Prosody and Recursion

by

Michael Wagner

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 2005

Author ....
Department of Linguistics and Philosophy
September 1, 2005

Certified by ....
Alec Marantz
Professor of Linguistics
Thesis Supervisor

Certified by ....
Donca Steriade
Professor of Linguistics
Thesis Supervisor

Accepted by ....
Irene Heim
Chair of the Linguistics Program

Prosody and Recursion

by

Michael Wagner

Submitted to the Department of Linguistics and Philosophy
on September 1, 2005, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis proposes a recursive mapping of syntactic derivations to prosodic representations. I argue that the prosody of an expression, just like its meaning, is determined compositionally, as originally proposed in Chomsky et al. (1957), Chomsky and Halle (1968). Syntactic structure are cyclically spelled out and assigned a semantic and phonological interpretation. The cyclic approach is motivated based on data from the prosody of coordinate structures, integrating insights from syntax, combinatorics, and semantics.

The algorithm distinguishes two ways of prosodically relating the output of cyclic domains: they can either be mapped to prosodic domains that are on a par and match in prosodic status: PROSODIC MATCHING; or the output of one cycle can be prosodically subordinated to another cycle: PROSODIC SUBORDINATION. Together, they derive a metrical structure that encodes information about phrasing, accent placement, and prominence. Scope relations, argument structure, and information structure affect prosodic phrasing indirectly by determining which of the two principles applies and when a syntactic cycle is spelled out.

The derived metrical representation is a relational grid (Liberman, 1975). It encodes syntactic structure and also the derivational history of how it was assembled. The theory attempts to incorporate insights from recent work on stress and prominence (Cinque, 1993, Arregi, 2002) and prosodic recursion Ladd (1988), Dresher (1994), as well as insights from the research on prosodic phrasing and phrasal phonology (Gussenhoven, 1984, Selkirk, 1986, Truckenbrodt, 1995). Phonetic evidence from on-line production is presented to show that speakers implement the predicted metrical relations and scale boundaries later in the utterance relative to boundaries already produced, a phenomenon dubbed BOUNDARY STRENGTH SCALING.

Thesis Supervisor: Alec Marantz
Title: Professor of Linguistics

Thesis Supervisor: Donca Steriade
Title: Professor of Linguistics

"There is nothing more vivifying than a hypothesis."

Primo Levi, The Periodic Table

In memory of Mirco Ghini,

and his contagious passion for linguistics.

Acknowledgments

Over the last five years I experienced every day that the recommendation that Heinrich von Kleist gives at the beginning of his essay 'Concerning the gradual formulation of thoughts while speaking' was a good one: He recommends to us that if we want to understand something and just cannot figure it out to start talking about it with others, even if the thoughts are still inarticulate. I am fortunate in having had the opportunity to talk to the teachers, students, and visitors here at MIT during this period, it was a great privilege. There are many people that I would like to thank for their support, and probably more than I can think of right now.

First my advisors, Alec Marantz and Donca Steriade, who had the generosity to help me whenever I needed with their own ideas, their time, their encouragement, and their comments and criticisms; and also to let me make my own mistakes when I insisted on making them. They were great advisors, also in preventing me from going down many garden paths but never trying to push me down well-trodden or personally favored ones. They let me make up my own mind and develop ideas at my own pace.

I would also like to thank Kai von Fintel, Danny Fox, Irene Heim and David Pesetsky, who have always had an open door and have been a constant source of support and inspiration over the last years. They have supported me working on various projects over the time and encouraged various expeditions into the world outside phonology, which have left quite an impact on this dissertation. Irene agreed to be on my committee when nobody including me was quite sure what the thesis was going to look like and has been helpful all along.

This thesis would have looked completely different had I not discussed aspects of it on many, many occasions with Karlos Arregi, Jon Gajewski, Florian Jaeger, Aditi Lahiri and Hubert Truckenbrodt, all of whom have taught me a lot.

I am also indebted to Asaf Bachrach, Andrew Nevins, and Andres Salanova whose ideas on phonology and its relation to other domains have been an inspiration; and to Heejeong Ko and Joey Sabbagh for their syntactic insights.

Adam Albright, Sabine Iatridou, Michael Kenstowicz, Shigeru Miyagawa and Norvin Richards have been very helpful when I met with them about various issues at various points over the years. And Morris Halle has been very supportive all along, especially when I was lost among syntactic trees and didn't see the phonological forest.

