but not all of the girls, and today every boy talked to some but not all of the girls

[19]Though I do not know why the result is odd.

Judgments about (81) are far from clear. To the extent that the sentence is at all acceptable, though, it seems to imply that yesterday at least one boy talked to all of the girls and that at least one boy didn't. Similar effects arise with other non-scalar items. For example, (83) suggests that yesterday John was required to talk to at least three girls, and that he could have satisfied this requirement by talking to exactly three girls or by talking to more.[20]

(83) Yesterday, John was required to talk to three girls, and today John was required to talk to exactly three girls

(84) a. Yesterday, John was required to talk to four girls, and today John was required to talk to exactly three girls

b. Yesterday, John was required to talk to exactly three girls, and today John was required to talk to exactly three girls

It appears, then, that monotonicity plays no role at all in implicatures. It restricts neither the triggers nor the alternatives, as we discussed in the context of strictly simpler alternatives in section 2.5.3, and now we just saw further evidence that it does not restrict the alternatives. The only relevant factor, it seems, is structure.

2.7 Summary

We started with the symmetry problem, which arises when every stronger alternative can be negated, as in the naïve conversational principle (40), repeated here.

(85) Conversational principle (naïve version, repeated from (40)): do not assert ϕ if there is another sentence ϕ' such that both

a. ϕ' is more informative than ϕ, and

b. ϕ' is weakly assertable

We followed the neo-Gricean idea of restricting the set of alternatives, though we saw evidence against the more familiar implementations in terms of scales and monotonicity.

[20] I use a modal rather than a quantifier over individuals as the universal operator in (83) following a suggestion by Danny Fox, p.c., who points out that modals help sharpen the judgments about the inferences in these sentences. I have no account for the difference between the operators.

Instead, a structural condition on alternatives seemed to work: the alternatives of ϕ are all the structures that are at most as complex as ϕ. We noticed that in many cases, simpler alternatives are also weaker, and hence irrelevant for the computation of implicatures. When the simpler alternatives are stronger, as with disjunction or with weak alternatives in downward-entailing contexts, the predicted inferences arise. Moreover, we found inferences of this kind that arise with respect to non-monotonic expressions. Our definitions made no reference to the notion of scales, and in fact we noticed that deriving the alternatives directly helped us avoid some of the challenges that are faced by accounts that use that notion. Finally, in addressing Matsumoto's argument against complexity, we had to revise our definition of substitution source. This revision led to certain predictions that were different from those made by Matsumoto's account. As it turned out, the correct predictions were the ones of the current account, providing an additional argument in favor of the complexity-based approach and against the use of scales and monotonicity.

Coming back to the general claims of this dissertation, the current chapter has provided evidence that, at least in the domain of implicatures, structural complexity matters. This point was made independently of choices of locality and without committing to a particular preference for simplicity. I also tried to isolate the discussion from considerations that arise by looking at alternatives elsewhere. These are messier issues, and it was convenient that we could get away without discussing them. But we cannot avoid these issues much longer if we want to get beyond the general claim that complexity matters.

Chapter 3

Implicatures and Focus

(Joint work with Danny Fox)

In chapter 2 we discussed the way alternatives are constrained when used for the implicature computation. It is standardly assumed, however, that implicatures are only one instance of a more general family of computations that make reference to alternatives. In particular, focus semantics is thought of being sensitive to alternatives in much the same way as implicatures are. Both phenomena involve associating a linguistic object ϕ with a set of alternatives $A(\phi)$, and both involve strengthening the meaning of ϕ by negating elements in $A(\phi)$. We will make this idea more precise as we proceed, but for now the following vague statement will suffice.

(86) Background assumption: implicatures and focus are sensitive to alternatives in similar ways

A second assumption that is standardly made concerns the roles of grammar and context in determining the alternatives.

(87) Relevant, formal, and actual alternatives:[1]

a. Context provides a set of relevant alternatives, R

[1]The formulation in (87) follows the general spirit of the proposals of Westerstahl (1984), Rooth (1992), and von Fintel (1994), among others, though we have abstracted away from important specifics. For example, Rooth analyzes F as part of the asserted content and R as presuppositional. We hope that our arguments below are not affected by such distinctions. The discussion in the literature about the division of labor is mostly

b. Grammar provides set of formal alternatives, F

c. The actual alternatives, A, are defined as $A = R \cap F$

We will accept the general picture expressed in (86) and (87). Our goal in this chapter is to sharpen this picture by making two further points regarding the division of labor between context and grammar. First, we will propose a constraint on the set of relevant alternatives R, both for implicature and for focus. Specifically, we will argue that R is closed under negation and conjunction. Second, we will present evidence that F, too, is the same for both processes.

(88) Claim: for both implicature and focus,

a. R, is closed under $\neg$, $\wedge$

b. F is determined in the same way

Since the set A of actual alternatives is defined as the intersection of R and F, our claims will entail that the actual alternatives are the same for implicature and for focus. This, in turn, will provide concrete content to the vaguely stated (86). The particular balance between R and F we will arrive at will involve a certain shift in the division of labor that is often assumed.

3.1 Relevance: a constraint on pruning the alternatives for implicature

3.1.1 Back to symmetry in implicature

Let us start by revisiting an argument that we discussed in chapter 2.

(89) It was warm yesterday, and it is a little bit more than warm today

(90) # Yesterday John did some of his homework, and today he did some but not all of it

concerned with the semantics of quantifiers and of focus. Including implicature in this discussion does not strike us as unnatural (see, in particular, Horn, 1989).

Sentence (89), due to Matsumoto (1995), gives rise to the inference that yesterday it was not a little bit more than warm. The superficially similar (90), on the other hand, does not give rise to an inference that yesterday John failed to do some but not all of his homework (that is, the sentence does not suggest that yesterday John did all of his homework).

We discussed two possible ways to account for the contrast between (89) and (90). The first option, following Matsumoto, as well as Horn (1989), is to say that the problem for (90) is one of insufficiently many alternatives. While *warm*, in the first conjunct of (89), has *a little bit more than warm* as an alternatives, *some of his homework* in the first conjunct of (90) does not have *some but not all of his homework* (or *some but not all of it*) as an alternative.[2]

The second option we mentioned is to say that the problem with (90) is not that it has too few alternatives but rather that it has too many of them. This was the option argued for in chapter 2. The idea was that in (90), as in (89), the more complex element in the second conjunct is available as an alternative for the first conjunct. That is, *some* will have *some but not all* as an alternative. However, *some* also has *all* as an alternative, as it does in familiar cases of implicature. Having both *some but not all* and *all* as actual alternatives to *some* gives rise to symmetry (cf. Fox, 2006, Spector, 2006a).

(91) a. Yesterday John did some but not all of his homework, ...

b. Yesterday John did all of his homework, ...

How do we choose between the two options? The option of too many alternatives makes the prediction that as soon as symmetry is broken, we should be able to see *some but not all* in action as an alternative to *some*. The option of too few alternatives predicts that *some but not all* would never be there, so breaking symmetry should not matter. Section 2.6.2 above presented evidence that the problem was one of too many alternatives:

(92) Yesterday John was required to do some of his homework, and today he was required to do some but not all of it

[2]As mentioned earlier, the question behind this discussion is whether the alternatives are constrained by complexity or whether, as Matsumoto (1995) argues, the appropriate constraint is semantic.

Sentence (92) gives rise to an inference that yesterday John was not required to do all of his homework, nor was he required to do some but not all of it. This is as predicted by the idea of too many alternatives: the universal modal (*require*) breaks the symmetry, so the negations of the two stronger alternatives below, (93a) and 93b, are compatible with each other and with the assertion.

(93) a. Yesterday John was required to do some but not all of his homework, ...

b. Yesterday John was required to do all of his homework, ...

It will be convenient in what follows to have names for the two situations of symmetry that we saw. For (90) we had an instance of what we will refer to as *true symmetry*: given the assertion, we could not negate the alternatives and stay consistent.[3] Embedding under a universal modal, as in (92) resulted in a more benign situation: negating the embedded alternatives was consistent with the assertion. We will refer to a context that embeds a symmetric set but is not necessarily fatal as *symmetry embedding*.

(94) Definition: Symmetry

Let α be a semantic object of a type τ that ends with t, and let $A = \{\alpha_1, \ldots \alpha_n\}$ be a set of semantic objects of type τ.

a. We will say that A is *truly symmetric* with respect to α if the set $\{\alpha, \neg\alpha_1, \ldots, \neg\alpha_n\}$ is inconsistent

b. If A is truly symmetric with respect to α, then for any semantic context $\phi[\cdot]$ that can embed elements of type τ we will say that the set $\Phi = \{\phi[\alpha_1], \ldots, \phi[\alpha_n]\}$ is *symmetry embedding* with respect to $\phi[alpha]$

[3]True symmetry does not necessarily block implicatures. In disjunction, for example, stronger alternatives to $\phi \vee \psi$ include $\phi, \psi, \phi \wedge \psi$, which cannot all be negated consistently with the assertion. However, while there appears to be no non-arbitrary way to choose between ϕ and ψ, the alternative $\phi \wedge \psi$ is stronger than either and can safely be negated (it is *innocently excludable* in the sense of Fox 2006). This explains the familiar exclusive reading of *or*. In (90) above, on the other hand, no alternative was innocently excludable. We may perhaps refer to true symmetry with no innocently excludable alternatives as *fatal symmetry*.

3.1.2 A problem with pruning

Showing that (90) involves true symmetry was as far as chapter 2 went in terms of arguing for the option of too many alternatives. Here we would like to discuss a remaining concern. In the literature on alternatives (see especially Rooth, 1992, as well as Horn, 1972, 1989), the set of actual alternatives A is not determined by F alone, and can be further restricted by context, which provides a set of contextually relevant propositions (or properties) R. For example, as Horn (1972, p. 112) discusses (though in a somewhat different context), the use of *some*, as in (95), can give rise to the implicature that the *most*-counterpart, as in (96a), is false, showing that (96a) is a possible alternative for (95). However, this alternative can be ignored while computing the implicatures of (96a): it is possible for (96a) to imply only that (96b) is false without suggesting anything about the truth of (96a).

(95) Some of my friends are Zoroastrians

(96) a. Most of my friends are Zoroastrians

b. All of my friends are Zoroastrians

In other words, even though (96a) belongs to the set F of formal alternatives for (95), as the possibility of the *not most* implicature shows, it does not have to be in A, the set of actual alternatives to (95), as the possibility of obtaining the *not all* implicature without the *not most* implicature. This can be taken to show that R, the set of contextually relevant alternatives for (95), can exclude (96a), ensuring that it will not make it to A.

If hearers of (90), like hearers of Horn's example, could choose to consider one of the alternatives and ignore the other, they would be able to infer that the other alternative holds. But these inferences are not available: (90) is just odd. We seem forced to say that both symmetric alternatives for (90) are unavoidable.

This conclusion is somewhat surprising. The problem is not in having *some but not all* and *all* in the set F of formal alternatives to *some*. The proposal in chapter 2, for example, provides us with a characterization of F that has this property. The problem is that we need the symmetric alternatives not just in F but in A, the set of *actual* alternatives. This, in turn, means that both alternatives are also in R, the set of contextually relevant alternatives. The problem, then, is in preventing contextual pruning from excluding one of

the symmetric alternatives from R, while still allowing it to exclude alternatives in cases like Horn (1972)'s example.

3.1.3 Proposal: a definability constraint on relevance

Here is a potentially important distinction that we may want to use. The alternatives for (90) are inter-definable: *all* = *some*∧¬*some-but-not-all*. In the cases where pruning is possible, on the other hand, inter-definability does not hold: we cannot define *most* in terms of *some* and *all*.

If this distinction is meaningful, we could state the following constraint on pruning of alternatives: if a set of alternatives $A \subseteq F$ is being considered, then any $\beta \in F$ that is defined using elements in A by $\neg$ and $\wedge$ must also be considered. We propose to state the constraint as a property of relevance:

(97) Proposal: R (the set of *relevant* propositions) is closed under $\neg$ and $\wedge$

There are reasons to think that this is a natural notion of relevance. Closure under conjunction seems innocuous: if two propositions are relevant, it is hard to imagine how their conjunction will not be relevant. Closure under negation follows if for a proposition to be relevant we want to know whether or not it is true (cf. Groenendijk and Stokhof, 1984). Note that we could not say that it is the set of formal alternatives F that is closed under $\neg$ and $\wedge$, since that would give rise to symmetry for sentences like *John did some of his homework*, contrary to fact. If our argument is correct, then, it supports the idea that relevance alone is not sufficient for defining the alternatives for implicature.

3.1.4 Interim summary

Let us summarize the argument. There is a theory of F for the implicatures we have been looking at. According to this theory, a sentence with *some*, call it ϕ[*some*], will always have the corresponding sentence with *all*, ϕ[*all*], as an alternative in F. Generally, F is not closed under boolean operations (this, we take it, is why implicatures are possible in the first place). In certain contexts, however, as in the Matsumoto (1995)'s example and its variants discussed above, ϕ[*some but not all*] can also be in F. This leads to *symmetry*

embedding, which can involve true symmetry or not, depending on the logical properties of the embedding contexts. For example, we saw that a universal quantifier can prevent true symmetry. The actual alternatives A are determined not only by F but also by R. Since R *is* closed under boolean operations, it can never help in avoiding true symmetry. Consequently, if the context makes ϕ[*some but not all*] $\in F$, and if the logical properties of the embeddining context $\phi[\cdot]$ make ϕ[*some*] $\wedge \neg\phi$[*some but not all*] $\wedge \neg\phi$[*all*] inconsistent, no impicature of $\neg\phi$[*some but not all*] will arise.

3.2 Symmetry and relevance in focus

3.2.1 The division of labor between R and F

As with implicatures, we may ask whether focus alternatives sometimes involve symmetry, and if so, whether this symmetry is fatal. In fact, given familiar proposals in the literature, we may expect symmetry to be more prevalent for focus, since there is usually no equivalent to scales that could rule out a symmetric alternative. We will suggest that the pattern of symmetry and its avoidance is accounted for in exactly the same way for focus and for implicatures.[4]

Here is a simple example where we might expect symmetry with focus:

(98) a. What did Mary do (for this class)?

b. (Very little.) She only $[_{VP}$ read three books$]_F$

Rooth (1992) suggests that the actual alternatives (A in our earlier notation) are determined by two factors: formal alternatives (F in our notation, which for Rooth would be defined semantically as the set of all property denotations); and contextual relevance (our R). For example, if *did her homework* and *read four books* are the contextually relevant alternatives for the focused VP in (98b), we could derive the meaning that Mary didn't read more than three books. Crucially, for this reasoning to go through, *read three books and not read four books* (amounting here to *read exactly three books*), which is symmetric to *read*

[4]If this is correct, the theory of focus alternatives will have to be revised in interesting ways; we will get to that below.

three books, cannot be in A. For Rooth (1992), this symmetric alternative is in F (since he takes the definition to be in terms of semantic type), so it is up to R to filter it out. If Rooth is right about this, R (as defined for focus) cannot have the same closure properties as the notion of relevance that we defined for implicatures above. Are we forced to say that there are two different notions of relevance?

3.2.2 An argument for having the same R and F for focus and for implicature

As mentioned in the introduction, we think that the notion of relevance for focus is the same as the one that we had for implicature. Evidence comes from the possibility of replicating with focus the same symmetry-based pattern that we saw earlier with implicatures:

(99) a. # In last week's robbery they only [stole the books]$_F$. In today's robbery they [stole the books but not the jewlery]$_F$.

b. By last week they only found out that the robbers [stole the books]$_F$. Today they found out that the robbers [stole the books but not the jewlery]$_F$.

The sequence in (99a), a focus-based variant of (90) above, cannot be taken to mean that in last week's robbery it was not the case that they did not steal the jewlery: this would have meant that in last week's robbery they stole the books *and* the jewlery, a reading that (99a) does not have. The sequence in (99b), on the other hand, a focus-based counterpart of (92), does entail that by last week they had not figured out that the robbers did not steal the jewlery.

If, as we are proposing, the same R and F we had for implicature are also used for focus, we will be able explain the contrast between (99a) and (99b) in the same way as we did for (90) and (92). In a configuration of this kind, the more complex alternative is part of F. This alternative, *stole the books but not the jewlery* in the current case, is also in R. By the closure property of relevance, the alternative *stole the books and the jewelry* is also in R. Since A is assumed to be the intersection of F and R, *stole the books but not the jewlery* and *stole the books and the jewlery* will both be among the actual alternatives, leading to fatal

symmetry. In (99b), on the other hand, embedding under the modality of *find out* ensures that, while the actual alternatives are symmetry including, no fatal symmetry will arise.

For an account such as the one proposed by Rooth (1992), where symmetry is broken by R, the contrast between (99a) and (99b) is surprising. On such an account, F is symmetric for any focused element of a type ending in t. This is the case for the focused VP in (99a) and (99b). The symmetry of F, on the semantic approach, is supported by the fact that (99b) implies that by last week they had not found out whether or not the robbers stole the jewelery. Since on such an account R is capable of breaking symmetry (and in particular it is not subject to our proposed closure condition), we would expect it to be able to rescue (99a) by filtering out the problematic alternative *stole the books and the jewelery*. This alternative is not explicitly mentioned, and its presence gives rise to fatal symmetry. However, (99a) is odd, suggesting that the offending alternative is not pruned.

Summing up, it appears that the situation for focus is parallel to what we saw earlier for implicatures. In particular, we conclude that both processes rely on the same division of labor between contextual relevance and formal alternatives: relevance is closed under negation and conjunction, rendering it useless for purposes of symmetry avoidance, while the set of formal alternatives is often, but not always, symmetry-free.

3.3 A unified theory of formal alternatives for focus and implicature

3.3.1 Using structurally defined alternatives for focus

The view that it is the role of F to break symmetry seems more natural under a structural conception of alternatives than under a semantic approach. In particular, for any type τ that ends in t the domain D_τ is closed under negation and conjunction, so defining F in terms of semantic type, along the lines of Rooth, will not have the required symmetry-breaking properties. For implicatures, we saw in chapter 2 that there are independent reasons for adopting a structural characterization of F, and we saw arguments for a definition in terms of complexity:

(100) STRUCTURAL COMPLEXITY (repeated from (100)):

Let ϕ, ψ be parse trees. If we can transform ϕ into ψ by a finite series of deletions, contractions, and substitutions of constituents in ϕ with constituents of the same category taken from $L(\phi)$, we will write $\psi \lesssim \phi$. If $\psi \lesssim \phi$ and $\phi \lesssim \psi$ we will write $\phi \sim \psi$. If $\psi \lesssim \phi$ but not $\phi \lesssim \psi$ we will write $\psi < \phi$.

(101) STRUCTURAL ALTERNATIVES (repeated from (54)):

Let ϕ be a parse tree. The set of *structural alternatives* for ϕ, written as $A_{str}(\phi)$ is defined as $A_{str}(\phi) := \{\phi' : \phi' \lesssim \phi\}$.

To allow F to break symmetry in focus, we can make use of the same definition of structural alternatives in (101). However, we should also talk about the embedding of a focus-marked phrase ϕ_F within a phrase ψ that is not focus-marked. As in the Roothian definition of F we would like to say that the focus semantic value of ψ is the set of meanings obtained by replacing ϕ_F with its alternatives. The only difference is that instead of replacing ϕ_F with every element of the same semantic category as ϕ we would like to restrict the alternatives to $A_{str}(\phi)$. It will be convenient to define the focus semantic value of ψ using an intermediate stage of focus syntactic value:[5]

(102) Let ϕ_F be a maximal focus-marked constituent embedded in ψ.

a. The focus *syntactic* value of ψ, $F(\psi)$, is the set of structures obtained from ψ by replacing ϕ with elements in $A_{str}(\phi)$: $F(\psi) := \{\psi[\phi \mapsto \phi'] : \phi' \in A_{str}(\phi)\}$

b. The focus *semantic* value of ψ, $[\![\psi]\!]^f$ is the image of $F(\psi)$ under the usual semantic interpretation function $[\![\cdot]\!]$: $[\![\psi]\!]^f := \{[\![\psi']\!] : \psi' \in F(\psi)\}$

3.3.2 Using focus values for implicature

We used A_{str}, as defined in (101), to restrict the formal alternatives for focus so as to break symmetry, along the same lines as discussed for implicatures. Having done this, we can now use these focus alternatives in those places where they were used under the semantic

[5]The definition in (102) assumes a single maximal focus-marked phrase. Generalizing to an arbitrary number of such phrases is straightforward: If $\phi_1, \ldots, \phi_k$ are the maximal focus-marked constituents embedded in ψ then $F(\psi) := \{\psi[\phi_1 \mapsto \phi'_1, \ldots, \phi_k \mapsto \phi'_k] : \phi'_1 \in A_{str}(\phi_1), \ldots, \phi'_k \in A_{str}(\phi_k)\}$

definition. One such place is the computation of scalar implicature. As noted by Rooth (1992), the placement of focus affects the possible inferences that an utterance gives rise to:

(103) How did the exam go?

a. Well, I [passed]$_F$

b. Well, [I]$_F$ passed

The two answers in (103), differing only in the placement of focus, license different inferences. (103a), with focus on the *VP*, suggests that the speaker only passed, rather than aced the exam; it suggests nothing about whether other people passed or not. (103b), with focus on the subject, suggests that some other people did not pass, without implying anything about whether the speaker did better than just pass. In each case, the alternative that is used to compute the implicature is obtained from the utterance by a substitution *within the focus-marked phrase*. Rooth suggests that this is true more generally, and that the alternatives for scalar implicature are derived from the assertion by substitutions that are confined to focused phrases.