Edward Flemming, Lisa Selkirk and Elan Dresher agreed to be on my committee despite the fact that it was rather late in the game, and have shared their expertise , their encouragement, and also their incisive criticism and different perspectives. It was fun and I learned a lot through this process, also about what I still have to learn.

Various people working on prosody and intonation have been very kind and helpful, especially David Adger, Daniel Büring, Abby Cohn, Caroline Féry, Bill Idsardi, John Kingston, Bob Ladd, Ted Gibson, Carlos Gussenhoven, Dan Jurafsky, Aritz Irurtzun, Shin Ishihara, Arsalan Kahnemuyipour, Manfred Krifka, Ad Neeleman, Jon Nissenbaum, Colin Phillips, Uli Sauerland, Stephanie Shattuck-Hufnagel, Kriszta Szendröi, Haruo Kubozono, Roger Schwarzschild and Draga Zec.

Cedric Boeckx, Gerlof Bouma, Julia Horvath, Richard Larson, Alan Munn, Charles Reiss, Barry Schein, Tony Kroch, Bob Franks, Tania Reinhart and Anna Szabolcsi have given me helpful feedback during the last months on some of the issues I was working on. Especially discussions with Barry Schein especially helped me a lot in developing some of the ideas in chapter 2.

In spring 2004 I had the opportunity to attend a seminar on intonation taught by Angelika Kratzer and Lisa Selkirk in Amherst which was very helpful; this also gave me the no less instructive opportunity to share rides with Luis Alonso-Ovalle. During this time and after, Chris Potts and Hubert Haider have been a great source of inspiration and productive feedback.

In 2003 I spent some time in Stanford, which was a great experience in many ways, thanks to Lev Blumenfeld, David Beaver, John Beavers, Cleo Condoravdi, Liz Coppock, Philip Hofmeister, Florian Jaeger, Paul Kiparsky, Beth Levin, Judith Tonhauser, and many others.

I am very indebted to the EGG Summar Schools, and everything I learned there.

Thanks to Klaus Abels, Boban Arrsenijević, Dea Garraffa, Nino Grillo, Aritz Irurtzun, Luisa Martí, Øystein Nilsen, Gillian Ramchand, Milan Rezac, Curt Rice, Michal Starke, Peter Svenonius, Irene Sanoudaki, Tobias Scheer, Hedde Zeijlstra and many others; and also to the experience at the Winter school in Campinas in 2004.

In the year before coming to MIT, I was in the good hands of the Linguistics department at Konstanz, who have been very hospitable in the time since, thanks to Melanie Hochstätter, Elisabeth Jönssen Steiner, Achim Kleinmann, Astrid Kraehenmann, Frans Plank, Henning Reetz, Dirk Saleschius, Mathias Scharinger, Thomas Schöneborn, Allison Wetterlin, and Frank Zimmerer. I am especially indebted to Aditi Lahiri, who was incredibly supportive over the entire time.

I also want to thank Keren Rice and Diane Massam and the department of UofT for somehow convincing me of linguistics, and several people in Toronto who have since been or become friends, Claude Desmarais, Chiara Frigeni, Maria Gentle and Sara Mackenzie, Abdel-Khalig Ali, and Taylor Roberts.

Thanks to David Adger and Anson, Suzana Bejar, Cristina Cuervo, Daniel Harbour, Stephanie Harves and Olivier Pauluis, and Toby Colding for helping me through first year and their friendship ever since.

I also want to thank Artemis Alexiadou, Arto Anttila, Elena Anagnastopoulou, Sigrid Beck, David Braun, Lisa Brunetti, Andreas Bulk, Nagore Calleja, Christopher Collins, Aniko Csirmaz, Gisbert Fanselow, Justin Fitzpatrick, Itamar Francez, Werner Frey, Marting Gärtner, Andreas Haida, Silke Hamann, Brent Henderson, Roland Hinterhölzl, James Isaacs, Junko Ito, Nina Kazanina, Nathan Klinedienst, Martin Krämer, Armin Mester, John McCarthy , Victor Manfredi, Gereon Müller, Winnie Lechner, Tom Leu, Pilar Prieto, Liina Pylkkänen, Tomas Riad, Henk van Riemsdijk, Tamina Stevenson, Tim Stowell, Linnaea Stockall, Shoichi Takahashi, Luis Vicente, Bert Vaux, Mary-Ann Walters, Susi Wurmbrand, Sabine Zerbian, Malte Zimmermann for help and comments at various points.