Here is another example that makes the same point. In (104), with focus on the *VP*, the speaker implies that they do not believe that John talked to both Mary and Sue yesterday. This is the familiar scalar implicature for disjunction. This implicature disappears, or is at least weakened, in sentences like (105) where the *VP* is not focused.

(104) a. What did John do yesterday?

b. He [talked to Mary or Sue]$_F$

(105) a. Who talked to Mary or Sue yesterday?

b. [John]$_F$ talked to Mary or Sue yesterday

3.3.3 Aside: a recursive definition of focus syntactic values

Definition (102) above leaves open the question of how the focus values of complex structures are arrived at compositionally. The same question has been discussed for semantic accounts of focus, and two general answers have been suggested in the literature. One ap-

proach, outlined by Chomsky (1976), assimilates the configuration of a focused phrase ϕ embedded in a larger structure ψ to that of variable binding. Informally speaking, ϕ can be thought of as contributing a variable, as would be the case if it were a quantifier or a *wh*-element. The other approach, proposed by Rooth (1985), builds the alternative set of focus semantics directly, using a recursive definition.[6]

As far as we can tell, our proposal is compatible with either approach. However, we would like to point out that the second approach, in which focus values are built recursively, suggests an interesting way to dispense with a separate definition of the set of structural alternatives $A_{str}(\cdot)$. The definition in (107), a variant of Rooth's, demonstrates this point. It constructs focus syntactic values directly, without recourse to $A_{str}(\cdot)$, by using substitutions at the level of terminal items only. The idea that (107) tries to capture is that we can simulate the deletion of a phrase ϕ by replacing all of the terminals dominated by ϕ with empty nodes, and that we can simulate the substitution of a phrase ϕ' for ϕ by substituting ϕ' for the head of ϕ and deleting the rest of ϕ. For purposes of presentation, we will make certain simplifying assumptions, summarized in (106), hoping that they would not affect the general point.

(106) Auxiliary assumptions for the recursive definition of focus

a. If ϕ is a terminal of category X, the possible substitutions for ϕ, $L(\phi)$ contains all lexical entries of category X, all salient constituents of category X, and the empty set $\emptyset$

b. If M is a syntactic mode of combination, and if $M(\ldots, \phi_i, \phi_{i+1}, \ldots)$ is well-defined, then $M(\ldots, \phi_i, \emptyset, \phi_{i+1}, \ldots)$ is also well-defined, and $M(\ldots, \phi_i, \emptyset, \phi_{i+1}, \ldots) = M(\ldots, \phi_i, \phi_{i+1}, \ldots)$

c. For any syntactic mode of combination M, $M(\emptyset) = \emptyset$

d. A nonterminal ϕ is focus-marked if and only if all of its daughters are focus-marked

e. Each phrase has a head

(107) A recursive definition of the focus syntactic value of ϕ, $F(\phi)$:

[6]A similar idea was explored by Hamblin (1973) in the context of the semantics of questions.

a. The focus syntactic value of a non-focused terminal ϕ is the singleton set of ϕ: $F(\phi) = \{\phi\}$

b. The focus syntactic value of a focused terminal ϕ_F is the set of possible substitutions for ϕ: $F(\phi) = L(\phi)$

c. The focus syntactic value of a nonterminal $\phi = M(\phi_1, \ldots, \phi_k)$, where $\phi_1, \ldots, \phi_k$ are the daughters of ϕ and M their mode of combination, is $F(\phi) = \{M(x_1, \ldots, x_k) : x_1 \in F(\phi_1), \ldots, x_k \in F(\phi_k)\}$

3.3.4 Summary

We saw that focus is similar to implicature in requiring symmetry to be broken in the set of *formal* alternatives rather than in the set of contextually relevant alternatives. We proposed using a structural characterization of the formal alternatives, originally developed for implicature, as a more general solution that applies also to the formal alternatives for focus. We then followed Rooth in the opposite direction, using focus as a constraint on the alternatives for implicature. The emerging picture is this:

(108) For both focus and implicature,

a. The set F of formal alternatives is defined as the focus (semantic) value, as in (102)

b. The set R of contextually relevant alternatives is subject to the closure condition in (97)

c. The set A of actual alternatives is defined as $F \cap R$

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 4

Blocking: a direct preference for simplicity

4.1 Review: blocking and structure

The goal of this chapter is to provide evidence for a preference for simpler structures over complex ones. So-called blocking effects, which we briefly discussed in the introduction, seem like a natural place to start. As we saw, many approaches endorse some notion of competition to explain these effects, and some of them express an intuition that is quite similar to simplicity (a listed form can be thought of as easier to use than a form that must be computed from scratch; a word is in some sense simpler than a phrase; and so on). As we saw, however, there are some compelling arguments against what seem to be the most straightforward implementations of the notion of simplicity. In particular, we reviewed Poser (1992)'s arguments against using phonetic effort (*children* is no simpler than **childs*) and against using a Gricean notion of pragmatic competition for morphology (the losers in morphology are not just odd but downright ungrammatical, and a winner can survive despite a significantly longer paraphrase). We also reviewed Embick and Marantz (2008)'s arguments against a global notion of competition as an explanation for *do*-support (the absence of a synthetic verbal form does not improve an otherwise inappropriate analytical form: **did forgo* is no better than **did walk*), as well as against Poser's more local notion of competition between words and phrases (PP-complements do not help an analytical form:

??more proud of his achievements is no better than the PP-less *??more proud*).

Embick and Marantz's conclusion was that competition exists, but only at the extremely local level of individual nodes, where something like the *Subset Principle* (Halle, 1997; cf. also Kiparsky, 1973) governs vocabulary insertion. Since under this view (which is largely also the general view within Distributed Morphology) competition is restricted to a single node, the notion of structural complexity defined in chapter 2 will not be able to distinguish between the candidates.

A brief clarification is in order. The Subset Principle can sometimes give rise to indirect structural-complexity effects. Since the Subset Principle prefers an item that matches more features of the node over an item that matches fewer features, the loser would always leave some extra features unmatched. If something like Noyer (1992)'s notion of *fission* is used, these unmatched features can then be realized by an additional morpheme. This means that if an otherwise better item cannot be inserted for some reason, the result would be the insertion of a less-good item plus an extra morpheme. This, in turn, means that the better item is at most as complex as (and sometimes strictly less complex than) a less-good item.

The preference obtained through fission is epiphenomenal, however, and in any case it remains a highly local matter. The main claim of this chapter is that, in some cases at least, blocking effects involve a preference for less complex structures that is more direct, more pervasive, and less local than the one suggested by Embick and Marantz. In addition to showing why this preference is needed, I will try to explain why the criticisms made by Poser and by Embick and Marantz do not apply to the proposal developed here.[1]

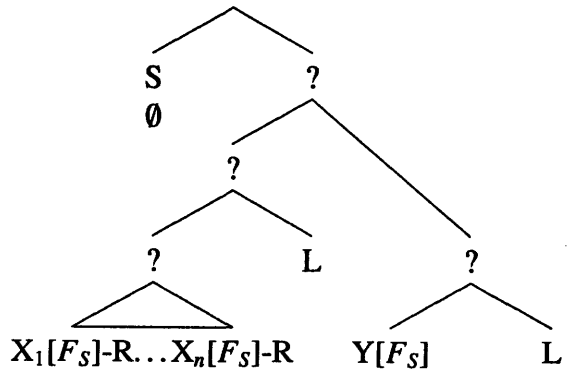
My evidence for a direct preference for simplicity in morphology comes from certain mismatches between morphology and semantics. I will be particularly concerned with mismatches that involve definiteness marking within noun phrases in different languages, where I will try to show that what looks like function heads bearing the semantics of definiteness are better treated as pieces of morphology that do not express definiteness (and might not have any meaning at all) and that are subject to a condition of structural econ-

[1]The preference for less-complex structures incorporated into the proposal below will also be more direct than the one argued for in chapter 2, and a natural question that will arise, though one that I will not answer here, is whether the two notions of complexity-sensitive competition are compatible with each other.

omy. These pieces of morphology, which I will refer to as *licensors* will be related to the semantics of definiteness only indirectly, through a mechanism that I will discuss in detail below. However, I will not try to derive the behavior of licensors from deeper principles, and I will simply stipulate their existence alongside the more familiar kinds of function elements: semantically interpreted function heads, which I will refer to as *spreaders* for reasons to be explained shortly; and meaningless elements expressing features in their immediate surroundings, which I will refer to as *realizers*.

The current proposal, then, involves a larger inventory of function elements than commonly assumed, including not only spreaders and realizers but also licensors. I will refer to this proposal as the SRL approach, and the general idea that I will try to pursue is that for the mismatches under discussion, spreaders, realizers, and licensors all play a role, and that the distribution of licensors is governed in part by structural economy. A potentially covert spreader S will attach at a high position, where it makes a semantic contribution and also spreads a feature F_S onto elements below. Elements that have received F_S from the spreader above will sometimes express it overtly with a semantically vacuous realizer R.[2] Finally, a licensor L will be able to attach at certain positions, each one licensing all instances of F_S within a structurally-defined domain. I will try to show that a particular notion of c-command provides a correct characterization of licensing domains. A schematic configuration for SRL is presented in (109).

(109)



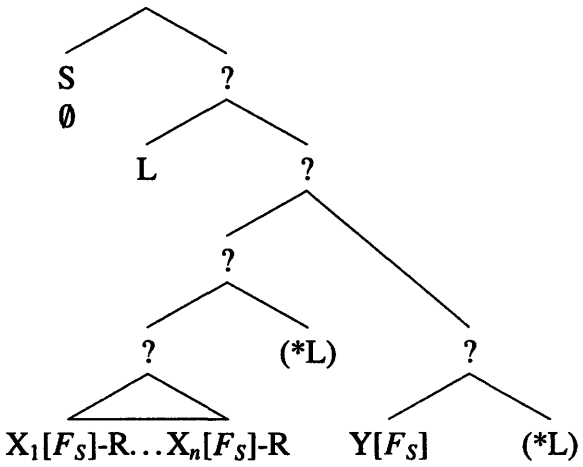
In (109), the spreader S has spread F_S onto $X_1, \ldots, X_n$, each of which surfaces with its own

[2]I will have little to say about what determines whether an element X with a feature F surfaces with a realizer, as $X[F] - R$, or without a realizer, as $X[F]$.

realizer R, and onto Y, which does not surface with a realizer.[3] The attachment of the first licensor allows it to c-command all of the instances of F_S on the various X_i's. For F_S on Y, a second licensor is needed.

Suppose now that we could attach, through whatever means, a single instance of L at a position that allowed it to c-command all the X_i's as well as Y. This would provide a more economic way to license the features without affecting the semantics. Consequently, the lower instances of L will become ungrammatical.

(110)



We should not be pleased by having to stipulate a new kind of function element and the conditions that govern its distribution, and I will devote the rest of this chapter to explaining why I think we are justified in doing so. The first set of data will come from definiteness marking in Danish, where I will try to show that the stipulations of the SRL approach are no worse than those that more conservative approaches have been forced to make, even while restricting our attention to the basic pattern of a single noun and up to one modifying adjective. As soon as the patterns get more complicated, by looking at other modifiers and at a partially-overlapping pattern of gender marking, the advantages of SRL become much clearer. Whereas licensor-free approaches have to stipulate increasingly many conditions to accommodate the familiar facts discussed in the literature, SRL captures these patterns with little or no modification, making new predictions along the way. These predictions extend beyond Danish, leading us, in section 4.6, to a straightforward (and to my knowledge novel) account of the complex paradigm of number, gender, case, and definiteness marking

[3]As mentioned, I am not taking the appearance or absence of a realizer as relevance to the current proposal.

in Icelandic in terms of a change in the labeling of a single node within the Danish noun phrase.

Finally, in section 4.7, I will try to address a familiar concern about morphological analyses in general. Even when we find a simple story about a complex paradigm, it is often hard to avoid the question of whether this story is really part of the grammar, or indeed whether there is any interesting system at all behind the pattern. For all we know, the whole paradigm could in principle be learned as an arbitrary collection of unrelated facts.

Looking at interactions of the paradigm with factors that are assumed to be independent can be of help. The interactions with different modifiers within the noun phrase, as well as with different marking patterns and different phrase-structural choices, will serve this purpose. These interactions will increase considerably the complexity of the patterns, stretching the plausibility of an account of the data in terms of arbitrary memorization. This increase, however, will still be limited by the inventory of available modifiers, different patterns, and choices about phrase structure. Moreover, it is not always obvious that manipulating these factors will keep all else constant. What is missing is the possibility, familiar from syntax, where it is the norm, to test a hypothesis (or rather, compare two competing hypotheses) by introducing arbitrary levels of complexity in one factor while keeping everything else fixed.

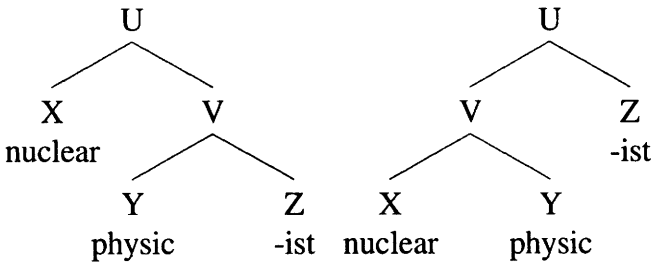
I will suggest that Greek allows us to submit SRL to this kind of test. In particular, I will try to show that if we analyze Greek as a variant of Danish in which we can move elements within the noun phrase, we can derive the full pattern of definiteness spreading, a pattern of involved morpho-semantic mismatches and their interaction with word order and discourse conditions that has received a certain amount of attention in the literature but has remained an open problem.

Before we turn to the details of SRL and its application to particular patterns, let us take a detour through the general phenomenon of mismatches.

4.2 Mismatches

Apparent conflicts between two sets of considerations are familiar from the domain of *wh*-movement (the *wh*-phrase appears at a high position but is subject to selectional and binding-theoretic restrictions that correspond to a different position, that of the 'gap' or 'trace'), from scopal interactions between quantifiers (a quantifier satisfies the selectional and binding-theoretic restrictions of its surface position, but its scope can correspond to either a higher or a lower position), and from many other places. In morphology, mismatches often take the form of so-called *bracketing paradoxes* (Pesetsky, 1979, 1985; Kiparsky, 1983; Williams, 1981), where two kinds of diagnostics suggest two distinct bracketings of the same surface string. Here is an example:

(111) (Morphology) (Semantics)



Morpho-phonological considerations suggest that *physic* first combines with the suffix *-ist*, resulting in the noun *physicist*, which in turn combines with the adjective *nuclear*. Semantic considerations, on the other hand, suggest that *physic* first combines with the adjective, and that the suffix is combined with the result. It is perhaps worth mentioning that the mismatch in (111) depends on our views on locality in both morphology and semantics. In particular, in drawing the two structures I made the (fairly common) assumption that constraints on morphological affixation and semantic composition are stated in terms of sisterhood in a binary-branching structure. If we wish to maintain these assumptions, we will have to find a way to explain how the mapping between the two structures in (111) works. For example, Pesetsky (1985) suggests that the representation corresponding to semantics is derived from the one corresponding to morphology through an LF-movement operation, akin to the rule of Quantifier Raising. This approach has the advantage of preserving a predictive theory of both morphology and semantics, and of reducing the mismatch to an independently moti-

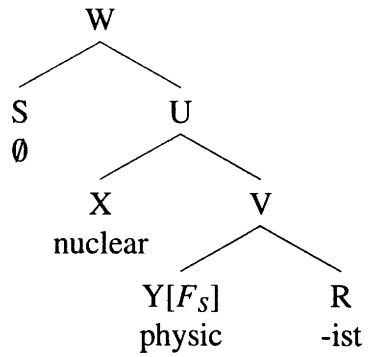
vated operation that is fairly well understood (though, as noted by Pesetsky, the bracketing variant of LF-movement appears to be sensitive to string adjacency). More recently it has been suggested that some mismatches can be resolved by movement at PF. In particular, Embick and Marantz (2008) argue that apparent mismatches in Scandinavian languages involve a post-syntactic operation of Local Dislocation (LD; see also Embick and Noyer, 2001 and Embick, 2007). We will examine the Scandinavian data and the operation of LD in detail below.

An alternative route would be to loosen our assumptions about either morphological affixation or semantic composition. For example, we could say that the only structure available for *nuclear physicist* is the one corresponding to sisterhood constraints in morphology, and that semantic interpretation can combine elements that are not sisters. This would have the advantage of eliminating the mismatch altogether, though it runs the risk of making semantics less predictive. An attempt to loosen the semantic component in a principled way is provided by Moortgat (1987), who argues for the introduction of function composition into the semantic component as a way to maintain a compositional account of surface morphological representations.

The two approaches to bracketing paradoxes sketched immediately above differ in many essential details, but they have the following in common: at one point or another, each piece of morphology gets interpreted. In particular, the function element *-ist* was assumed to make a semantic contribution that corresponds to the interpretation that is intuitively associated with this suffix. In the introduction I referred to such elements as *spreaders*.

The assumption that *-ist* is a spreader is of course perfectly plausible, but one could imagine things being different. For example, one might consider an analysis along the lines of den Dikken (2002), in which the semantic denotation that we usually associate with the suffix *-ist* is contributed by a covert head that is higher than both the noun and the adjective, and that the suffix *-ist* is just a meaningless piece of morphology. In terms of our earlier discussion, *-ist* will now be a *realizer* rather than a spreader, and it will express a feature F_S on *physic* that has been spread onto the noun from the covert spreader S above:

(112) (Indirect association)



The realizer analysis of *-ist*, as schematized in (112), involves a less direct association between *-ist* and the semantics than what was assumed under the spreader analyses mentioned earlier. As hinted at in the introduction, one could imagine loosening the association between *-ist* and the semantics (and the feature F_S) further. The idea, if we were to apply it to *nuclear physic-ist* (though I will not try to claim that this is the correct analysis for this particular example), would be that, as in the structure for the realizers approach in (112), a covert spreader S spreads F_S onto *physic*, but differently from the realizers approach, we could say that F_S requires some form of *licensing*, and that *-ist* is a *licensor* that can provide it.

A licensor is different from a realizer in being less local: a realizer is typically a sister of the element that bears the features that it expresses; a licensor, on the other hand, can be further away from the elements bearing the licensed features (though it can also be close, as in (112). A related difference is that a realizer can express at most the features of one element, while a licensor can license the features of as many elements as there are within its domain. A licensors analysis of *-ist* would lead us to expect that there may be other positions at which we would find *-ist*, and that, if *physic* combined with some other element that also required *-ist*, then it might be possible for a single instance of *-ist* to license F_S on both elements. The realizer account, on the other hand, would predict that each element would surface with its own instance of *-ist*, each instance realizing the F_S of its sister. Finally, and most importantly for our general claims, licensors are special in being subject to a condition of structural economy. Under a licensors analysis of *-ist*, if a single instance of it can license two distinct instances of F_S, an otherwise grammatical

structure in which each F_S has its own licensor would become ungrammatical.

I used *nuclear physic-ist* to illustrate what I meant by mismatches. In order to justify adopting the SRL idea, the mismatches under discussion will need to have several properties which might not hold in the case of *nuclear physic-ist*. First, we should have a fairly clear idea about the semantics of the relevant elements so as to be convinced that the mismatch is real to begin with. Second, we should make sure that we cannot relate the morphological form with the semantic interpretation by means of movement or the relaxation of the semantic component without doing violence to the predictiveness of the system as a whole. Finally, it is not enough to show that licensor-free approaches face challenges; we should also show that SRL does better in accounting for the distributional facts.

As mentioned earlier, I believe that definiteness marking provides the right kind of evidence for the indirect association approach of SRL. We have a reasonably clear understanding of the semantics of definiteness, so we can easily convince ourselves that the mismatches are real. There are also cross-linguistically robust patterns of definiteness marking that involve definiteness-related markers that seem difficult to account for in terms of movement or of enriching the semantics. And a rather elementary implementation of the SRL idea seems to make the right predictions.

Definiteness, as expressed by the English definite article *the*, for example, can be described (informally) as a function that takes a set of individuals, denoted by a noun phrase, and returns the unique salient element in that domain, if such an element exists.[4] For example, a sentence like *The dog is happy* is true exactly in those conditions where there is a unique salient dog and that dog is happy. If there is a unique salient dog and that dog is not happy, the sentence is false. If there is no unique salient dog (either because there is no salient dog or because there are two or more such dogs), the sentence is certainly not judged to be true, though whether it is false or not has been a matter of considerable debate in the literature. The outcome of this debate makes little difference for the argument in this chapter, and I will assume, following Frege and many others, that when the uniqueness

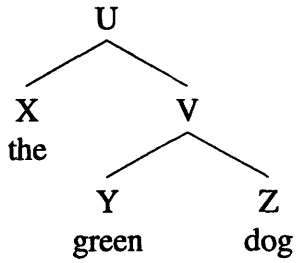
[4]To simplify the presentation, I will gloss here over issues concerning the interaction of definiteness and plurality. Following Sharvy (1980) and Link (1983), plurality is usually analyzed in terms of maximality, and the same could be done here.

conditions are not satisfied, the sentence is neither true nor false. Using the notation of Heim and Kratzer (1998), with ':' introducing a presupposition and C denoting the set of contextually salient individuals, the following entry is roughly what we need:

(113) $[\![\text{ the }]\!]^C = \lambda f_{et}$: there is exactly one $x \in C$ s.t. $f(x) = 1$. the unique $x \in C$ s.t. f(x)=1

When a noun is modified by an adjective, the domain of the definite article is the result of restricting the denotation of the noun with that of the adjective (again, subject to salience). *The green dog is happy*, for example, is true in contexts where there is exactly one salient dog that is green and where that dog is happy. It is false when that dog is not happy. And it is undefined (and in any case not true) if the set of contextually salient dogs has anything other than a unique member. Whether there are other salient dogs or green things in the context makes no difference: the only thing that counts is the set of green dogs. This means that, from the point of view of semantic interpretation, the noun and the adjective must combine together first, and that the definite article combines with the result. For a language like English, these semantic considerations lead to a bracketing that matches the bracketing suggested by syntactic considerations, so no compositionality issues arise:

(114) (Semantics~Syntax)



The semantic considerations we just discussed apply not only to adjectives but to any element that restricts the denotation of the noun. In particular, restrictive relative clauses (RRCs) and PPs should combine with the noun before the definite article does. It is not entirely straightforward to ascertain that RRCs and PPs in English are attached below the definite article, but a general bracketing like the following seems quite plausible:[5]

[5]The question of whether semantic compositionality should make us adopt this structure in syntax has been the center of the debate between Partee (1973) and Chomsky (1975). See Carlson (1983) and Heim

(115) [*def* [*AP* ... *N* ... *PP/RRC*]]

English, then, seems to suggest a compositional approach to definiteness marking, one in which the constraints imposed by semantic interpretation are fully satisfied by morpho-syntactic evidence. Other languages tell a different story.

(116) a. gris-en med blå pletter
pig-DEF with blue spots
'the pig with blue spots' (Danish, Hankamer and Mikkelsen, 2005, ex. 49)

b. skapar-en av skådespel-et
creator-DEF of play-the
'the creator of the play' (Swedish, Delsing, 1993, ex. 117)

c. to megalo to kokkino to vivlio
the big the red the book
'the big red book' (Greek, Alexiadou and Wilder, 1998a, ex. 39)

The definiteness marker appears between the noun and a modifier PP in (116a), between the noun and a complement PP in (116b), and multiple times, including one occurrence between an adjective and the noun and another between two adjectives in (116c). Assuming that the semantics of definiteness in Danish, Swedish, and Greek is the same as in English (and I am not aware of anything that would suggest otherwise), the examples in (116) are a clear compositionality puzzle.

My goal is to use the patterns that generate the apparent mismatches in (116) as evidence for SRL, and consequently also for the use of complexity-based competition in grammar. My plan, as hinted at above, will be to provide evidence showing that what we referred to above as definiteness markers are neither function heads with the denotation in (113) (that is, they are not *spreaders*), nor are they semantically vacuous elements expressing a definiteness feature in their immediate environment (that is, they are not *realizers*). Spreaders and realizers will both play a role in the patterns of definiteness marking, but I will try to show that the definiteness markers themselves are best analyzed in terms of *licensors*. If the arguments are correct, we will have evidence for the SRL approach, which in turn would support the notion of structural economy in grammar.

---

and Kratzer (1998, 82–3) for discussions both of the general problem and of the significance of definiteness marking in Scandinavian languages to the issue of compositionality.

I will introduce the details of SRL through the interaction of definiteness marking and adjectival modification in Danish, to which we now turn.

4.3 Definiteness in Danish

4.3.1 The basic pattern

Like most Scandinavian languages, Danish can mark definiteness either with a nominal suffix or with an independent, pre-nominal form.[6] The nominal suffix is *-en* in the common gender and *-et* in the neuter. The pre-nominal form is *den* in the common gender and *det* in the neuter.[7]

The choice of suffix or pre-nominal form is not free. At a first approximation, the generalization is this:[8]

(117) Basic generalization: the pre-nominal marker is required if the noun is modified by an adjective and is disallowed with unmodified nouns.

The following examples, from Hankamer and Mikkelsen (2005, pp. 87–88), illustrate the generalization:

[6]The pre-nominal form is often referred to as the definite article, which may suggest a treatment in terms of a spreader with the semantics of definiteness, as in (113). If SRL is right, however, we should also consider the possibility that the pre-nominal form is a licensor, and that it does not have the meaning of definiteness. Below I will claim that treating both forms of definiteness marking as licensors is better than analyzing one of them (or both) as spreaders.

[7]The distinction between the two genders, the common gender and the neuter, will play an important role later, but for now I will mostly ignore it and use the common gender wherever possible.

[8]For the moment, I will focus exclusively on adjectives in describing the Danish pattern of definiteness marking. We already saw, in (116a) above, that PPs behave differently than adjectives. As discussed at length by Hankamer and Mikkelsen (2005), PPs and RRCs interact with definiteness marking in rather subtle ways, raising challenges for many familiar accounts. We will get back to this point in section 4.3.4, where we will see that the current proposal generalizes straightforwardly to handle PPs and RRCs and correctly predicts their interaction. I should also mention that, as has often been observed, the pre-nominal forms *den* and *det* can be used not only for definiteness but also as demonstratives. In their demonstrative use, prosodically marked by stress, the following noun can be unmodified. Generalization (117), then, is about the use of *den* and *det* for definiteness and not as demonstratives.

(118) Unmodified *N*: obligatory suffix

a. hest-en
horse-DEF
'the horse'

b. * den hest
DEF horse

(119) Adjectival modification: obligatory pre-nominal marker

a. * gamle hest-en
old horse-DEF

b. den gamle hest
DEF old horse
'the old horse'

Most of the literature on this pattern has focused on the dependencies between the two definiteness markers, *-en* and *den*, and the rest of the phrase. This will be a good place to start our discussion of licensors, and I will get to that in a moment. Before that, though, I would like to point out that we can already see the less controversial function elements, which I referred to above as spreaders and realizers. Spreaders are the function elements that make a semantic contribution. In our case, the relevant spreader is the definiteness head, which takes the denotation of the whole noun phrase as its argument. We can imagine this spreader as a covert head that attaches as sister to the noun phrase.[9] The main thing that spreaders do, other than contribute to the semantics, is spread features on elements in their domain, which can then be expressed by the other kind of function element, which I referred to as *realizer* above. In our case we will talk about a definiteness feature F_{DEF}. A realizer for this feature is the final *-e* on the attributive adjective in (119). This ending changes the adjective 'old' from its base form, *gammel*, which is found in singular indefinite noun

[9]Alternatively, one can avoid the idea of a definiteness head and take a syncategorematic approach, as proposed by Carlson (1983), where the distribution of the definiteness-related elements within the noun phrase is related directly to the operation of interpreting the noun phrase as definite. If the current proposal is correct, though, the syncategorematic version will not determine the distribution of *-en* directly but through something like licensors. As far as I can see, the argument below can be stated equally well with covert heads and with syncategorematic entries, and I am not aware of other considerations that would support one approach over the other. For purposes of presentation I will stick to the terminology of covert heads.

phrases, to the form *gamle* that we saw. More generally, *-e* is added to most adjectives in Danish when appearing in definite or in plural noun phrases. When multiple adjectives modify the same noun, each will show up with its own *-e*:[10]

(120) a. en stor gammel hest
1-CG big-CG old-CG horse
'a big old horse'

b. den stor-e gaml-e hest
DEF-CG big old horse
'a big old horse'

What we have, then, is a spreader attaching where its semantic contribution is made and spreading its features below, and realizers expressing instances of this feature on the adjectives. This much, I take it, is shared by all accounts of the basic pattern, either explicitly or implicitly.[11] What remains controversial is the distribution of *-en* and *den*. Let us take a very quick look at what licensor-free accounts have to say about the distribution of these markers in (118) and (119) and then turn to our SRL account.

4.3.2 Sketches of licensor-free accounts

Licensor-free account I: the markers *-en* and *den* as spreaders

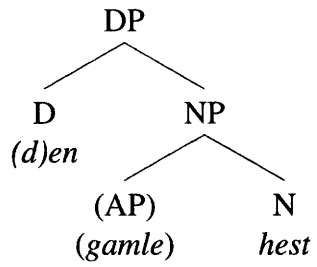
If definiteness markers in Danish are spreaders, the same considerations of compositionality that applied in our discussion of the English definite article in the introduction would lead us to something like the following:[12]

[10]I use 1 as a gloss for the common gender (CG) form of the indefinite article. This is also the form of the numeral 'one', orthographically written as *én*. I will have more to say about 1 when we discuss gender marking, in section 4.5 below.

[11]Minor details can change. For example, Börjars and Donohue (2000) treat the *-e* form as basic, analyzing what we referred to as the base form as the marked form expressing indefiniteness. For them, too, though, the relevant features are spread within the noun phrase. The difference is only in the particular features that are spread and in the choice of realizers.

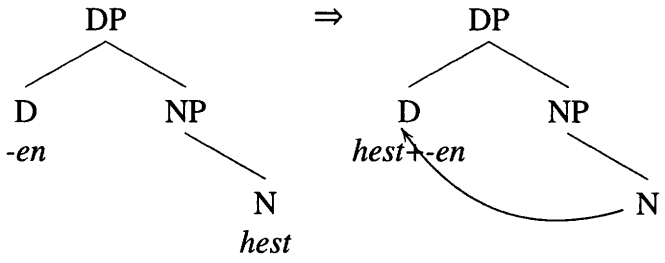
[12]The labeling of the non-terminal nodes can vary. Delsing (1993), for example, analyzed *NP* as a right specifier of the adjective. As far as the current compositionality considerations are concerned, this choice does not make a difference.

(121)

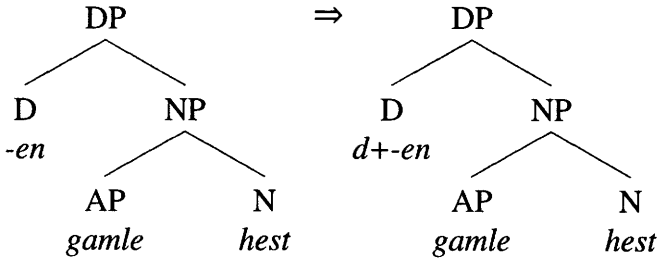


This structure corresponds directly to the surface form of noun phrases with adjectival modification, such as *den gamle hest.* It is also perfectly compatible with the unmodified form *hest-en*: (121) represents hierarchical relations; our compositionality considerations do not tell us whether *D* should appear to the left of *NP* or to its right, or whether *D* can form a word with *N*. However, to account for the dependency pattern in (118) and (119), most spreader accounts treat the linear order and affixation in *den gamle hest* (as accidentally drawn in (121)) as basic, attributing the post-nominal suffix configuration to a process that can only take place in the absence of an intervening adjective. This process can be the raising of *N* to *D*, as assumed by Delsing (1993) and Embick and Noyer (2001):

(122) a. Unmodified *N* raises to *D*



b. Intervening adjective blocks raising



Alternatively, as suggested by Embick and Marantz (2008), movement can go in the opposite direction, with *D* moving post-syntactically to *N* unless an adjective linearly intervenes.

As Hankamer and Mikkelsen (2005) discuss in detail, it is hard to use familiar kinds of movement as the basis for the dependencies between the markers in Danish. They also note that a movement account would encounter difficulties in accounting for certain lexical

exceptions, as well as in capturing the different effects on definiteness marking that various post-nominal modifiers have, an issue that we will come back to below. We will encounter further complications for a spreader account in what follows. However, while I agree with Hankamer and Mikkelsen that movement does not offer an adequate characterization of the dependencies under discussion, I believe that it does offer a key ingredient that should be preserved by any account. This ingredient is the identification of the nominal suffix *-en* with the pre-nominal form *den* (or with part of it). The two forms seem to be related, but capturing this relatedness can be a difficult task for certain approaches, such as the realizer accounts that we will shortly review. For a movement account, relatedness is expected: it is the exact same *D* that is sometimes attached to *N* and sometimes appearing in its base position, where it surfaces with an anchoring stem *d*. The nominal suffix is correctly predicted to be a subpart of the pre-nominal form.

Licensor-free account II: the markers *-en* and *den* as realizers

On a realizer account of the definiteness marker in Danish, *-en* and *den* are similar to the adjectival ending in that they all express a definiteness-related feature within the noun phrase. Since there is only one marker but possibly many adjectival realizers, something about the marker will have to be different. A familiar idea is that the adjectival endings express a feature that is related to *words* while the marker expresses a feature that is related to whole *phrases*. For example, Börjars and Donohue (2000) propose that each of the adjectives in an indefinite noun phrase, as in (120a), has a *word-level* null affix realizing an *word-level* indefiniteness feature, while the marker is a *phrasal affix* expressing a *phrasal* definiteness feature of the whole noun phrase. A similar idea is followed by Hankamer and Mikkelsen (2002), who treat *-en* as an affix that is attached by a lexical rule (their Rule D, p. 155) that turns a definite *N* into a *D*, effectively stating that *N* is a complete *NP* that cannot be further modified.[13]

To capture the generalization in (117), a natural idea for a realizer account of definiteness marking is that *-en* is the preferred form, used whenever possible, while *den* is the

[13]The other definiteness marker, *den*, is assumed by Hankamer and Mikkelsen to appear in *D*, as sister to a definite *NP*. It can be thought of as another realizer of the definiteness feature or as the spreader itself.

marked form, used only if *-en* is disallowed. This idea, which both Börjars and Donohue (2000) and Hankamer and Mikkelsen (2002) pursue, derives the badness of **den hest*, in (118) above, as the result of *blocking* by the preferred *hest-en*. Both accounts attribute the markedness of *den* with respect to *-en* to a preference for words to phrases. For Hankamer and Mikkelsen (2002), the preference is stated directly, following Poser (1992)'s proposal. For Börjars and Donohue 2000, p. 335, the preference is derived from an economy condition: *den* adds *syntactic* structure, leading to a violation of economy that is more severe than the violation that is caused by *-en*, which only adds *morphological* structure.

There are some aspects of the realizers approach that I think are correct. For example, the idea that the definiteness marker is not itself a spreader but rather an indirect reflection of definiteness, related to the features within the noun phrase. I also think that realizer accounts are right to distinguish between the local kind of realizers that we saw on the adjectives in (120) and the much less local behavior of markers. Finally, I agree with the realizer accounts on the significance of some structural markedness condition to the distribution of markers.

On the other hand, the realizers approach also gives rise to certain concerns, even if we restrict ourselves to the basic pattern above. First, the supposed distinction between words and phrases is suspect, and the very notion of a lexicon as a meaningful term has been shown to face nontrivial challenges. It would be good to avoid relying on a distinction between words and phrases as a basis for specifying features, affixes, and economy conditions.[14]

In addition to the inherent difficulty of basing an account on a distinction between words and phrases, there are specific generalizations about the Danish pattern that such an account obscures. For example, the form *den* properly contains the form *-en*, and it would seem reasonable to try to analyze *den* as bimorphemic: *d* + *-en*. This decomposition is easily accomplished within a spreaders approach, as we saw above. Within a realizer account, on

[14]Of course, it would be even better if we could reverse the dependencies assumed by the realizer accounts and actually *derive* the intuitive notion of word from independently needed primitives. I will make a preliminary and very tentative proposal in this direction in section 4.4.2 below, where I will suggest ways in which a syntactic notion of word might help us as a diagnostic of certain structural relations.

the other hand, decomposition would mean that *-en* can sometimes express a phrasal feature (as in *d-en gamle hest*), raising the question of why it was not able to do so when attached to the noun (**gamle hest-en*). Similarly, the proper containment of *-en* in *den* suggests an economy condition that is simpler than the ones used in the proposals discussed above. If *den=d+-en*, we could appeal to a uniform condition that penalizes structure rather than appeal to distinct evaluations of morphologic complexity and syntactic complexity. Again, this would not be available to an account that relies on *-en* being structurally different from *den*.

A related concern is the implementation of the two kinds of definiteness (or indefiniteness) features used to distinguish the unique occurrence of a definiteness marker within the phrase from the variable number of adjectival affixes. The proposals mentioned above do not provide any details about what they had in mind, and it seems to me that doing so would require complicating the machinery in various ways. In any case, I think it would be good if the account of *-en* and *den* and the account of adjectival endings would use the same features and the same spreading operation.

4.3.3 Licensing

The SRL idea makes use of a two-step architecture. The first step, which I will refer to as GRAMMAR, generates structures and enforces the usual well-formedness conditions on phrase structure, selection, phonology, semantics, and so on. For example, the following aspects of the distribution of *-en* will be managed by GRAMMAR:[15]

[15]I am assuming that *den* is by morphemic and that its decomposition is *den = d + en*. I don't know what *d* is, and I will treat it as a dummy element for purposes of exposition. More significantly, as (123b) makes clear, I am assuming that the *-en* in *den* is the same element as the nominal suffix *-en*. This morphological identity seems straightforward for the singular, common gender form, but it will get obscured once we move to the neuter gender, and even more so in the plural, where, as pointed out to me by Line Mikkelsen and others, the pre-nominal form is *de* and the post-nominal one is *-ne*. However, it seems to me that the idea of decomposing *de* into *d* and *-ne* is not entirely far-fetched, especially given the absence of *dn* as an onset cluster in Danish (see Basbøll, 2005 p. 206). Decomposition will be the least obvious for Icelandic, which we will study in detail in section 4.6. Interestingly, once we peel off a few phonological processes that are needed for an account of Icelandic independently of our account, the form identity predicted by the current

(123) a. *-en* is a suffix

b. *-en* can attach to *N* and *d*, but not to *A*

GRAMMAR also takes care of spreading, realizing, and licensing. For the moment, we will focus on the following conditions:

(124) Spreading: the head noun in a noun phrase and all modifying adjectives have the feature [F_{DEF}] *iff* the noun phrase is interpreted as definite

(125) Realizing: F_{DEF} is realized as *-e* on all modifying adjectives (a different instance of *-e* for each instance of F_{DEF})

(126) Licensing: each instance of F_{DEF} is c-commanded by some instance of *-en* (possibly one *-en* for several instances of F_{DEF})

In everything we have in GRAMMAR so far, only (126) is special to SRL. The rest, as mentioned above, is shared in one way or another by all current proposals that I am aware of.

The second step of the SRL architecture is ECONOMY, where outputs of GRAMMAR are evaluated, and those that have unnecessarily many licensors (and maybe other dummy elements) are ruled out.[16]

(127) ECONOMY: If S_1 and S_2 are identical except for licensors, and if S_1 has strictly fewer licensors than S_2, then S_2 is ungrammatical

Here is how it would work for the basic pattern above. In (118) and (119) the noun phrase is definite, so *N* and all modifying *A* (when present) bear F_{DEF}. *-en* licenses the single instance of F_{DEF} in (118a), *hest-en*, and since there is no well-formed candidate that satisfies ECONOMY better, (118a) is grammatical.[17] Note that the bare form *hest* is more economical than *hest-en*, but it is ill-formed due to its unlicensed instance of F_{DEF}, and consequently it is not part of the candidate set evaluated by ECONOMY.

---

proposal will turn out to be complete.

[16]This two-step evaluation, where some operations can take place freely but are subject to an economy condition, bears obvious resemblance to the framework of Fox (2000).

[17]I am currently assuming that spreading domains are the same as licensing domains. This makes sense here, but one could imagine things being different.

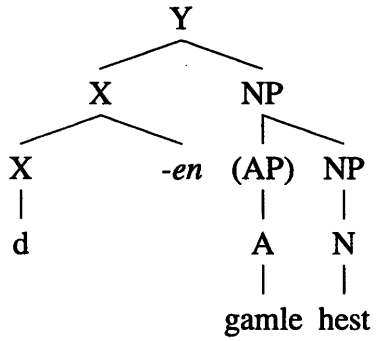
There are two possible explanations for why (118b), **den hest*, is ungrammatical. If *d* is a dummy element (another licensor?), the structure will be ruled out by ECONOMY: (118a), *hest-en*, is identical modulo dummy elements and is more economical. Or, if we can find a more meaningful role for *d* in the way modification is implemented, (118b) could be ruled out by GRAMMAR.