I am lucky to have had friends in Cambridge that kept me (sort of) sane, and helped me in work and play: Marta Abrusan, Luis Alonso-Otovalle. Pranav Anand, Rajesh Bhatt, Asaf Bachrach, Jon Gajewski, Elena Guerzoni and Martin Hackl, Andrea Gualmini, Valentine Hacquard, Sarah Hulsey, Ivona Kucerova, Luisa Meroni, Andrew Nevins, Alex Perovic, Amelia Reis Silva, Joey Sabbagh, Andres Salanova, Amelia Reis Silva and Vina Tsakali. I also want to thank Effi Georgala for her constant friendship and help in the last years, and also many other friends here and abroad who have been faithful despite being so neglected by me at times.

I wish I could thank at this point my godfather Reinhart Wagner†, who has supported me all my life, and who I miss even more in this moment; and I want to thank his family, who have been there for me since I can think. The final thanks to go to my (extended) family, Wilbert, Thomas, Christoph, and my mother and their unconditional love.

Contents

Chapter 1

Introduction

This dissertation is about the relation between syntax and prosody. The challenge in this area of research is that not only are there many unresolved questions about the phonological and syntactic facts themselves, it is also an issue where to draw the line between phonology and syntax. While probably every generative linguist will agree that the feature [coronal] is not a syntactic one but the feature [1st person] is, the domains are otherwise very intricately intertwined in ways that are ill understood.

The measure of success for a theory of the syntax-phonology mapping is whether it turns syntactic facts into phonological facts and phonological facts into syntactic facts, in the following sense: the theory should be specific enough to make inferences about the syntactic structure by looking at its prosody, and conversely allow inferences about prosody by looking at the syntactic derivation.

This is the standard of how to deal with facts about linear precedence, one of the main sources of evidence in syntactic research. Linear precedence itself is a phonological notion, and yet it can be used as evidence in syntactic argumentation. The reason is that there are explicit assumptions that allow inferences about syntactic constituency and hierarchical relations (e.g. c-command) based on linear precedence.[1]

[1]An example of a mapping theory is the LCA, in Kayne (1994) and subsequent work in the antisymmetric framework; other work on syntax–linear order include current versions of OT syntax (e.g. Sells, 2001), representation theory (Williams, 2003), or the theory of cyclic linearization in Fox and Pesetsky (2005), Ko (2005), Sabbagh (to appear), Takahashi (2005).

In current research on syntax, prosodic structure does not have the same status as linear order as a source of evidence. Only occasionally are prosodic facts used in syntactic arguments. The reason is that the relation between syntax and prosody is less well understood, and inferences about syntax just based on prosodic evidence are often not possible. The goal of this dissertation is to provide a mapping function that allows such inferences. The argument is based on a limited set of data from a very limited set of languages, and it remains to be seen how well it extends beyond.

The presented algorithm assumes that prosody is compositional, analogous to Fregean compositionality in semantics. Grammar assembles complex expressions, and assigns both a semantic interpretation (a denotation) and a phonological interpretation (among other things a metrical structure) by compositionally combining the denotations and the phonological representations of its parts.

The idea of a compositional theory of prosody goes back to the transformational cycle in Chomsky et al. (1957) and Chomsky and Halle (1968). The algorithm presented in this thesis is a version of the transformational cycle. The precise assumptions made about syntax and about phonology differ of course from the ones made in SPE. One crucial difference is that I incorporate an idea from Bresnan (1971) and Chomsky (2001) who propose to intersperse the phonological cycle with the syntactic cycle. Grammar assembles atomic elements into syntactic trees.

The process of assembling a phrase marker in a syntactic workspace and its subsequent spell-out is called a 'cycle'. An output of a cycle is an output of the grammar. It receives a semantic and a phonological interpretation. The mapping principle, presented in chapter 2, has the effect that the elements within a cycle are mapped to a single prosodic domain in a metrical grid upon spell-out.

After spell-out, a cycle can enter a new cycle as an atom. This is the recursive step, that lies at the heart of semantic and prosodic compositionality. This recursion together with a simple mapping principle derive the prosody of intricate prosodic patterns in coordinate structures.

A crucial question is how grammar decides whether to continue building structure in the same cycle, or whether to spell-out and start a new cycle.