As for (119a), **gamle hest-en*, I would like to say that it is ruled out by GRAMMAR: *-en* is attached too low to c-command the adjective, leaving an instance of F_{DEF} unlicensed. (119b), *den gamle hest*, on the other hand, is grammatical: all the instances of F_{DEF} are licensed (details soon), and there is no well-formed candidate that is more economical. By ECONOMY (and the well-formedness of (119b)) we correctly predict that double-definiteness should be ungrammatical:[18]

(128) * den gamle hest-en
DEF old horse-DEF

In order to actually derive the asymmetry between the bad (119a) and the good (119b) we should be more particular about our structural relations. What we need is to ensure that *-en* c-commands both *N* and *A* when attached to *d* but not when attached to *N*:

(129) Good: *-en* on *d* c-commands both *A* and *N*



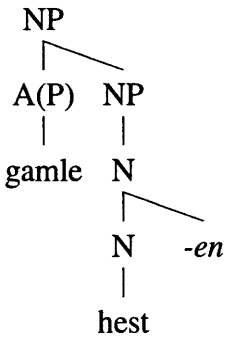
If c-command is defined in terms of first (branching) node up, (129) will not have the desired c-command relations. We need a looser notion of c-command, ensuring that something like the following holds:

(130) An affix c-commands everything its attachment site does

[18]Many other candidates should be considered. **en-hest*, for example, and **gaml-en hest-en*, and so on. I assume that they are all ruled out by GRAMMAR.

In section 4.4.2 below I will try to show that we can derive condition (130) if we follow May, 1985 and Kayne, 1994 in using a notion of c-command that is sensitive to the distinction between categories and segments (Chomsky, 1986). I will also try to show that this notion will give us a handle on the wordhood, which in turn will restrict our syntactic possibilities and give rise to some new predictions. For now, though, I will treat (130) as an arbitrary stipulation and simply assume that it holds. If (130) is guaranteed, we obtain the licensing of F_{DEF} on the adjective and the noun by the definite article. For the bad case, N must be too low for a sister of $A(P)$

(131) Bad: post-nominal *-en* does not c-command the adjective



The dependence of our story on the particular way that $A(P)$ attaches within the noun phrase should be alarming. We have to make what looks like a completely arbitrary choice with respect to phrase-structure. I believe, though, that this is actually a fortunate state of affairs, and that treating this choice as arbitrary allows us to account for the superficially different distributional patterns of marking in Icelandic in terms of a variant of Danish in which this choice was made differently. I will discuss that in detail in section 4.6.

4.3.4 Other modifiers

Prepositional phrases do not trigger the pre-nominal definiteness marker. In the absence of other modifiers within the definite noun phrase, only the nominal suffix is possible, and the PP appears after the suffixed noun:[19]

[19]We already saw some of the relevant facts in the introduction when we talked about compositionality puzzles.

(132) PP in a definite noun phrase with no other modifiers (Hankamer and Mikkelsen, 2005:111)

a. gris-en med blå pletter
pig-DEF with blue spots
'the pig with blue spots'

b. * den gris med blå pletter
DEF pig with blue spots

The current account has little of interest to say about (132). The spreading rule (124) said that F_{DEF} spreads onto the head noun and all modifying adjectives. Without further modification, nothing within a PP will get F_{DEF} from the modified noun. This means that nothing within the PP would need licensing from the outside. In (132), only the noun will have F_{DEF} related to the definiteness of the whole noun phrase, and as before, this F_{DEF} can be licensed using nothing more than the post-nominal *-en*. More generally, we would expect that, all things being equal, PPs would be inert with respect to definiteness marking: the definiteness marking of a noun phrase with a PP should be exactly the same as that of a noun phrase without it. For example, we expect, correctly, that adding an adjective should trigger the use of the pre-nominal marker:

(133) * gamle gris-en med blå pletter
old pig-DEF with blue spots

(134) den gamle gris med blå pletter
DEF old pig with blue spots

Non-restrictive relative clauses behave like PPs in not triggering the pre-nominal marker. For them, too, nothing has to be changed in the current proposal. Restrictive relative clauses (RRCs), on the other hand, are different from PPs (and from non-restrictive relatives clauses) and similar to adjectives in triggering the pre-nominal marker, at least optionally:

(135) (Examples from Hankamer and Mikkelsen, 2005:108)

a. hest-en som vandt løb-et
horse-DEF that won race-DEF
'the horse, which won the race' (*non-restrictive, all speakers*)

'the horse that won the race' (*restrictive, some speakers*)

b. den hest som vandt løb-et
DEF horse that won race-DEF

'the horse that won the race' (*restrictive, all speakers who get the definite DP reading*)

To account for definiteness marking with RRCs, then, we need to update our definition (124). We could say, for example, that F_{DEF} spreads onto the relative pronoun in an RRC. Here is the revised spreading rule:

(136) Spreading (revised): the head noun in a noun phrase, all modifying adjectives, and the relative pronoun in a restrictive relative clause have the feature [F_{DEF}] *iff* the noun phrase is interpreted as definite

Other than this minor modification, nothing more needs to be said about relative clauses to get their interaction with definiteness. For example, an adjective should trigger the pre-nominal marker, regardless of whether the relative clause is restrictive or not. And since, on our account, PPs are inert and RRCs trigger the pre-nominal marker, we predict that a definite noun phrase with a PP and an RRC (and no other modifier) should have a pre-nominal marker. This is correct:

(137) den gris med blå pletter som vi fik af nabo-en
DEF pig with blue spots that we got from neighbor-DEF
'the pig with blue spots that we got from the neighbor' (Hankamer and Mikkelsen, 2005:112)

From the perspective of SRL, all these facts are unremarkable. Our earlier system, designed to account for the basic pattern, was able to accommodate these new interactions straightforwardly. The only reason I have mentioned PPs and RRCs at all is that they have posed problems for all the other major approaches that I am aware of. It is not difficult to see why. For a realizer account, the choice between *-en* and *den* boils down to a distinction between words and phrases: *-en* goes on words; *den* goes on phrases. RRCs seem well-behaved in this respect. Like adjectives, they make the noun phrase too big to fit into a single word, so *den* is the marker of choice. PPs, on the other hand, are mysterious. I take it that modifying a noun with a PP is no more word-like than modifying it with an adjective, and yet a PP appears with *-en* and not *den* (in the absence of other modifiers). Could one say that PPs attach so high that they do not really belong to the noun phrase? Hankamer

and Mikkelsen (2005) suggest this, but the position of the PP between the noun and the RRC in (137) makes this proposal look unappealing.[20]

As for a spreader account, PPs and RRCs pull such an account in two opposite directions. In terms of structure, the attachment of a modifier PP is presumably not all that different from that of an attributive adjective. The inertness of PPs to definiteness, then, can be seen as further evidence against an analysis in terms of structural intervention, where *-en* on *D* and *N* below meet through movement unless there is more structure in the middle. In the space of currently available spreader accounts, this amounts to further evidence for an analysis in terms of linear intervention, like that of Embick and Marantz (2008). This makes it all the more surprising that a post-nominal RRC can also trigger *den*.[21]

4.3.5 Interim summary

We saw how introducing the notion of licensors allowed us to account for the basic pattern of definiteness marking in Danish. Stipulating a new kind of function element is hardly a pleasing move, but we noted that previous attempts to account for the pattern in a more conservative way have ended up making considerably complex stipulations of their own even for the simple case of a single noun and a single adjective. Moving on to other elements within the noun phrase, we observed that our SRL account extended naturally to capture the interactions of those elements and their combinations with definiteness, while licensor-free accounts had a harder time. We will soon see another advantage of using licensors. The same mechanism that accounted for the definiteness pattern will allow us to account for a seemingly very different pattern concerning gender marking, as well as for the interaction of the two patterns. Additional support for the current proposal will come from examining marking patterns in other languages, with particular focus on Icelandic and

[20]Hankamer and Mikkelsen try to account for these facts in terms of a raising analysis for relative clauses (as well as various assumptions about definiteness and phonological content of the elements participating in this construction). They do not, however, provide any independent evidence for this analysis, and as far as I have been able to establish, the relevant cases pattern with what has been analyzed as matching relative clauses and not with raising ones. See Hulsey and Sauerland (2006).

[21]Embick and Marantz refer to Hankamer and Mikkelsen (2005)'s proposal, mentioned above, that what we have here is a raising relative clause.

Greek.

At this point it might be useful to remind ourselves how the current discussion relates to the general claims of this dissertation. The connection is through structural economy. The component that we referred to as GRAMMAR constrained the placement of licensors within the structure, but often there were still several structures that were all well-formed and differed only with respect to how many times a licensor occurred within a structure and where those occurrences were attached. In those cases, ECONOMY ruled out all those structures that did with more what could have been done with less.

The significant point here is that this final evaluation is based on structural competition alone. All the candidates for ECONOMY are syntactically well-formed and have the same meaning. The only criterion that distinguishes between them is structural complexity, and simpler counts as better.

4.4 Danish definiteness marking: some remaining issues

4.4.1 Evaluating candidates

The account outlined above involves structural competition. As mentioned in chapter 1, however, there are good reasons, discussed by Poser (1992) and Embick and Marantz (2008), to be wary of such forms of evaluation. While I find the arguments provided by these authors convincing, I believe that the structural competition in our SRL account is not subject to their objections. For example, Poser notes that a comparison that is based on phonetic effort, measured in terms of overt phonological material, would fail to account for the shorter *child-s* being blocked by *children*. In our account, phonetic effort played no role – our measures here and elsewhere in this thesis are syntactic.

We also mentioned two further problems that Poser points out for a pragmatic approach. First, the judgments for supposed violations of Grice's Maxim of Manner seem weaker than the intuitively much clearer grammaticality contrasts in blocking effects: *??pale red* might be blocked by *pink*, but its oddness is milder than the badness of **man-s*. Second, Poser observes that the evaluation for blocking seems to be subject to both syntactic and semantic

restrictions: *John is smarter than Tom* may perhaps block **John is more smart than Tom* but it clearly does not block *John has greater intelligence than Tom.*

Starting with the syntactic restriction, our account compares candidates that are almost entirely identical, differing only in the number of occurrences of licensors within the relevant spreading and licensing domains. This allows us to avoid the issue of form-relatedness. In addition, the domains in which the candidates can differ are quite small. The relevant domain for *-en*, for example, seems to be bounded by the spreader above and the head noun below, and other than these two elements it includes only modifiers and possibly complements of the head noun (of which we only discussed adjectives so far).[22] Material that is higher up or further embedded does not play a role in the distribution, and we do not make reference to it in the evaluation. Our account, then, is quite conservative with respect to locality and to form-relatedness.

As for the semantic computations involved, the elements that make the main semantic contribution are identical for all the candidates. The difference is in elements that are either entirely devoid of meaning or with a meaning that is impoverished enough so that no occurrence beyond the first one will make a semantic contribution. This makes it possible for economy to know that if the candidates differ only in licensors, they all have the same meaning. Competition, then, can take place before any semantic computation has taken place, and certainly before computations that refer to lexical material and world knowledge have been performed.[23]

We can now return to the difference between pragmatic oddness and strict ungrammaticality. As we just saw, the current approach makes reference to competition that evaluates candidates that are near-identical in form, differing only in the distribution of occurrences of one kind of element within a syntactically circumscribed domain. Ensuring that the candidates have the same meaning requires very little semantics and no world-knowledge. It seems quite reasonable to me that this evaluation takes place within the grammar, before

[22]This roughly coincides with the domain of spreading as diagnosed by realizers like *-e*.

[23]This fits in with a modular architecture of the kind argued for in Fox (2000), where certain semantically-sensitive operations are isolated within a module that has no access to full semantic and pragmatic knowledge. See Gajewski (2002); Fox and Hackl (2006); Magri (2006); Abrusán (2007) for further arguments for such an architecture.

general pragmatic reasoning is invoked, opening the way for a source of badness for **d-en gamle hest-en* that was not available for cases like *pale red.*

4.4.2 Licensing, c-command, and wordhood

We noted above that we would need a particular notion of c-command to make licensing do what we want. We stated this, in (130), as a condition saying that affixes c-command whatever their hosts do. Is it possible to derive this condition rather than stipulate it?

First attempt: c-command through wordhood

Considerations of wordhood suggest an approach we might explore for deriving (130). We will look at this approach very briefly and then reject it in favor of a more syntactic alternative.

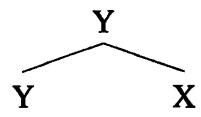
Affixes combine with their attachment sites to form one word. Perhaps we could say, then, that each element inside a word c-commands everything that the containing word c-commands. For an affix that does not take scope over another element (in the sense relevant for syntax above the word level), wordhood will be provide a special way to be considered as c-commanding that other element. This would build an asymmetry between scope relations above and below the word level, a direction that might make sense within a framework that assigns a special status to the notion of word.[24] From the perspective of our framework so far, however, the opposite direction seems more promising. We have been developing a syntactic account of affixation. It would seem more natural to look for a way to relate (130) to other syntactic phenomena, and to try to obtain the notion of wordhood as a derived notion.

[24]A possible concern is that in such frameworks, word-parts do not usually interact with sentence-level syntax, making the notion of c-command by word-parts for the purposes of licensing an exception.

Second attempt: c-command and labels

A definition of c-command that derives (130) was already provided and used for a variety of purposes by May (1985) and Kayne (1994). This definition makes use of a distinction between *categories* and *segments* (cf. Chomsky, 1986). A category is a maximal set of contiguous nodes that have the same label. The individual nodes in a category are its segments. In (138), for example, there are two categories. One is composed of the single segment labeled *X*, while the other is composed of the two segments labeled *Y*.

(138)



Following May (1985) and Kayne (1994) we will assume that the notion of dominance that is relevant for c-command is sensitive to the distinction between category and segment: a category *X* will be said to dominate a node *n iff* every segment of *X* dominates *n*.[25] In (138), for example, the category *Y* does not dominate *X*, since only the higher segment of *Y* dominates *X*, while the lower segment of *Y* does not. The definition of c-command will now look like this:

(139) *X c-commands Y iff*

a. *X* and *Y* are categories

b. $X \neq Y$

c. Every category that dominates *X* dominates *Y*

A consequence of definition (139) is that adjuncts, such as our definiteness suffix, now c-command outside of their attachment sites. A simple adjunction structure has the schematic form of (138) above. As we just saw, the category *Y* does not dominate *X* (only one segment of *Y* does). Consequently, every category that dominates *X* also dominates *Y*.[26] If (138) is embedded inside a bigger structure, definition (139) guarantees that *X* will c-command everything inside the sister of the highest segment of *Y*. In other words, *X* will

[25]Segments are individual nodes, and for them I assume the usual definition of dominance as the transitive closure of the *mother-of* relation among nodes in the tree.

[26]This is not necessarily true if multiple-dominance is allowed, a possibility that I will ignore here.

c-command everything that its attachment site does, which is precisely what is needed for (130).

Wordhood within a syntactic framework

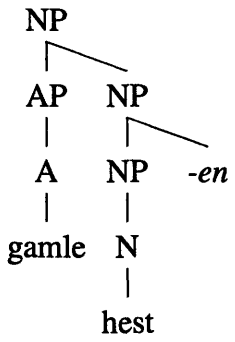
With the help of definition (139) we now have a reasonably principled way to derive the scopal relations of those elements within the Danish noun phrase that we have looked at. Moreover, our account is uniformly syntactic: all the relevant relations are defined in terms of dominance and labeling within the tree. In particular, no reference is made to the notion of word. Of course, some things are words and others are not, at least as far as speaker intuitions and certain phonological processes are concerned. For example, *-en* combines with the element to its left to form a single word, while the noun and the adjective do not. One could say, perhaps, that *-en* is specified as an affix, and that this forces it to surface as part of a bigger word more or less independently of its syntactic configuration. This would hardly be a predictive theory, though. Fortunately, we can do better. Using the definition of c-command in (139), we can define a derived notion of wordhood as follows:

(140) Two different *terminal* categories u and v will be said to belong to the same *word* *iff* (a) u c-commands v and (b) v c-commands u

When *-en* adjoins to *hest*, for example, the first category dominating either is *NP*. The affix and the noun, then, c-command each other and therefore belong to the same word. When *-en* adjoins to *d* the first category that dominates either is *DP*, and again we predict the two to be part of the same word. On the other hand, the first category dominating the adjective *gamle* is *AP*, so it does not c-command the noun, the affix, or *d*, and we correctly predict that the adjective will form an independent word.

Definition (140) allows us to impose further constraints on phrase structure. Consider the following structure (the relevance of this example was pointed out to me by Sabine Iatridou, p.c.):

(141) Bad:



Like (131), the structure in (141) is an analysis of the ungrammatical *gamle hest-en.* As opposed to (131), however, our system would predict (141) to satisfy all the marking constraints: *-en* is adjoined to *NP* rather than dominated by it, and every category that dominates *-en* also dominates the adjective *gamle.* Until now we had no independent reason to rule out (141) as a possible analysis for *gamle hest-en.* Wordhood provides a reason to reject this structure. By (123a), the definiteness marker *-en* is a suffix. This means that *-en* must form part of the same word as the element that precedes it, in this case the noun *hest.* Consequently, using (140), the two have to c-command each other. In (141), however, the first category dominating *hest* is *NP*, which does not c-command *-en*, ruling out the structure.[27]

Using definition (139) we have derived the ability of affixes to c-command everything their attachment site does, as required by our account of licensing. The same definition also provided us with a derived notion of wordhood that accounts for the mapping from those structures that we saw into words and provides some further constraints on possible phrase structures. We can now return to our affixation patterns and examine the pattern of gender marking and its interaction with definiteness.

[27]Note that the consideration above applies only for the suffixal use of *en*, which seems to correspond to its use as a definiteness marker. As we have already seen, Danish uses *en* for purposes other than definiteness marking, and at least some of those uses involve an unbound *en*. As far as I can tell, the free occurrences of *en* are all pre-nominal. Since we have (141) available for a post-nominal, non-affixal *en*, we would need a separate explanation for why this parse is never used.

4.5 Gender in Danish

Our examples have so far used only the Common Gender (CG). As mentioned above, Danish has a second gender, the Neuter (Nt). The distinction between the two genders is expressed in several different places within the noun phrase.[28] Orthographically, the usual way to mark the distinction is by ending the Nt form in *t* instead of the final *n* (or ∅) in the CG form:[29]

(142) a. Definiteness suffix: *-en* for CG and *-et* for Nt

b. Pre-nominal definiteness marker: *den* for CG and *det* for Nt

c. Indefinite article: *en* for CG and *et* for Nt

d. 1: *én* for CG and *ét* for Nt

e. Adjective: ∅ for CG and *-t* for Nt

For now I will simply assume that *-t* is a marker of Nt, as suggested by (142e), and that the *et* forms in the other cases listed in (142) are the result of decomposition: *et* = *en* + *-t*. Earlier we mentioned that the pre-nominal CG definiteness marker *den* is already decomposable into *d* + *-en*; so for the Nt marker we will now have *det* = *d* + *-et* = *d* + *-en* + *-t*.

The following paradigm suggests that *-t* has the distributional properties of a licensor.

(143) Indefinite: article and adjectives both marked for gender

a. en stor gammel hest
1-CG big-CG old-CG horse
'a big old horse'

b. et stor-t gammel-t hus
1-Nt big-Nt old-Nt house

---

[28]But only in the singular. The distinction is neutralized in plural noun phrases.

[29]Phonetically, all the *-t* endings are not alike, as has been pointed out to me on several occasions. I believe that these differences between the different surface forms make sense once certain facts about Danish phonology are taken into account, and that the orthography is quite faithful to the underlying morphological forms. I will not try to go into detail in this case, but a similar issue on a larger scale will soon emerge as a challenge to our account of Icelandic in section (4.6), and for that language I will try to show that a set of indendently supported phonological processes account for all the apparent counterexamples.

'a big old house'

(144) Definite: contrast neutralized on adjectives (but remains on the article)

a. den store gamle hest
DEF-CG big old horse
'a big old horse'

b. det store gamle hus
DEF-Nt big old house
'a big old house'

The Nt marker *-t* appears on every adjective in an indefinite Nt noun phrase but only on the definiteness marker *-en* within a definite Nt noun phrase. This seems mysterious if *-t* is a spreader. The pattern is no less puzzling if *-t* is a realizer, since whatever gender/number features spread onto the adjectives is presumably the same in the definite and the indefinite form.[30] If *-t* is a licensor, on the other hand, things are simpler. What we would say is that there is some feature, call it F_{Nt}, which, judging from the places in which *-t* appears in (143b), spreads onto the adjectives and *en* when the noun phrase is Nt.

(145) *en* and all modifying *A* have [F_{Nt}] *iff* the noun is neuter singular

And just as *-en* licensed F_{DEF}, so does *-t* license F_{Nt}, and we can have either licensor on its own or stack *-t* on top of *-en* (though there is no stacking of *-en* on top of *-t*):[31]

(146)

	–	*def*
–	∅	*-en*
Neut	*-t*	*-en+-t*

As for the pre-nominal definiteness marker, we have already decomposed *den* into *d* + *-en*, so for the Nt marker we will have the decomposition *det* = *d* + *-et* = *d* + *-en* + *-t*.

[30]Börjars and Donohue (2000) offer an interesting proposal for dealing with the gender pattern within a realizers framework. They suggest that the *-t* form encodes indefiniteness, so that forms ending in *-t* are incompatible with a definite noun phrase, resulting in the emergence of the less fully specified *-e* form. This makes the appearance of *-t* on the definite article itself somewhat surprising. It also raises questions regarding the appearance of *-t* on predicative adjectives that are predicated on definite subjects: *hus-et er stort* 'the house is big'. Hankamer and Mikkelsen (2002, 2005) avoid the issue of gender marking altogether, and it is not obvious to me how the distributional facts can be made to fit with their account.

[31]Note that neither indefiniteness nor CG get marked. F_{DEF} and F_{Nt} appear to be privative features. I do not think this is a complete accident, but I will not pursue this matter here.

We also said that *-en* would need to c-command the adjectives from its attachment point on *d* in the pre-nominal marker. We now predict that when *-t* attaches on top of *d* + *-en*, c-commanding *-en*, it will also c-command all the adjectives below. This will allow the *-t* on the pre-nominal marker to license every instance of F_{Nt} on the adjectives below, just as *-en* on the pre-nominal marker licenses every instance of F_{DEF} on those adjectives (as well as on the noun). Consequently, any additional *-t* would be unnecessary, and due to ECONOMY only the highest *-t*, the one on the pre-nominal definiteness marker, will survive.

(147) a. det store hvide hus
DEF-t big white house
'the big, white house'

b. * det store hvid-t hus
DEF-t big white-t house

c. * det stor-t hvide hus
DEF-t big-t white house

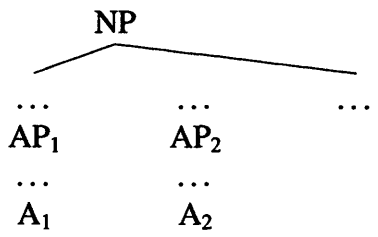
d. * den stor-t hus
DEF big-t house

In other words, if we are right about our treatment of *-t* as a licensor, we are not surprised to see all but one of the adjectival instances of *-t* obligatorily disappearing when preceded by the pre-nominal definiteness marker, which in turn surfaces with a final *-t*.

Turning our attention to the marking of Nt in the indefinite, as in (143b) above, we notice that each adjective bears its own *-t* marker. On our earlier assumptions, this can only mean that *-t* on one adjective does not c-command the other adjective. (If it did, one *-t* would be enough, and the second occurrence of *-t* would be ruled out by ECONOMY.) This, in turn, means that the attachment site of *-t* within one adjective does not c-command the other adjective. We have already seen one case where a licensor attaches too low to c-command anything outside of its host, and we have also seen a case where it attached high enough to c-command more material. The former case was the attachment of *-en* to the noun, and the latter was its attachment to *d*. It seems, then, that when *-t* attaches to the adjective the configuration is in some sense closer to the attachment of *-en* to the noun than to its attachment to *d*. This would make sense if, as has often been suggested, the Danish adjective is not a head that takes the noun phrase as its sister but part of a phrase

that attaches as an adjunct within the noun phrase.[32] The change of label from *A* to *AP* will ensure that a *-t* on *A* will only c-command elements that are lower than *AP*. In particular, it will not c-command adjectives other than its host. Any other adjective will require its own *-t* for licensing of F_{Nt}:

(148)

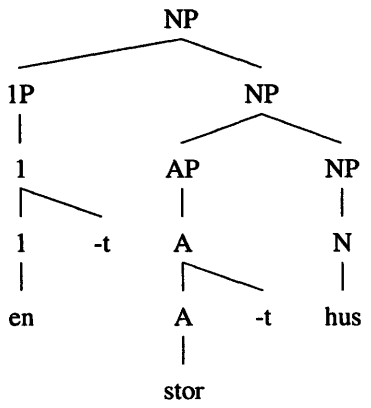


The appearance of *-t* also on the indefinite article 1 (*et*) suggests that 1, too, does not c-command the adjectives (and that the adjectives do not c-command 1). This means that the structural position of 1 is closer to that of adjectives than to the pre-nominal definiteness marker *den* (or *det*). Here is a possible structure where 1 attaches as an adjective:[33]

[32]See Svenonius (1994), Julien (2002), and Hankamer and Mikkelsen (2005) for discussion of the main considerations, all of which seem quite independent of our current concerns. To the evidence from the literature we can also add considerations of wordhood, as presented in section 4.4.2 above: if adjectives could c-command each other, we would expect them to cluster together to form a single word. Such clustering occurs within compounds, which are widespread in Germanic languages, including Danish, but crucially it never applies to a series of attributive adjectives.

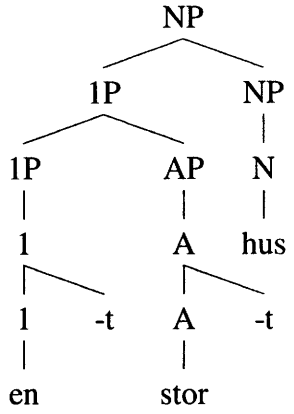
[33]This is not the only structure that allows 1P to adjoin to the noun phrase at a lower position than *den*. For example, we could have tried to use the following structure instead:

1. *AP* adjoined to *NP*



For our immediate purposes, I believe that the particular choice does not matter.

(149) A possible structure (there are others): *AP* adjoined to 1*P*



This point of treating 1 as an adjective is somewhat more controversial than the analysis of adjectives as phrasal, but I think it can justified independently of our proposal. The evidence comes from the ability of the definite article and of 1 to appear together in the same noun phrase. This happens, as has been pointed out to me by Torben Thrane, (p.c.), in certain contexts that correspond to a partitive reading for the noun phrase:[34]

(150) Den ene kop er forsvundet
DEF 1 cup is disappeared
'One of the cups has disappeared'

Significantly for our discussion of phrase structure, 1 in these cases appears between the definite article and the noun. Moreover, the 1 surfaces with what we referred to above as the realizer *-e*, just like an adjective under a definite article. This supports our choice of treating 1 as something similar to an adjective rather than as a counterpart of *den*. To summarize the last few steps, our earlier assumptions about licensing have pushed us in a particular direction with respect to phrase structure. This direction, in turn, allowed us to discover an interesting fact, for which we found independent support, about the position of so-called definite and indefinite articles within the Danish noun phrase.

And one final step. If the indefinite Nt article *et* is really an adjectival 1 with a licensor *-t*, and if this article can appear below the definiteness marker, we should expect that the *-t* on 1 would obligatorily disappear in such contexts, just as it disappeared on adjectives, and that instead of it we would find the realizer *-e*, again, in analogy with adjectives. This is correct:

[34]I will ignore here the precise meaning of this construction and how it may arise compositionally.

(151) a. det ene hvide krus er forsvundet
DEF 1 white mug is disappeared
'One of the white mugs has disappeared'

b. * det ete/et hvide krus er forsvundet
DEF 1 white mug is disappeared

c. * det ene hvidt krus er forsvundet
DEF 1 white mug is disappeared

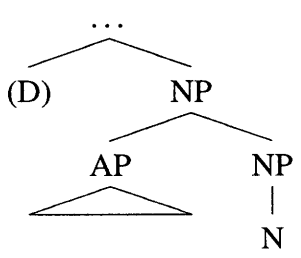
4.6 Icelandic

4.6.1 Changing the lower segment in the Danish *NP*

In the previous sections we examined two distributional patterns within the Danish noun phrase. One pattern had to do with the marking of definiteness and involved *-en* hopping; the other pattern had to do with the marking of the the neuter singular and involved *-t* spreading. We noted that analyses that treat these markers as spreaders or realizers face a variety of challenges, and we saw how introducing a third kind of function element, which we referred to as licensors, allowed us to account for these patterns and their interactions. It also led us to some new facts about the attachment site and the internal structure of the definiteness marker 1. Indirectly, we developed an argument for structural competition in which simpler counts as better.

Enriching our ontology of function elements is not something to celebrate. It seems to me, though, that this is a reasonable price to pay for the ability to capture the marking patterns that we saw. As for our other assumptions, I tried to show that, with one exception, they all seem plausible enough quite independently of the current proposal. The exception had to do with the attachment of *AP* as an adjunct to *NP*:

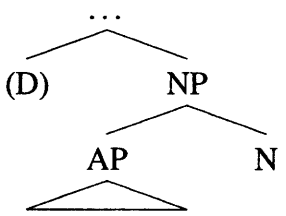
(152)



Nothing about the attachment of *AP* in (152) is particularly implausible (and in fact a

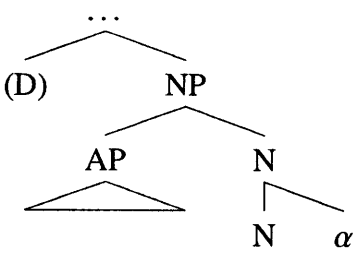
similar structure has been suggested by Svenonius, 1993, though for different reasons and within a different theoretical setting). However, our only motivation for having this attachment site came from our need to prevent the post-nominal definiteness marker from c-commanding into the adjective. I can think of no general reason to expect this kind of configuration and must therefore treat this as an arbitrary choice of Danish. But if it is an arbitrary choice, we should expect to find a variant of Danish, call it Danish′, where a different choice has been made. I believe that Icelandic is precisely this kind of Danish′, and that where Danish has *NP* as the sister of *AP*, Icelandic has *N*:

(153) Danish′ (= Icelandic)



As far as licensing is concerned, the main difference between the structure for Danish, in (152), and the structure for Danish′, in (153), is that in the former *N* does not c-command the adjective and the latter it does. This means that if an affix α is adjoined to *N* in Danish′, as *-en/et* did in Danish, it would c-command the adjective and license occurrences of F_α on it, contrasting with its inability to do so in Danish. For purposes of licensing features on adjectives, then, a post-nominal attachment site would suffice:

(154)



And as before, if we attach an additional suffix, call it β, on top of α, it will c-command both α and everything that α c-commanded. In our discussion of Danish earlier we talked about *-t* attaching on top of *-en* and licensing F_{Nt} both on *-en* and on everything that *-en* c-commanded. Assuming that Danish′ has counterparts to the Danish F_{DEF} and F_{Nt}, and that these features are spread in the same way in both languages and have licensors that

behave like the Danish ones, we would expect Danish′ to have exactly the same pattern of definiteness and F_{Nt} as in Danish with the exception that there will be no need for a pre-nominal definiteness marker. The post-nominal suffix will license F_{DEF} on N and on the adjectives, and the post-definiteness suffix corresponding to *-t* will license the counterpart of F_{Nt} on the definiteness marker and on the adjectives. And since the post-nominal suffix will be good enough, ECONOMY will rule out the more complex pre-nominal form. To complete the analogy with Danish, where the realizer *-e* expressed either definiteness or plurality, Danish′ might use its own realizers to express various combinations of features on the adjectives. Schematically, and using for the moment the actual forms of the Danish suffixes, we would expect Danish′ to look like this:

(155) a. Indefinite: $[Adj-t]\ldots[Adj-t]\ [N]$

b. Definite: $[Adj-e]\ldots[Adj-e]\ [N-en-t]$

Real Icelandic, to which we now turn, will follow exactly the pattern in (155), but seeing that will require some effort, both because of its rich morphological paradigm and because of various phonological processes that sometimes obscure the underlying morphological structure.

4.6.2 Overview

Like Danish, Icelandic marks definiteness and does not mark indefiniteness (in fact, it does not even have an indefinite article). Like Danish, it also marks a distinction between singular and plural. The gender system of Icelandic is richer than that of Danish, though: it distinguishes between masculine, feminine, and neuter genders, distinctions that survive, at least in part, in the plural part of the paradigm. Finally, Icelandic differs from Danish in that it marks grammatical case, distinguishing between nominative, accusative, dative, and genitive. The distinctions are marked both on nouns and on adjectives, though in different ways, as we will shortly see.[35]

Focusing first on definiteness, it has been observed before that definiteness is marked

[35]There are further complications: phonological processes, which I will discuss, and inflectional paradigms, which I will not.

on the noun, regardless of whether it is a bare noun or whether it is modified by adjectives. Thus, for example, the Nom. Masc. Sg. noun for 'horse' is *hestur* when indefinite and *hesturinn* when definite, and the forms for the noun remain the same in the presence of adjectival modification: *gulur hestur* 'a yellow horse' and *guli hesturinn* 'the yellow horse'. The invariance of the nominal form under adjectival modification remains the same in the plural, as well as under changes of gender and case. With respect to definiteness, then, Icelandic follows our characterization of Danish'.[36]

Turning our attention to the marking of the various combinations of number, gender, and case, we observe that there are three different places within the noun phrase where elements vary according to these factors. One place is the adjectival ending. The two other places appear after the noun, one between the noun and the definiteness marker, and the other following the definiteness marker. Let us stay with the masculine singular '(yellow) horse,' and let us treat *-in* as the definiteness marker. The following preliminary decomposition of the surface forms arises:[37]

(156) *gul* + *hest* 'yellow horse' (masc.)

	-DEF		+DEF	
	Adj	*N*	Adj	*N*
Nom.	gul-ur	hest-ur	gul-i	hest-ur-in-n
Acc.	gul-an	hest	gul-a	hest-in-n
Dat.	gul-um	hest-i	gul-a	hest-in-um
Gen.	gul-s	hest-s	gul-a	hest-s-in-s

The position right after the noun (and before *-in* in the definite) will not concern us too much in what follows. I will call the affix that appears in that place as *C*1, just to have a name for it, but as far as I have been able to establish it does not participate in any pattern that might be relevant to the current proposal.[38]

The remaining two positions, the post-adjectival one and the one right after *-in*, will be

[36]Icelandic does have a pre-nominal definiteness marker, often described as part of a formal register. I will discuss it in section 4.6.4 below.

[37]The underlying forms are somewhat different. More on that in a moment.

[38]*C*1 seems to be absent from the definite form in the definite in (156). As discussed by Orešnik (1972), however, there are phonological reasons to expect the combination *i-in* to surface as *in*.

directly relevant to the analysis of Icelandic as Danish'. In Danish these were the positions that could host *-t*, and I will claim that in Icelandic they can host the affixes that correspond to the Danish *-t*. I will refer to these counterparts of *-t* collectively as $C2$, or, when it will be useful to refer to the $C2$ affix that corresponds to a particular combination ξ ($=< n, g, c >$) of number, gender, and case, I will write $C2_\xi$. My claim, then, will be that the adjectival suffix in an indefinite noun phrase is the same affix as the one that follows the definiteness marker in a definite noun phrase, and that, more generally, (156) instantiates the pattern in (155). Using $C1$ and $C2$ and *-in*, and writing *v* for the adjectival ending in the definite (corresponding to the Danish $-e$ in the same configuration), we can now restate (155) in Icelandic-related terms:[39]

(157) a. Indefinite: $[Adj - C2] \ldots [Adj - C2]\ [N - C1]$

b. Definite: $[Adj - v] \ldots [Adj - v]\ [N - C1 - in - C2]$

The pattern in (157) is what Icelandic would look like if it were our Danish'. How closely does this pattern match the actual data in (156)? The distribution of $C1$ and of the realizer *-v* does not seem directly relevant, and with respect to definiteness I already mentioned above that the distribution of *-in* is as predicted. So we are left with the question of whether the adjectival suffix in the indefinite is indeed the same as the affix that follows *-in* in the definite. The data in (156) close to what we want, but identity is not complete. For two of the rows, those of Dat. and Gen., the indefinite adjectival suffix is exactly the

[39]We can also use these new terms to restate our SRL story about the mechanism underlying the predicted pattern. In a definite noun phrase, a feature F_{DEF} will appear on the head noun and on any modifying adjective. The licensor for F_{DEF} is *-in*. In addition, for each combination $\xi =< n, g, c >$ of number n, gender g, and case c, there will be a corresponding feature (or combination of features) F_ξ that will appear on each adjective, as well as, in definite noun phrases, on *-in*. The licensor for F_ξ is the appropriate $C2_\xi$. As in Danish, each adjective in an indefinite noun phrase will need its own instance of $C2_\xi$ to license its F_ξ. For a definite noun phrase, again in analogy with Danish, only one instance of $C2_\xi$, the one on *-in*, will be needed: *-in* c-commands the adjectives (in order to license the instances of F_{DEF} below), and an occurrence of $C2_\xi$ c-commands *-in* (in order to license F_ξ on *-in*); consequently, the occurrence of $C2_\xi$ on *-in* also c-commands the adjectives below, licensing their instances of F_ξ. Finally, in the absence of a licensor, a realizer can appear on an adjective in a definite noun phrase, expressing the features that are present there (F_{DEF}, F_ξ, and perhaps other features as well).

same as the one that follows *-in* (*-um* for Dat. and *-s* for Gen.). For Nom. and Acc., on the other hand, the relevant surface forms are not identical. For Nom. we have an adjectival *-ur* in the indefinite and *-n* following *-in*, and for Acc. we have a post-adjectival *-an* in the indefinite and *-n* following *-in*.

To maintain the position that Icelandic is Danish', and that (157) is a correct characterization of the morphological reality, I would need an explanation for the apparent exceptions in the Nom. and in the Acc. One could appeal, perhaps, to contextual allomorphy, but it would be better if we could avoid this option and find an independent explanation that would allow us to actually predict where and how the surface forms for *C*2 differ.

Phonology offers precisely this kind of independent explanation. After certain adjectives that end with *in*, such as *heiðin* 'heathen' and *heppin* 'lucky,' the *C*2 endings in the indefinite are *-n* for both the nominative and the accusative, rather than the usual *-ur* and *-an* (cf. Einarsson, 1945, p. 53). Significantly, the *-in* in these adjectives has nothing to do with definiteness. Some processes, then, make the *C*2 forms for Nom. and Acc. surface as *-n* (instead of *-ur* and *-an*) after the *-in* in *heiðin* and *heppin*. It seems plausible enough that the same processes also apply to these *C*2 forms after the definiteness marker *-in*, changing them in the same way.

The phonological processes behind these changes are not entirely surprising. For example, what surfaces as the ending *-ur* has been argued by Orešnik (1972) (following Anderson, 1969a) to be underlyingly *-r*, with *u*-epenthesis taking place under certain conditions.[40] Another familiar process is the assimilation of *r* to *n* following *n* (cf. Anderson, 1974; Kenstowicz, 1994). We can state the relevant rules as follows:

(158) $\emptyset \rightarrow u \,/ C\_r\{C, \#\}$

(159) $r \rightarrow n \,/ n\_$

Ordering (158) before (159) ensures that the Nom. *C*2 ending would surface as *-n* after a

[40]Orešnik provides a broad range of phenomena that argue against having *u* in the underlying form of the suffix. For example, stem-final *j* drops unless it is followed by a vowel. Thus, masculine singular forms of the adjective stem *miðj* 'in the middle' (Einarsson, 1945) surface as *miðj-an* in the accusative and *miðj-um* in the dative but as *miðs* in the genitive; crucially, the nominative form is *mið-ur*, suggesting that at some level of representation, the vowel *u* was absent. Other diagnostics include *v*-deletion and interaction with *u*-umlaut.

final *-in*, regardless of whether it is attached to one of the adjectives we mentioned or to the definiteness marker. Elsewhere, in contexts that satisfy the conditions for epenthesis are met, *u* would intervene between the previous *n* and the following *-r*, bleeding assimilation.

For Acc., I can find no similar evidence for epenthesis.[41] I will tentatively assume that the Acc. form is due to a process of *a*-deletion:

(160) $a \rightarrow \emptyset \; / Cin\_n\#$

4.6.3 The rest of the paradigm

Given (157) and the indefinite forms, we can already predict correctly what the affix that follows *-in* will be for the Fem. Sg. part of the paradigm.

(161) *gul* + *kinn* 'yellow cheek' fem. sg.

	-DEF		+DEF	
	Adj	*N*	Adj	*N*
Nom.	gul	kinn	gula	kinn-*in*
Acc.	gul-**a**	kinn	gulu	kinn-*in*-**a**
Dat.	gul-**ri**	kinn	gulu	kinn-*in*-**ni**
Gen.	gul-**rar**	kinn-ar	gulu	kinn-ar-*in*-**nar**

The indefinite adjectival endings and the post-definiteness endings are identical for Nom. and for Acc. For Dat. and for Gen. the the only difference is that where the indefinite adjectival ending begins with *r* the post-definiteness begins with *n*. This is what we expect given (159), and we expect it to be a phonological effect: the indefinite adjectival ending where the stem ends with *in* should similarly be *n*-initial. Thus we are not surprised to find that the Fem. Sg. forms of the adjective 'heathen' are *heiðin-ni* in Acc. and *heiðin-nar* in Gen., rather than the usual *-ri* and *-rar* endings (cf. Einarsson, 1945, p. 53).

For the neuter singular we will need one final phonological rule.

[41] In fact, the diagnostics that show that *u* is epenthetic in *-ur* suggest that the *a* in *-an* is present underlyingly. For example, stem-final *j* does not drop before the *C*2 Acc. ending in forms like *miðj-an*, mentioned above.

(162) *gul* + *barn* 'yellow child' neut. sg.

	-DEF		+DEF	
	Adj	*N*	Adj	*N*
Nom.	gul-**t**	barn	gula	barn-i-**ð**
Acc.	gul-**t**	barn	gula	barn-i-**ð**
Dat.	gul-**u**	barn-i	gula	barn-i-***n*-u**
Gen.	gul-**s**	barn-s	gula	barn-s-***in*-s**

For Dat. and Gen., the indefinite adjectival ending is identical to the post-*in* ending, as expected. For Nom. and for Acc. we find *-ið* instead of the predicted *-in-t*. As with the Masc. endings for Nom. and Acc., what we want is a process of assimilation. I will suggest the following rules, with (163) ordered before (164):

(163) $t \rightarrow$ ð /*Cin*_

(164) $n \rightarrow \emptyset$ /_ð#

Again, adjectives ending with *in* provide evidence that this is indeed a phonological process: the Nom. and Acc. forms of 'heathen' (neut. sg.) are both *heiðið* (Einarsson, 1945, p. 53).

And we can already predict the correspondence in the plural part of the paradigm without further modification:

(165) Masculine: *gul*+*hest*+Pl.

	-DEF		+DEF	
	Adj	*N*	Adj	*N*
Nom.	gul-**ir**	hest-ar	gulu	hest-ar-***n*-ir**
Acc.	gul-**a**	hest-a	gulu	hest-a-***n*-a**
Dat.	gul-**um**	hest-um	gulu	hest-u-***n*-um**
Gen.	gul-**ra**	hest-a	gulu	hest-a-***n*-na**

Feminine: *gul*+*kinn*+Pl.

	-DEF		+DEF	
	Adj	*N*	Adj	*N*
Nom.	gul-**ar**	kinn-ar	gulu	kinn-ar-*n*-**ar**
Acc.	gul-**ar**	kinn-ar	gulu	kinn-ar-*n*-**ar**
Dat.	gul-**um**	kinn-um	gulu	kinn-u-*n*-**um**
Gen.	gul-**ra**	kinn-a	gulu	kinn-a-*n*-**na**

Neuter: *gul+barn*+Pl.

	-DEF		+DEF	
	Adj	*N*	Adj	*N*
Nom.	gul	börn	gulu	börn-*in*
Acc.	gul	börn	gulu	börn-*in*
Dat.	gul-**um**	börn-um	gulu	börn-u-*n*-**um**
Gen.	gul-**ra**	barn-a	gulu	barn-a-*n*-**na**

4.6.4 Independent definiteness marking

The structure of the Icelandic noun phrase makes it unnecessary to use a free article-like definiteness marking, and because of ECONOMY we expected that such marking would in fact be impossible. But ECONOMY is relevant only as long as everything other than licensors remains fixed. If we could find a non-vacuous pre-adjectival head in Icelandic that can combine with *-in* the result would not mean the same as the usual definite form with a post-nominal *-in*, and so ECONOMY will not rule it out. In such case, we expect two further things to happen. First, the post-*D -in* will make the post-*N* occurrence superfluous. Second, the post-*in C*2 would c-command the adjectives, making lower occurrences of *C*2 superfluous.

Icelandic has a pre-nominal marker that allows us to test these predictions. The precise meaning of this marker is somewhat unclear (my informants reported that the use of this marker implied some emphatic or emotional value, in addition to definiteness). Morphologically, the marker is composed of a base *hin*, which I will analyze as *h+in*, and a suffix, which we expect to be *C*2. Once the phonological rules above are taken into account, the entire paradigm is derived:

(166) a. Singular:

	M.	F.	N.
Nom.	hin-**n**	hin	hið
Acc.	hin-**n**	hin-**a**	hið
Dat.	hin-**um**	hin-**ni**	hin-**u**
Gen.	hin-**s**	hin-**nar**	hin-**s**

b. Plural:

	M.	F.	N.
Nom.	hin-**ir**	hin-**ar**	hin
Acc.	hin-**a**	hin-**ar**	hin
Dat.	hin-**um**	hin-**um**	hin-**um**
Gen.	hin-**na**	hin-**na**	hin-**na**

And as predicted, if the pre-nominal marker is used, the definiteness suffix cannot appear on *N*, and there are no occurrences of *C*2 below:

(167) a. hinn góði hestur
the good horse

b. * hinn góði hestur-in-n
the good horse

c. * hinn góð-ur hestur
the good horse

4.6.5 Interim summary

We started with the idea of Icelandic as Danish′, where we changed nothing other than the the label of the sister of *AP* (*NP* in Danish, and *N* in Danish′). By using five independently motivated phonological rules for Icelandic, we derived the entire correspondence between indefinite adjectival endings and post-definiteness endings in the full paradigm. Our ability to make these predictions supports both the general SRL idea, and so strengthens our case for structural economy in grammar.

4.7 Greek

4.7.1 Moving things inside the noun phrase

We saw that the marking patterns in Danish and Icelandic noun phrases lend themselves to an SRL analysis, indirectly supporting a notion of structural competition in grammar. The new part of the analysis was the introduction of *licensors*, a family of function elements that associate with one or more instances of a feature, possibly at a distance. The domain of licensing appeared to be structurally defined (specifically, an adjunction-sensitive version of c-command seemed to provide the right results). The evidence for licensors and their domains came from the accumulation of distributional facts that they helped predict. The core Danish facts already suggested that an SRL approach would have an advantage over licensor-free accounts. Many further data points, both language-internally and cross-linguistically, provided further support for this idea.

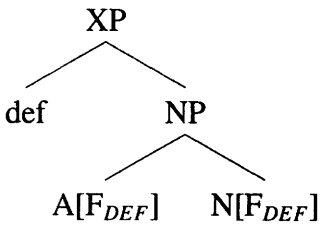
As mentioned in the introduction, however, there remains a general concern that should be addressed. As is often the case in morphology, the Germanic data that we have tried to analyze made it hard to evaluate competing hypotheses with respect to arbitrarily complex configurations. We were able to increase the level of complexity by considering further elements within the phrase, partially-overlapping marking patterns, and different choices with respect to phrase structure. Eventually, however, we will run out of new modifiers, patterns, and phrase-structural choices, and even before that we might run out of languages that would allow us to test the relevant combinations. In addition, when we switch to different constructions and different languages it is often hard to ensure that all other factors of relevance remain without change. We could try to show, as we did, that our assumptions have at least some independent plausibility, either by using familiar structural diagnostics or by pointing out new predictions, but ultimately all our evidence in favor of SRL came from the ability of the system as a whole to predict a finite collection of paradigms. What is missing is the ability to take an initial set of configurations and start complicating them in a controlled (and meaningful) way.[42] If we could move elements within the noun phrase

[42]By *meaningful* I mean relevant for comparing competing theories. Stacking additional adjectives before a Danish noun or PPs after it would make the noun phrase more complex in some sense, but all current

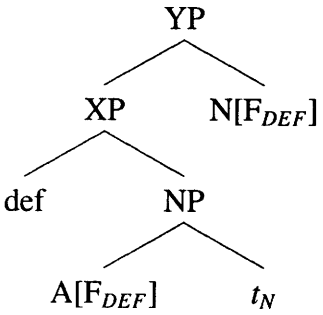
we would be able to complicate our configurations in the way we need. Unfortunately, however, Danish and Icelandic have a fairly rigid word order within the noun phrase, and the same seems true for other Germanic languages as well.

So what would Danish look like, for example, if it allowed elements within the noun phrase to move around? The answer, assuming that our analysis of the Germanic data is correct, depends on what can move where and on what positions are available for licensors. Suppose, first, that both *A(P)* and *N* can right-extrapose to a position higher than the position corresponding to the Danish *den*.[43] Suppose further that when the adjective moves, its features must be licensed in their *surface* position (that is, c-commanding the trace is insufficient for licensing the moved element). (168) shows the base, Danish-like, configuration. The configuration for right-extraposition is shown in (169a) for *N*, string-vacuously, and in (169b) for *A(P)*.

(168)



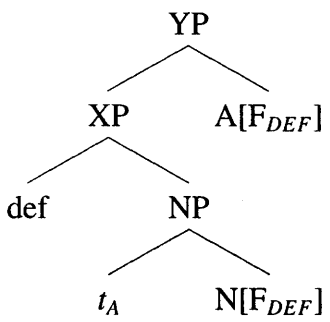
(169) a.



theories predict that the choice between *den* and *-en* will not change when moving from one adjective to more or from one PP to more.

[43]Or perhaps more naturally, especially from the point of view of structural economy, we could assume that movement just has to be structurally non-vacuous, cached out in terms of changing hierarchical relations between the moved constituent and at least one other (terminal) node. This would allow *N* to move above *A* but below *den*. Below, when we get to possessive clitics, we will see that we can diagnose such this kind of local movement. Until then, though, either of these two notions of movement would be equally applicable.

b.

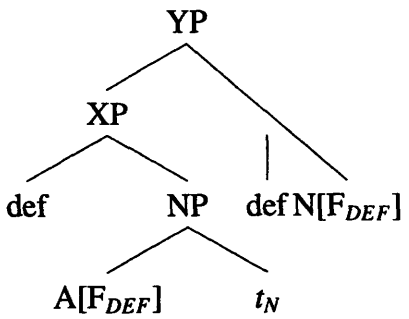


In the basic case, shown in (168), the licensor *def* is sister to *NP* and can license F_{DEF} both on the adjective and on the noun. In each of the dislocated structures, however, we now have an unlicensed instance of F_{DEF}. If *def* can attach not only where we see it above, as sister to *NP*, but also as sister to *YP* or higher, doing that would resolve the licensing issue, making the lower *def* redundant. All we would see, unless there were some intervening element between the two positions, would be a Danish-like language with the possibility of reordering the noun and the adjectives. The higher attachment site, then, makes the pattern very similar to the basic Germanic pattern.[44]

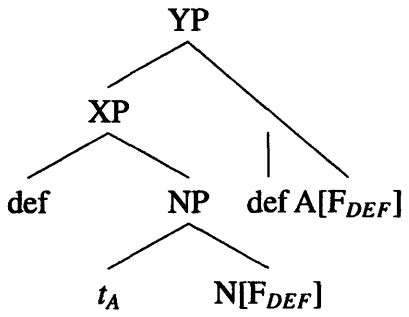
A more interesting pattern would emerge if such a higher position were not available, but if lower, more local attachment sites were available instead. For example, imagine that *def* could take not only *NP* but also *N* and *A(P)* as its sister. In the basic configuration, (168), we would have no need for those positions, since *def*'s higher attachment site (as sister to *NP*) would suffice to license all the occurrences of F_{DEF}. In the dislocated structures, however, the local attachment sites of *def* can come to the rescue. The dislocated element can have a licensor as its sister, licensing its F_{DEF}, while the original occurrence of *def* licenses F_{DEF} in the remnant, as before. Neither attachment site c-commands the other, though, so we would get two instances of *def*. Here are the two structures:

[44]The pattern just discussed bears a certain resemblance to various patterns of adjectival modification in Romance. It would be interesting to see whether this resemblance is significant, and whether our approach can account for the Romance data, but I will not pursue this idea here.

(170) a.



b.



For a language that has the attachment possibilities just mentioned, but which is otherwise like Germanic, our predictions so far are the following. The basic word order is [*def A(P) N*], but where dislocation takes place we will find [*def A(P) def N*] and [*def N def A(P)*]. A post-nominal occurrence of *A(P)* is obligatorily preceded by *def*, and if we have independent ways to detect movement, we should be able to confirm that the order [*def A(P) def N*] arises exactly when dislocation takes place. We can also say something about indefinite noun phrases. If the current language is like Danish in treating definiteness as a privative feature, with no indefinite counterpart for F_{DEF}, dislocation in an indefinite noun phrase should not disrupt any licensing relations. What we expect, then, is that along with a basic [(1) *A(P) N*] we should also find [(1) *N A(P)*], but crucially neither [(1) *N* 1 *A(P)*] nor [(1) *A(P)* 1 *N*].

The predictions we have are elaborate and quite specific, and the pattern they lead us to expect is superficially quite different from Germanic. More generally, our hypothetical pattern would be surprising for licensor-free accounts: if *def* is a realizer we would not expect an element to acquire a new occurrence of it by leaving the domain of definiteness (if anything, we might expect to find the opposite); and if *def* is a spreader, any additional occurrence of it would be at odds with everything we believe about the semantics of definiteness. From our current perspective, on the other hand, the pattern is an almost immediate relative of the Germanic marking patterns that we saw.

Normally, this state of affairs would be a cause for concern. We are committed to a particular and unusual set of predictions, and if we fail to find a language that bears these predictions out (or reasons to rule out such a language on independent grounds), we would be at a disatvantage compared to the rest of the world. Of course, if we do find a language that behaves like Germanic with dislocation, the situation would be reversed, and we will be alone in predicting yet another complex pattern, using little more than what we already had in place long ago. As the reader may have already concluded, I believe our current situation to be of the second kind. The language I have in mind is Modern Greek, and its intricate pattern of definiteness marking has generated a fair amount of theoretical interest, though, to my knowledge, the proper way to account for it has remained an open question.[45]

Here is the basic pattern of word order and definiteness in Modern Greek:

(171) a. to megalo (to) vivlio
the-NEUT.SG big-NEUT.SG (the-NEUT.SG) book-NEUT.SG
'the big book'

b. to vivlio *(to) megalo
the-NEUT.SG big-NEUT.SG *(the-NEUT.SG) book-NEUT.SG
'the big book'

(172) a. ena megalo (*ena) vivlio
1-NEUT.SG big-NEUT.SG (*1-NEUT.SG) book-NEUT.SG
'a big book'

b. ena vivlio (*ena) megalo
1-NEUT.SG big-NEUT.SG (*1-NEUT.SG) book-NEUT.SG
'a big book'

The definiteness marker usually precedes the adjective, which in turn precedes the noun.[46] It is possible to have a post-nominal adjective, though this requires a second definiteness

[45]For discussion of the facts and for some of the main proposals see Androutsopoulou (1996), Alexiadou and Wilder (1998b), Manolessou (2000), Kolliakou (2004), Ioannidou and den Dikken (2007), and Lekakou and Szendrői (2007), Marinis and Panagiotidis (2007), as well as references cited therein. The terms *Determiner Spreading* (from Androutsopoulou) and *Polydefiniteness* (from Kolliakou) are often used to refer to one of the main aspects of the pattern. Mine is by no means the first attempt to bring together the Greek patterns with the Germanic ones. See Alexiadou (2003) for very relevant discussion of the cross-linguistic setting. See also Leu (2007) for an interesting, though preliminary account that treats Greek and Germanic in parallel using remnant movement.

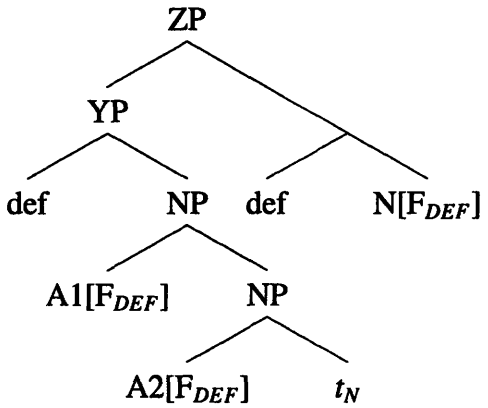
[46]In the absence of an adjective, the definiteness marker still precedes the noun. There is no post-nominal

marker immediately preceding the adjective. And it is possible, though not necessary, for the noun to have its own marker even in the basic order. For indefinite noun phrases, the basic word order is as in the definite case, but a post-nominal adjectival position is also available. In contrast with the definiteness marker, we never find multiple occurrences of 1. As I said, just what our Germanic-based story predicts.

What happens when more than one thing moves? Let us look at a noun phrase with two attributive adjectives. As in Germanic, the adjectives typically stack up on top of the noun and appear to its left. On our assumptions, a single instance of *def* attached as sister to *NP* should suffice to license F_{DEF} on all the adjectives and on the noun. This is all as one might expect by extrapolating from the case of a single adjective or by using the parallelism with Germanic.

Suppose now that we try to right-dislocate the noun. If we could do that and change nothing else, we would find something like this:

(173) Hypothetical structure for dislocating *N* and keeping the rest fixed



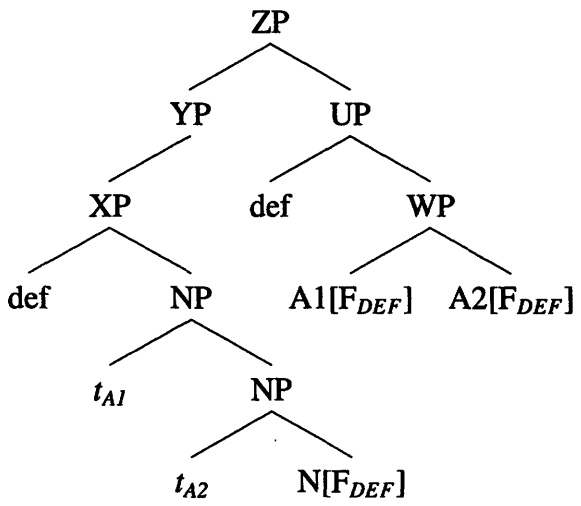
The original *def* would license the instances of F_{DEF} on *A*1 and *A*2, while, as in the case of modification by a single adjective, an extra *def* would be needed to license F_{DEF} on the dislocated *N*. The bi-definiteness in (173), however, relies on the structure being well-formed (up to licensing), and (173) has one property that makes it an unlikely configuration. In the remnant *NP*, we have two adjectives stacking without a noun at the bottom. Cross-linguistically, however, adjectives can only stack on top of a noun, and Greek appears to be no exception. For example, two adjectives in a predicative position require overt

suffix option. In this respect, definiteness in Greek more similar to languages like German and Dutch than to their Scandinavian cousins. None of this matters here.

conjunction (the equivalent of *John is tall *(and) stupid*). It also seems that what is required is more than having a noun in the right place at some point in the derivation: light nouns, which some accounts treat as moving up across the adjective, do not allow stacking (the Greek equivalent of *something big *(and) black* also requires conjunction).[47] The noun can be overt (as it is in most cases), or it can be covert, as in noun ellipsis (the Greek and Hebrew equivalents of *I'll have the big red one* do not have an overt counterpart for *one* and are grammatical), but a noun must be there at the end of the derivation. We would expect, then, that (173) would be ungrammatical, quite independently of issues of licensing. And it is: [*def A*1 *A*2 *def N*] is not a possible order.[48]

A similar problem prevents the two adjectives from appearing on the right and sharing a single licensor.

(174)



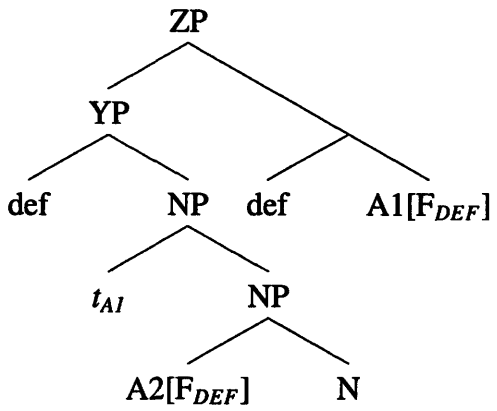
[47]See Kishimoto (2000) for an analysis along these lines. See Larson and Marušič (2004) for why it doesn't work.

[48]I don't know why adjectives cannot stack on their own, but it seems to me that type mismatches, along lines discussed by Irene Heim in a couple of unpublished handouts, would be a promising way to go. In the framework of Heim and Kratzer (1998) adjectives are of the same type as nouns, and their combination is intersective, using Predicate Modification (PM). If PM is indeed available, we would predict stacking as a general phenomenon, contrary to fact. On the other hand, if PM was not available (let alone some of the new modes of combination proposed recently, such as Chung and Ladusaw, 2004, 2006), and if Function Application (FA) were be the only option for combining two elements, we would predict that two items of the same type could never combine. This would correctly rule out stacking in the absence of an element of a different type down below. For our cases here, and in order to rule out stacking with *N*-movement more generally, one would also have to prevent the moved noun from leaving a trace of the type of nouns.

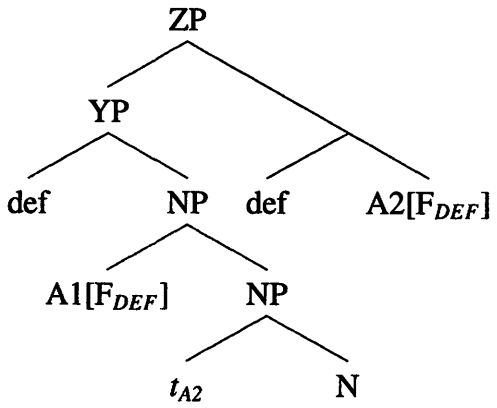
(174) is an attempt to cluster *A*1 and *A*2 together on the right. This would allow them to use a single instance of *def* to license their both of their occurrences of F_{DEF}. Even if we could find some way to move the adjectives to the relevant positions, the prohibition on noun-less adjectival stacking would rule out the structure. We should therefore not be surprised to discover that [*def N def A*1 *A*2] is not a possible order.

So those were two things that cannot happen with multiple adjectives. Let us now look at some things that can happen. One thing we predict should be okay is the movement of a single adjective to a post-nominal position. As before, this adjective will have an unlicensed F_{DEF}, so it will cause an extra licensor to appear. Either adjective can do that.

(175) a. *A*1 moves



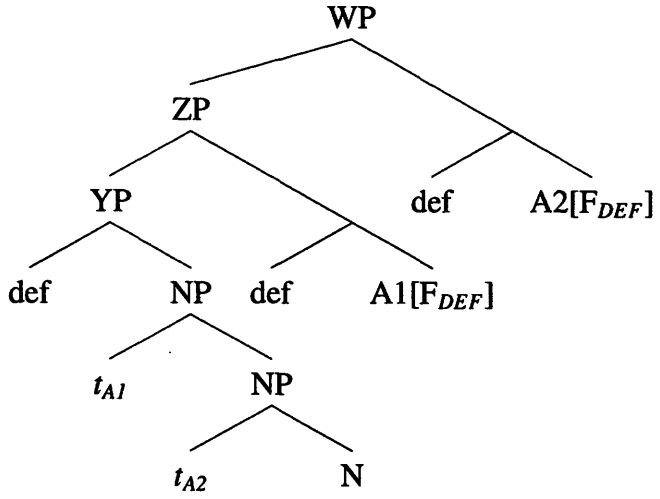
b. *A*2 moves



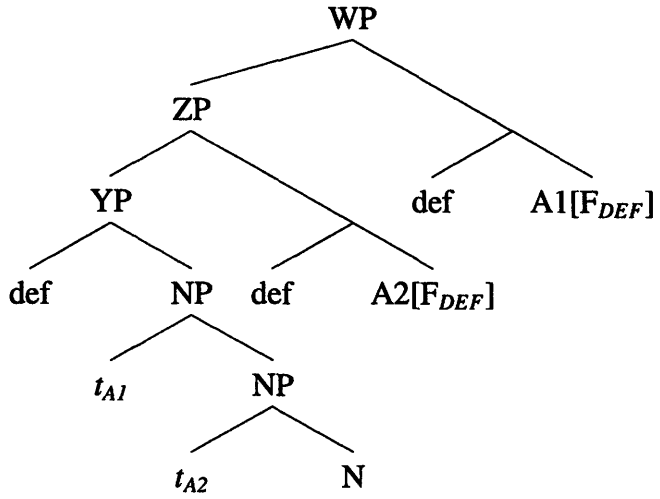
Or both adjectives could move, one after the other, stacking above *N*. Either adjective can be the first to move.[49]

[49]This assumes that the operation that dislocates the adjectives is not subject to intervention effects and does not render the remnant domain opaque.

(176) a. *A*1 moves first

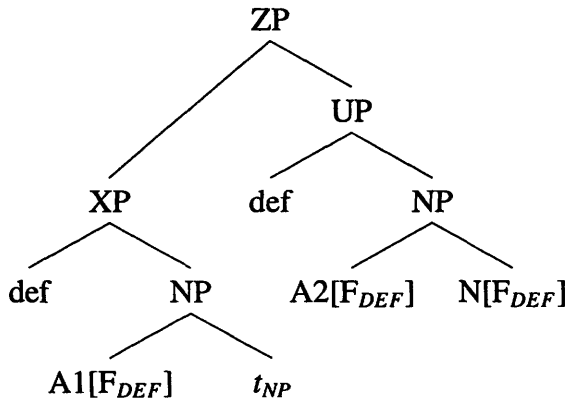


b. *A*2 moves first



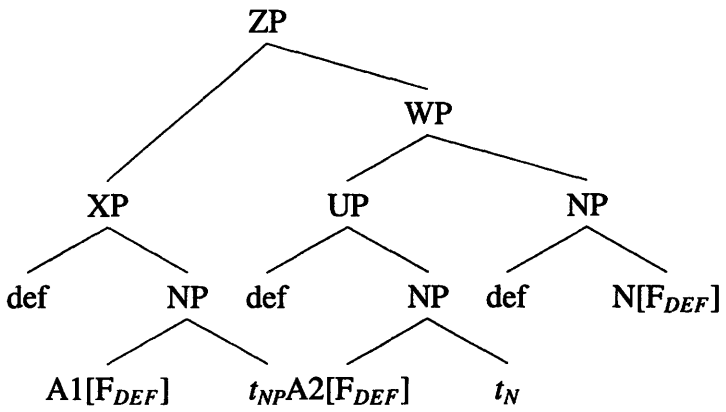
Another thing that could happen is that the whole *A*2 + *N* combination would move:

(177)



Finally, the movement in (177) can be combined with a local movement of *N* above *A*2 or of *A*2 above *N*. As long as *N* lands below *ZP*, no adjective stacking would occur.

(178)



In other words, any number of adjectives can stack above the noun and appear to its left, sharing a single licensor; any other adjective, either pre- or post-nominally would require its own licensor. Using the terminology of Kolliakou (2004), we could say that an obligatory *monadic* structure (which I take to consist of *def* on the left, *N* on the right, and zero or more adjectives in between) is embedded within an optional *polydefinite* structure (which I take to consist of pairs of *def* and *A*). Schematically, we could write it like this.

(179) Predicted word order possibilities in definite Greek noun phrases:

$(Def\ A)*\ Def\ A*\ N\ (Def\ A)*$

As far as I can see, our predictions capture precisely the patterns described in the literature by Alexiadou and Wilder, Kolliakou, Marinis and Panagiotidis (2007), and others.[50]

4.7.2 Further evidence for movement

We started the section on Greek by noting that our Danish-based system makes predictions that are difficult to test in Germanic languages, due to the rigidity of word order. We were able to use indirect tests, like switching from definite to indefinite, combining two marking

[50]Disturbingly, I have not been able to replicate these reported judgments in person. In particular, Sabine Iatridou and Giorgos Spathas, in independent p.c.'s, reject out of hand any combination of a non-trivial monadic domain with a polydefinite one, in either order. In the current system it might be possible to state this as a requirement, active in some dialects, that if something is dislocated within the phrase, everything else that can be dislocated should also move. I will have a bit more to say about this variation in what follows, but I would like to state at this point that I know of no reason to expect such a requirement of all-or-none in the first place. The variation in combining polydefinite and monadic domains may well be teaching us that something serious is wrong in the current proposal.

patterns, adding adjectives, or looking at another language. But we wanted something more direct than that. If we could fix one construction in one language and play with hierarchical relations using a mechanism independent of licensing, we would have a way to test our predictions on arbitrarily complex configurations (or at least come closer to this goal). We looked at what these predictions would be in a language that had the single mechanism of right-dislocation for rearranging hierarchical relations within the noun phrase. And we saw that those predictions are borne out in their entirety if we assume that Greek is Danish with this kind of right-dislocation.

I find it reassuring that all we needed to use Danish to predict Greek was the assumption of right-dislocation. It would be better still if we could provide further motivation for right-dislocation, especially since this is not a particularly wide-held belief about the proper treatment of Greek noun phrases.[51] I can think of several good reasons to like the idea of right-dislocation, all having to do with the interaction of the word order facts that we have seen above for adjectives and nouns with factors other than definiteness marking. In the remainder of this subsection we will look at interactions with indefiniteness, adjectival hierarchies, and possessive clitics. Then, in the next and final subsection of our discussion of Greek, we will look at the interaction of word order and information structure. This will provide further evidence for right-dislocation, but more importantly, it will offer a preliminary answer to why movement takes place in the Greek DP in the first place.

Adjectival orderings

Above we focused on the ordering of adjectives with respect to elements of a different category (*N*, as well as the category of *def*, whatever that might be). In addition, Greek adjectives follow a cross-linguistically common (though still poorly understood) pattern of respecting certain ordering constraints with respect to other adjectives modifying the same

[51]Though differing quite widely in their proposals, Androutsopoulou (1996), Alexiadou and Wilder (1998b), Leu (2007), and Ioannidou and den Dikken (2007) all assume that movement is strictly to the left. Kolliakou (2004) and Lekakou and Szendrői (2007) assume that polydefiniteness involves no movement at all.

noun.[52] For monadic definites, the facts look similar enough to English.

(180) Monadic definites: rigid adjectival order

a. the big red book *to megalo kokkino vivlio*

b. * the red big book * *to kokkino megalo vivlio*

As noted by Androutsopoulou (1996), these restrictions remain more or less intact in polydefiniteness when the adjectives are pre-nominal, but they seem to become inactive in post-nominal positions:

(181) Polydefinites: rigid adjectival order pre-nominally

a. the big the red the book *to megalo to kokkino to vivlio*

b. (*) the red the big the book *(*) to kokkino to megalo to vivlio*

(182) Polydefinites: free adjectival order post-nominally

a. the book the big the red *to vivlio to megalo to kokkino*

b. the book the red the big *to vivlio to kokkino to megalo*

Let us follow Androutsopoulou (1996) and Alexiadou and Wilder (1998b) in taking the monadic configuration as basic, and let us follow them in assuming that (180a) is base generated and that (180b) is not.[53] Differently from them, however, we can already derive the pattern of acceptability. (180b) is bad because it is not base-generated, and because right-dislocation cannot move *big* past *red* without also crossing *book*, and it cannot cross *book* without incurring an extra definiteness marker. With polydefinitess, we can get each of the post-nominal orders, by first moving *big* and then moving *red* (for (182a)) or the other way around (for (182b)), as we already discussed above. We get (181a) by dislocating the

[52]See Sproat and Shih, 1988 for discussion of the general setting, and Androutsopoulou, 1996 and Alexiadou and Wilder, 1998b for discussion of adjectival orderings in Greek

[53]Other proposals typically either remain silent on the issue of adjectival orderings or dispute the unacceptability of (180b) *to kokkino megalo vivlio*. I am not aware of speakers who agree about the badness of (180b) but find (181b) acceptable, once the relevant discourse conditions are satisfied. I am also not familiar with speakers for whom the post-nominal position is sensitive to orderings (that is, speakers who get a contrast between (182a) and (182b) in some direction). If it turned out that such speakers exist, I would have to revise the current proposal. Otherwise, the existence of some speakers who get the contrasts above are sufficient for our purposes.

lower noun phrase [*red book*] and then dislocating the noun above *red*.[54] Finally, we have no way to generate the unacceptable (181b).[55]

As Alexiadou and Wilder (1998b) observe, the adjectival ordering pattern is not limited to definite noun phrases.

(183) Indefinites: rigid adjectival order pre-nominally

a. a big red book *ena megalo kokkino vivlio*

b. (*) a red big book *(*) ena kokkino megalo vivlio*

(184) Indefinites: free adjectival order post-nominally

a. a book big red *ena vivlio megalo kokkino*

b. a book red big *ena vivlio kokkino megalo*

For us, again, this is as expected. (183a) is base-generated. (183b) is not, and we cannot reverse the order of the two adjectives without also crossing the noun. And we can derive each of the post-nominal orders by choosing which adjective to move first.

There is another ordering-related fact about adjectives that has been discussed extensively in the literature. While the monadic construction can use any adjective in Greek, polydefiniteness appears much more selective. For example, non-subsective adjectives like *ipotithemenos* 'alleged' (an alleged murderer is not necessarily a murderer) are degraded when used in polydefiniteness, as are adjectives that, informally speaking, provide an argument of the noun rather than a modification, as in the use of *italiki* 'Italian' as the agent of an action nominal (an Italian invasion can mean an invasion *by* Italy and not just an invasion in an Italian style). Alexiadou and Wilder (1998b) use this restriction, along with a similar restriction on predicative positions, to argue for two domains within the noun phrase: a lower domain, where all adjectives can attach, stacking above the noun and below the first determiner, and a higher domain of clausal structures, each clause hosting its own predicative adjective. It is the higher domain, on their account, that gives rise to polydefiniteness, which explains the restriction of the adjectives in polydefiniteness to those that can appear in predicative positions.

[54]This is actually more complicated than I'm trying to make it appear.

[55]Or maybe we do.

Interestingly, as as been discussed in the literature, the restriction of polydefinites to predicative adjectives is not absolute. Kolliakou (2004) mentions polydefiniteness in proper names (*i Maria i Papadopoulou* 'the M. the P.), and Leu (2007) reports that some speakers, at least, accept *o proigumenos o prothipurgus* 'the previous the president' in the relevant context.[56] What makes these apparent exceptions to the generalization of Alexiadou and Wilder (1998b) significant for our proposal is that, as far as I have been able to determine, reversing the order in these polydefinite constructions often leads to degradation. Thus, even speakers who accept the examples above either reject their reversed variants or find them marked: *?i Papadopoulou i Maria* '?the P. the M.' and **o prothipurgus o proigumenos* '*the president the previous' is out.[57] For an account such as Alexiadou and Wilder (1998b) this contrast is surprising: non-predicative adjectives are supposed to be generated within the lower domain, and their response to the original counter-examples would probably involve allowing those adjectives (or names) to be exceptionally generated higher up under the relevant discourse conditions; but once this step is done, moving the noun would be no different than in any other polydefinite construction. It is the adjective that is special, not the noun. It seems much more natural to derive the unacceptability of the reverse order from the reluctance of the relevant adjectives to move. In our account, the basic order results from movement of the noun, and only the reverse order requires the adjective to move.

The possessive clitic

A more direct diagnostic for right-dislocation within the Greek noun phrase is offered by the possessive clitic.[58] For an unmodified noun, the possessive appears post-nominally and does not take a separate definiteness marker. The only possibility for a definite unmodified

[56]Both authors use contrastive or corrective contexts to facilitate the judgment. We will be able to say more about this in section 4.7.3 below. See Lekakou and Szendrői (2007) for further discussion of non-predicative adjectives in polydefinite constructions.

[57]My survey included one speaker.

[58]Our concern here will be the distribution of the possessive with respect to the other elements within the noun phrase. See Kolliakou (2004) for a discussion of the morphology of the various possessive forms and Alexiadou (2005) for a discussion of their semantic interaction with definiteness.

noun is in (185). Some of the bad options are listed in (186).

(185) to vivlio mu
the book mine
'my book'

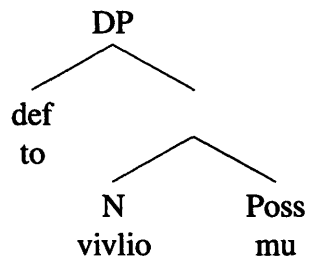
(186) a. * to mu vivlio
the mine book

b. * to vivlio to mu
the book the mine

c. * vivlio to mu
book the mine

It seems reasonable to assume that the only grammatical possibility, (185), corresponds to the base-generated structure, and that the possessive clitic is attached somewhere between *def* and *N*. Here is a schematic representation of the general idea:

(187)



If this is correct, the possessive clitic should be something of a right-boundary marker for the Greek noun phrase, allowing us to diagnose movement to the right. To use it, though, we would need more material within the phrase. Adding an adjective to (185) makes possible two different monadic forms:

(188) a. to kokkino vivlio mu
the red book mine
'my red book'

b. to kokkino mu vivlio
the red mine book
'my red book'

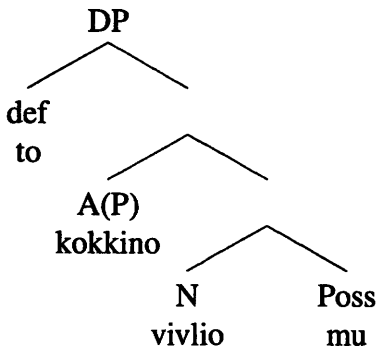
We noted above that if right-dislocation is governed by some kind of non-vacuity condition, it might be possible for the noun to move above the adjective without leaving the licensing domain of their joint *def*. The availability of (188b) provides evidence for that idea.

The noun appears to the right of *mu*, suggesting that it has moved. And yet it has no additional definiteness marker. On the same assumptions, the contrast between the grammatical (188b) and the ungrammatical (186a) seems to teach us either that the possessive does not count for the relevant economy condition or that it is attached at most as high as the noun. The latter option is supported by the ability of the noun and the possessive to appear after a separate definiteness marker:

(189) to kokkino to vivlio mu
the red the book mine
'my red book'

This indicates that the noun and the possessive can form a constituent together, to the exclusion of the adjective, and that (189) is the result of dislocating this constituent above the original definiteness marker. It appears, then, that the basic structure for a definite noun phrase like (188a), with an adjective and a possessive clitic is something like this:

(190)



What else would we expect to see? First, it should be possible for the noun to right-dislocate to a position outside the licensing domain of *def*, incurring its own marker. And it should be possible for the adjective to cross the noun. In other words, both of our earlier polydefinite options should still be available. This seems correct:

(191) a. to kokkino mu to vivlio
the red mine the book
'my red book'

b. to vivlio mu to kokkino
the book mine the red
'my red book'

Those are all the possibilities that we predict for a definite noun phrase. Nothing else should be grammatical, and as far as I have been able to establish, nothing else is. Moving on to

indefinites with a possessive clitic, we predict that the grammatical options will be all those that correspond to one of the grammatical definite ones, without the definiteness markers, and with an indefinite article on the left. This, too, seems to be correct.

The interaction of possessive clitics with word order in definite and indefinite noun phrases seems almost trivial from the perspective of our Danish-based assumptions and the idea of right-dislocation as the only reorganizating process within the Greek noun phrase. For other accounts, where movement is usually to the left, accounting for the distribution of possessive clitics seems much more complicated, even if we ignore the problems of polydefiniteness. I take this as further evidence in favor of right-dislocation and of the current proposal more generally.

4.7.3 Why should things move inside the noun phrase?

We have seen how Greek allows us to test the SRL theory in more or less arbitrary configurations. What we needed to add to our earlier system was the idea of right-dislocation within the Greek noun phrase. This was enough to predict the complex pattern of polydefiniteness, thus not only providing support to our proposal but also offering a solution to an open problem. The assumption of right-dislocation received further support from the word order fact in indefinite noun phrases and in noun phrases with possessive clitics.

In this subsection I will try to sketch one final argument for right-dislocation. We will look at the interaction of polydefiniteness and information structure, discussed by Kolliakou (2004) and others, and we will see that right-dislocation makes it possible to capture the tight correspondence of the distributional word-order facts and the semantic/pragmatic properties of polydefiniteness. When this is spelled out, it would allow us to use meaning as another diagnostic for movement, helping us convince ourselves further that our earlier claims were reasonable. More significantly, though, it will provide us with a simple answer to the question of why things should ever move within the noun phrase to begin with. Since the other movement approaches on the market have no answer to this question, we will gain some additional points. One we collect those points we will stop.

As a rough generalization, using a polydefinite noun phrase felicitously requires a con-

text that makes the noun phrase or parts of it anaphoric or contrastive in some sense.[59] More significantly for our purposes, polydefiniteness is also bad in all-given contexts:

(192) (Ioannidou and den Dikken, 2006, p. 7, based on Kolliakou, 2004, p. 271)

a. Context: o Janis taise tis mikres gates
the Janis fed the small cats

b. #i gates i mikres itan pinasmenes
the cats the small were hungry

c. #i mikres i gates itan pinasmenes
the small the cats were hungry

It seems that some kind of NP-internal contrast with respect information structure is needed to make polydefiniteness felicitous. From the current perspective, where polydefiniteness involves movement outside the domain, this requirement is vaguely reminiscent of the sensitivity to information structure in Slavic split nominals (Gouskova, 2001; cf. also Fanselow and Ćavar, 2002). The idea that I will try to pursue here is that Greek is similar to Slavic languages in separating given material and new material within the noun phrase. But whereas Slavic languages can move the given part of the noun phrase all the way to the top of the clause, Greek only moves material locally, attaching the given part within the noun phrase (or very close to it). More generally, the current proposal will join the proposals of Arregui-Urbina (2002), Wagner (2007), and Kučerová (2007), according to which movement sometimes serves the role of removing given material from the focus domain. These works, which I will refer to as *interface-based* approaches to information structure, share a perspective that assigns a lesser role for syntax and a greater role for the interfaces (the phonological interface in the case of Arregui-Urbina; the semantic interface for Wagner and Kučerová) in constraining movement than is assumed under the so-called *cartographic* (or *feature-based*) approach to information structure of Koopman and Szabolcsi (2000) and others. It will be useful to take a quick detour and say a bit more about this distinction and about why the interface-based approach is better.

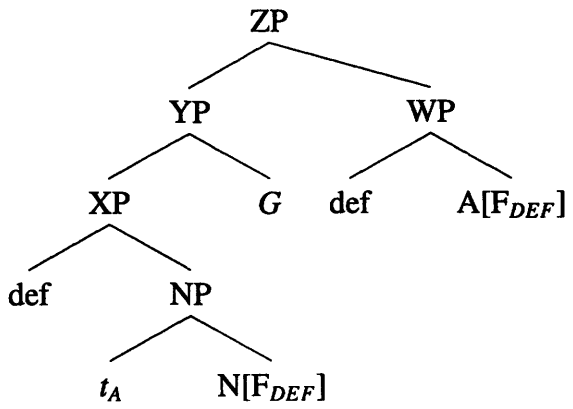
Cartographic approaches follow the Minimalist idea that movement is triggered by par-

[59]The condition of anaphoricity or contrast is discussed by Kolliakou (2004) and others. A variety of exceptions have been pointed out by Manolessou (2000) and Marinis and Panagiotidis (2007), among others.

ticular features that are visible to syntax.[60] This leaves for the interfaces the role of filtering out derivations where movement triggers have been ignored. For the most part, if a derivation has survived syntax and is free of active syntactic features, the interface will consider it well-formed.[61] The interface-based approach, on the other hand, follows Fox (2000) in allowing the interface to rule out an otherwise well-formed derivation because of a better alternative derivation. On Fox's proposal, some kinds of movement have the property of being both *free* and *costly*. Saying that movement is free means that it is untriggered, or at least that a derivation with movement and the derivation without it are close enough to be considered together at the interface. Saying that movement is costly means that a derivation with movement will be dispreferred unless it accomplishes something that could not have been done without movement.[62]

Following Wagner (2007) and Kučerová (2007, 2008), I will assume that one of the interface conditions that can justify movement has to do with information structure. Specifically, I will adopt Kučerová's proposal that moving given material into a givenness domain can justify movement, and that this can be implemented using an operator, *G*, that introduces a givenness presupposition to everything above it within the domain. In our case the domain is the noun phrase, and a relevant configuration could look like this:

(193)



As far as material within the noun phrase is concerned, the configuration in (193) is nothing

[60]Following Cinque, 1999, these features are assumed to be hosted by particular projections within a fine-grained functional hierarchy.

[61]The main exception is possible attempts to reorganize material from earlier stages of the derivation, stated as the *Phase Impenetrability Condition* within Minimalism.

[62]Note that the idea that movement it costly is itself a kind of structural economy, where the simpler, pre-movement form is preferred when possible.

new. The adjective right-dislocates across the noun and attaches right above the original noun phrase, incurring an extra licensor along the way. What is new in (193) is the attachment of the givenness operator *G* right below the intended landing site for the adjective. Earlier we used movement to explain the interaction of word order and the distribution of *def*, but we had no reason to expect movement in the first place. The possibility of attaching *G* gives us a reason to move. If the adjective is given and the noun is not, moving just the adjective to a higher position will create the right configuration for the semantics. Movement across *G* makes it semantically non-vacuous, so the *in situ* variant will not block it. And while movement of the noun across *G* would also be non-vacuous, our assumption that the noun is new means that such movement would result in presupposition failure. I haven't said yet what I mean by saying that the adjective is given or that the noun is new, but we can already take a look at the facts and notice with some relief that when the noun is explicitly mentioned in the context and the adjective is nowhere to be seen, the word order corresponding to (193) is infelicitous:

(194) a. Context: o Janis taise tis gates
the Janis fed the cats

b. # i gates i mikres itan pinasmenes
the cats the small were hungry

c. i mikres i gates itan pinasmenes
the small the cats were hungry

Two other scenarios have been used in the discussion of the discourse effects of polydefiniteness. Here again are Ioannidou and den Dikken's scenarios, based on Kolliakou, along with the judgments provided by the authors. Differently from the two scenarios above, however, speakers seem to differ quite substantially with respect to (195) and (196).

(195) a. Context: o Janis taise ta zoa
the Janis fed the animals

b. i gates i mikres itan pinasmenes
the cats the small were hungry

c. # i mikres i gates itan pinasmenes
the small the cats were hungry

(196) a. Context: o Janis taise ta mikra zoa
the Janis fed the small animals

b. i gates i mikres itan pinasmenes
the cats the small were hungry

c. # i mikres i gates itan pinasmenes
the small the cats were hungry

Bibliography

Abrusán, Márta. 2007. Contradiction and grammar: The case of weak islands. Doctoral Dissertation, MIT, Cambridge, MA.

Alexiadou, Artemis. 2003. Adjective syntax and (the absence of) noun raising in the DP. In *Proceedings of the Workshop on Head-Movement*, 1–39. UCLA.

Alexiadou, Artemis. 2005. Possessors and (in)definiteness. *Lingua* 115:787–819.

Alexiadou, Artemis, and Chris Wilder. 1998a. Adjectival modification and multiple determiners. In *Possessors, predicates and movement in the determiner phrase*, ed. Artemis Alexiadou and Chris Wilder, 303–332. John Benjamins.

Alexiadou, Artemis, and Chris Wilder. 1998b. Adjectival modification and multiple determiners. In *Possessors, predicates and movement in the determiner phrase*, ed. Artemis Alexiadou and Chris Wilder, 305–332. Amsterdam: John Benjamins Publishing Company.

Anderson, Stephen. 1969a. An outline of the phonology of Modern Icelandic vowels. *Foundations of Language* 5:53–72.

Anderson, Stephen. 1969b. West Scandinavian vowel systems and the ordering of phonological rules. Doctoral Dissertation, MIT, Cambridge, Mass.

Anderson, Stephen. 1974. *The organization of phonology*. New York: Academic Press.

Androutsopoulou, Antonia. 1996. The licensing of adjectival modification. In *Proceedings of WCCFL 14*, ed. José Camacho, Lina Choueiri, and Maki Watanabe, 17–31.

Aronoff, Mark. 1976. *Word formation in Generative Grammar*. Cambridge, Mass.: MIT Press.

Arregui-Urbina, Karlos. 2002. Focus on Basque Movements. Doctoral Dissertation, MIT, Cambridge, Mass. URL `http://hdl.handle.net/1721.1/8159`.

Atlas, Jay David, and Stephen Levinson. 1981. It-clefts, informativeness, and logical form. In *Radical pragmatics*, ed. Peter Cole. Academic Press New-York.

Basbøll, Hans. 2005. *The phonology of Danish*. Oxford University Press.

Blutner, Reinhard. 2000. Some aspects of optimality in natural language interpretation. *Journal of Semantics* 17:189–216.

Börjars, Kersti, and Mark Donohue. 2000. Much ado about nothing: Features and zeroes in Germanic noun phrases. *Studia Linguistica* 54:309–353.

Bresnan, Joan. 2001. Explaining morphosyntactic competition. In *Handbook of contemporary syntactic theory*, ed. Mark Baltin and Chris Collins. Oxford: Blackwell.

Carlson, Greg. 1983. Marking constituents. In *Linguistic categories: Auxiliaries and related puzzles*, ed. F. Heny and B. Richards, volume 1, 69–98. Dordrecht: Reidel Publishing.

Chierchia, Gennaro. 2004. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In *Structures and beyond*, ed. A. Belletti. Oxford University Press.

Chierchia, Gennaro, Danny Fox, and Benjamin Spector. 2008. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. Ms., MIT and Harvard, February 2008.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

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

Chomsky, Noam. 1975. Questions of form and interpretation. *Linguistic Analysis* 75–109.

Chomsky, Noam. 1976. Conditions on rules of grammar. *Linguistic Analysis* 2:303–351.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chung, Sandra, and William Ladusaw. 2006. Chamorro evidence for compositional asymmetry. *Natural Language Semantics* 14:325–357.

Chung, Sandra, and William A. Ladusaw. 2004. *Restriction and saturation*. MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. New York: Oxford University Press.

Cohen, Laurence Jonathan. 1971. Some remarks on Grice's views about the logical particles of natural language. In *Pragmatics of natural language*, ed. Yehoshua Bar-Hillel, 50–68. Reidel.

Delsing, Lars-Olof. 1993. On attributive adjectives in scandinavian and other languages. *Studia Linguistica* 47:105–125.

den Dikken, Marcel. 2002. Lexical integrity, checking, and the mirror: A checking approach to syntactic word formation. *Journal of Comparative Germanic Linguistics* 6:169–225.

Einarsson, Stefán. 1945. *Icelandic*. Baltimore and London: Johns Hopkins University Press.

Embick, David. 2007. Blocking effecs and analytic/synthetic alternations. *Natural Language and Linguistic Theory* 25:1–37.

Embick, David, and Alec Marantz. 2008. Architecture and Blocking. Ms., UPenn and MIT/NYU, to appear in Linguistic Inquiry 39:1.

Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.

Fanselow, Gisbert, and Damir Ćavar. 2002. Distributed deletion. In *Theoretical approaches to universals*, ed. Artemis Alexiadou, 65–107. Amsterdam: Benjamins.

von Fintel, Kai. 1994. Restrictions on quantifier domains. Doctoral Dissertation, University of Massachusetts-Amherst.

von Fintel, Kai, and Danny Fox. 2002. Classnotes for 24:954: Pragmatics in linguistic theory. DSpace. URL `http://hdl.handle.net/1721.1/36355`.

Fox, Danny. 2000. *Economy and semantic interpretation*. MIT Press.

Fox, Danny. 2006. Free choice and the theory of scalar implicatures. Ms., MIT.

Fox, Danny, and Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29:537–586.

Gajewski, Jon. 2002. On analyticity in natural language. Ms., MIT.

Gamut, L. T. F. 1991. *Logic, language, and meaning*. Chicago University Press.

Gazdar, Gerald. 1979. *Pragmatics: Implicature, presupposition and logical form*. Academic Press New-York.

Geurts, Bart. 2007. Scalar implicature and local pragmatics. URL `http://www.ru.nl/ncs/bart/papers/globallocal.pdf`, ms., University of Nijmegen, May 2007.

Gouskova, Maria. 2001. Split scrambling: Barriers as violable constraints. In *Proceedings of HUMDRUM 2000*.

Grice, Paul. 1967/1975. Logic and conversation. In *Speech acts*, ed. P. Cole and J. L. Morgan, volume 3 of *Syntax and Semantics*, 41–58. New-York Academic Press.

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

Groenendijk, Jeroen, and Martin Stokhof. 1984. On the semantics of questions and the pragmatics of answers. In *Varieties of formal semantics*, ed. Fred Landman and Frank Veltman, volume 3 of *Groningen-Amsterdam Studies in Semantics*, 143–170. Dordrecht: Foris.

Halle, Morris. 1997. Distributed Morphology: Impoverishment and fission. In *Papers at the interface*, ed. B. Bruening, Y. Kang, and M. McGinnis, 125–149. Cambridge, Mass.: MIT Working Papers in Linguistics.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from Building 20*, ed. Kenneth Hale and Jay Keyser, 111–176. MIT Press.

Hamblin, C. L. 1973. Questions in Montague English. *Foundations of Language* 10:41–53.

Hankamer, Jorge, and Line Mikkelsen. 2002. A morphological analysis of definite nouns in Danish. *Journal of Germanic Linguistics* 14:137–175.

Hankamer, Jorge, and Line Mikkelsen. 2005. When movement must be blocked: A response to Embick and Noyer. *Linguistic Inquiry* 36:85–125.

Hawkins, John A. 1991. On (in) definite articles: Implicatures and (un) grammaticality prediction. *Journal of Linguistics* 27:405–442.

Heim, Irene. 1991. Artikel und Definitheit. In *Semantik: Ein internationales handbuch der zeitgenössischen forschung*, ed. A. von Stechow and D. Wunderlich. Berlin: De Gruyter.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA: Blackwell.

Hirschberg, Julia. 1985/1991. *A theory of scalar implicature*. New York: Garland. PhD Thesis, University of Pennsylvania.

Horn, Laurence. 1969. A presuppositional analysis of only and even. In *CLS*, volume 5, 97–108.

Horn, Laurence. 1972. On the semantic properties of the logical operators in English. Doctoral Dissertation, UCLA.

Horn, Laurence. 1978. Lexical incorporation, implicature, and the least effort hypothesis. In *Papers from the Parasession on the Lexicon*, 196–209. Chicago: Chicago Linguistics Society.

Horn, Laurence. 1984. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicatures. In *Meaning, form, and use in context*, ed. D. Schiffrin, 11–42. Washington: Georgetown University Press.

Horn, Laurence. 1989. *A natural history of negation*. Chicago: University of Chicago Press.

Horn, Laurence. 2000. From IF to IFF: Conditional perfection as pragmatic strengthening. *Journal of Pragmatics* 32:289–326.

Householder, Frederick. 1971. *Linguistic speculations*. Cambridge University Press.

Hulsey, Sarah, and Uli Sauerland. 2006. Sorting out relative clauses. *Natural Language Semantics* 14:111–137.

Ioannidou, Alexandra, and Marcel den Dikken. 2006. P-Drop, D-Drop, D-Spread. Ms., CUNY, April 2006.

Ioannidou, Alexandra, and Marcel den Dikken. 2007. P-Drop, D-Drop, D-Spread. To appear with MITWPL, November 2007.

Julien, Marit. 2002. Determiners and word order in Scandinavian DPs. *Studia Linguistica* 56:264–314.

Kayne, Richard. 1994. *The antisymmetry of syntax.* Cambridge, MA: MIT Press.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar.* Blackwell.

Kiparsky, Paul. 1973. "elsewhere" in phonology. In *A festschrift for Morris Halle*, ed. Stephen Anderson and Paul Kiparsky, 93–106. Holt, Rinehart and Winston, Inc.

Kiparsky, Paul. 1983. Word-formation and the lexicon. In *Proceedings of the 1982 Mid-America Linguistic Conference.*

Kishimoto, Hideki. 2000. Indefinite pronouns and overt N-raising. *Linguistic Inquiry* 31:557–566.

Kolliakou, Dimitra. 2004. Monadic definites and polydefinites: Their form, meaning and use. *Journal of Linguistics* 40:263–323.

Koopman, Hilda Judith, and Anna Szabolcsi. 2000. *Verbal complexes.* Cambridge, Mass.: MIT Press.

Kroch, Anthony. 1972. Lexical and inferred meanings for some time adverbials. Quarterly Progress Reports of the Research Laboratory of Electronics 104, MIT, Cambridge, Mass.

Kučerová, Ivona. 2007. The syntax of givenness. Doctoral Dissertation, MIT, Cambridge, Mass.

Kučerová, Ivona. 2008. Givenness and maximize presupposition. To appear in Proceedings of Sinn und Bedeutung 12.

Landman, Fred. 2000. *Events and plurality: The Jerusalem lectures.* Dordrecht: Kluwer Academic Publishers.

Larson, Richard, and Franc Marušič. 2004. On indefinite pronoun structures with APs: Reply to Kishimoto. *Linguistic Inquiry* 35:268–287.

Lekakou, Marika, and Kriszta Szendrői. 2007. Eliding the noun in close apposition, or Greek polydefinites revisited. In *Ucl working papers in linguistics*, ed. Richard Breheny and Nikolaos Velegrakis, volume 19, 129–154. UCL.

Leu, Thomas. 2007. From Greek to Germanic: Poly-(*in)-definiteness and weak/strong adjectival inflection. URL http://ling.auf.net/lingBuzz/000392, ms., NYU, January 2007.

Levinson, Stephen. 2000. *Presumptive meanings*. Cambridge, Mass.: The MIT Press.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretic approach. In *Meaning, use, and interpretation of language*, ed. R. Bäuerle, C. Schwarze, and A. von Stechow, 303–323. De Gruyter.

Magri, Giorgio. 2006. The blindness hypothesis and individual level predicates. In *Proceedings of SALT XVI*, ed. M. Gibson and J. Howell, 102–119. Ithaca, NY.

Manolessou, Ioanna. 2000. Greek noun phrase structure: A study in syntactic evolution. Doctoral Dissertation, University of Cambridge, Cambridge, UK.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st Annual Penn Linguistics Colloquium*, ed. A. Dimitriadis, L. Siegel, C. Surek-Clark, and A. Williams, volume 4.2 of *UPenn Working Papers in Linguistics*, 201–225.

Marinis, Theodoros, and Phoevos Panagiotidis. 2007. Determiner spreading as DP-predication. Ms., University of Reading and Cyprus College, Under review.

Matsumoto, Yo. 1995. The conversational condition on Horn Scales. *Linguistics and Philosophy* 18:21–60.

May, Robert. 1985. *Logical form: Its structure and derivation*. Cambridge, Mass.: MIT Press.

McCawley, James D. 1978. Conversational implicatures and the lexicon. In *Syntax and semantics*, ed. Peter Cole, volume 9: Pragmatics, 245–259. New York: Academic Press.

Moortgat, Michael. 1987. Compositionality and the syntax of words. In *Foundations of pragmatics and lexical semantics*, ed. J. Groenendijk, D. de Jongh, and M. Stokhof, 41–62. Foris, Dordrecht.

Noyer, Rolf. 1992. Features, positions and affixes in autonomous morphological structure. Doctoral Dissertation, MIT, Cambridge, MA.

Orešnik, Janez. 1972. On the epenthesis rule in Modern Icelandic. *Arkiv för nordisk filologi* 87:1–32.

Parikh, Prashant. 2000. Communication, meaning, and interpretation. *Linguistics and Philosophy* 23:185–212.

Partee, Barbara. 1973. Some transformational extensions of Montague Grammar. *Journal of Philosophical Logic* 2:509–534.

Pesetsky, David. 1979. Russian morphology and lexical theory. URL http://web.mit.edu/linguistics/people/faculty/pesetsky/russmorph.pdf, ms., MIT.

Pesetsky, David. 1985. Morphology and logical form. *Linguistic Inquiry* 16:193–248.

Poser, William J. 1992. Blocking of phrasal constructions by lexical items. In *Lexical matters*, ed. Ivan Sag and Anna Szabolsci, 111–130. Stanford, CA: CSLI Publications.

Prince, Alan, and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Technical report, Rutgers University, Center for Cognitive Science.

van Rooij, Robert. 2004. Signalling games select Horn strategies. *Linguistics and Philosophy* 27:493–527.

van Rooij, Robert, and Katrin Schulz. 2004. Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information* 13:491–519.

Rooth, Mats. 1985. Association with focus. Doctoral Dissertation, University of Massachusetts, Amherst.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.

Russell, Benjamin. 2006. Against grammatical computation of scalar implicatures. *Journal of Semantics* 23:361–382.

Sauerland, Uli. 2002. The present tense is vacuous. *Snippets* 6:12–13. URL http://www.ledonline.it/snippets/allegati/snippets6004.pdf.

Sauerland, Uli. 2003a. Implicated presuppositions. URL http://www.zas.gwz-berlin.de/mitarb/homepage/sauerland/milanoHandout.pdf, handout of talk presented at the University of Milan Bicocca, Milan, Italy, June 2003.

Sauerland, Uli. 2003b. A new semantics for number. In *Proceedings of SALT 13*, 258–275. Cornell University, Ithaca, NY: CLC Publications.

Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27:367–391.

Sevi, Aldo. 2005. Exhaustivity: A semantic account of 'quantity' implicatures. Doctoral Dissertation, Tel-Aviv University.

Sharvy, Richard. 1980. A more general theory of definite descriptions. *The Philosophical Review* 89:607–624.

Soames, Scott. 1982. How presuppositions are inherited: A solution to the projection problem. *Linguistic Inquiry* 13:483–545.

Spector, Benjamin. 2006a. Aspects de la pragmatique des opérateurs logiques. Doctoral Dissertation, Université de Paris 7, Paris.

Spector, Benjamin. 2006b. Aspects of the pragmatics of plural morphology: On higher-order implicatures. URL http://www.cognition.ens.fr/ bspector/SemNet$_S$pector$_F$inal.pdf, ms., Ecole Normale Supérieure.

Spector, Benjamin. 2007. Scalar implicatures: Exhaustivity and Gricean reasoning. In *Questions in dynamic semantics*, ed. Maria Aloni, Alistair Butler, and Paul Dekker, 229–254. Elsevier.

Sproat, Richard, and Chilin Shih. 1988. Prenominal adjectival ordering in English and Mandarin. In *Proceedings of NELS 18*, 465–489.

Svenonius, Peter. 1993. Selection, adjunction, and concord in the DP. *Studia Linguistica* 47:198–220.

Svenonius, Peter. 1994. The structural location of the attributive adjective. In *Proceedings of the West Coast Conference on Formal Linguistics*, ed. Erin Duncan, Donka Farkas, and Philip Spaelti, volume 13, 439–454.

Wagner, Michael. 2007. Givenness and locality. In *Proceedings of SALT 16*, ed. Jonathan Howell and Masayuki Gibson, 295–312. Ithaca, NY: CLC Publications. URL http://hdl.handle.net/1813/7594.

Westerstahl, Dag. 1984. Determiners and context sets. In *Generalized quantifiers in natural language*, ed. Johan van Benthem and Alice ter Meulen, 45–71. Dordrecht: Foris.

Williams, Edwin. 1981. Argument structure and morphology. *The Linguistic Review* 81–114.

Zipf, George Kingsley. 1949. *Human behaviour and the principle of least effort*. Cambridge, MA: Addison-Wesley.