Coordinate structures sometimes receive a 'flat' prosody, in which each conjunct is prosodically on a par, and sometimes a more articulated prosody that reveals the syntactic bracketing as either left- or right-branching. The bracketing is reflected in the strength of boundaries that follow conjuncts. I represent the strength of boundaries by the number of pipe symbols (| , || , ||):

(1) a. $A \lor B \lor C$ 'Flat' Prosody: A | ∨ B | ∨ C

b. $A \lor (B \land C)$ 'Articulated' Prosody: A || ∨ B | ∧ C

c. $(A \lor B) \land C$ 'Articulated' Prosody: A | ∨ B || ∧ C

In chapter 2 I present a derivational theory of syntax—phonology mapping that accounts for the prosody of the fragment of English consisting only of proper names and the functors 'and'/'or'. The generalization about when grammar decides to close off a cycle is the same as the generalization about when in logical notation one closes off an expression by parentheses: when the associative law does not hold. Associativity of composition is a necessary property for elements to form a single cycle.

(2) Associative Domains

Cycles are assocative domains.

Since after each cycle spell-out maps the output to a prosodic foot in a new top grid line, the prosody reflects the way the structures are put together by syntax. A flat prosody is assigned exactly in associative domains, since they are built in a single cycle. A more articulate prosody results in cases where the associative law does not hold, and the prosody reflects the syntactic bracketing.

The grammar of coordinate structures involving proper names and 'and'/'or' is used as a baseline for the investigation of other types of constructions. In chapter 3 I present evidence that the analysis carries over to other domains of the grammar.

The prosodic structure derived by the algorithm is a metrical grid that is purely relational. Prosodic domains are differentiated by whether or not they occur on a higher or lower line, but not by their category (e.g. intonational phrase, phonological phrase). This relational notion of a 'metrical grid' was first introduced in generative grammar in Liberman (1975, 73): "The Grid, like time, is in principle infinite – we can pick any place in it to start, and have available sequences forward and backward, or up and down in the hierarchy, as we choose."

This relational notion of the metrical grid is also assumed in the approaches to sentence stress in Selkirk (1984), Halle and Vergnaud (1987), Cinque (1993), Arregi (2002) and Ishihara (2003). It contrasts with the assumption of a fixed and non-recursive prosodic hierarchy Selkirk (1978, et seq.). Recursion in prosody was also argued for in Ladd (1986), Ladd (1988), Kubozono (1989), and Dresher (1994). More recently, non-recursiveness has been proposed to be a violable constraint (Selkirk, 1995b, Truckenbrodt, 1995) even within the framework of the prosodic hierarchy theory.

In chapter 4 I report evidence form phonetics and phonology for a more recursive structure. The proposal in this thesis is compatible with a differentiation of prosodic categories, but it is not compatible with tieing these prosodic categories directly to certain syntactic categories. This is discussed in detail in chapter 4.

In chapter 5, I report experimental evidence that prosodic domains are scaled relative to earlier produced domains, which fits the relational understanding of metrical structure. The evidence comes from the adjustment of boundary strength as a reflex of the hierarchical organization of metrical structure. I will refer to this phenomenon as 'Boundary Strength Scaling'.

The second part of the thesis address cases of 'prosodic subordination'. Under certain circumstances, constituents become heavily pitch-reduced if not altogether deaccented. Prosodic subordination is at the heart of the generalization about nuclear stress. It is *the last unsubordinated constituent* that is perceived with main prominence in any complex phrase marker.

In chapter 6, I look at the conditions under which heads are subordinated. The generalization is that they are obligatorily subordinated whenever they *follow* they complements. This is easily observable in the prosodic difference between OV and VO structures. An asymmetry parallel to the one observed for VO *vs.* OV can be observed in a wide variety of other cases involving functors and complements. I propose the principle of PROSODIC SUBORDINATION in order to capture this generalization. Again, this principle is recursively applied in order to derive the final metrical structure.

Further evidence discussed in chapter 6 involve predicate clusters, from function words, affixation, and paratactic constructions, each illustrating the same asymmetry. Moreover, it accounts for a parallel in the prosody of predication and modification structures. Both predicates and modifiers are functors, and they show the same prosodic asymmetry.

The recursive approach to syntax—prosody mapping allows for a unified explanation. Since all cases involve functors that either precede or follow their complement , they are expected to display the same asymmetry.

In chapter 7, I present evidence that the same principle is at play in givenness marking and contrast. Imagine in the 'OV' derivation above, we replace the object 'O' with 'A and B', and replace 'V' with 'or C'. The metrical output of the derivation leads to a structure in which the last conjunct is not accented:

(3) Á and B́ or C

This prosody is not the neutral one. It can only be used in certain contexts, that satisfy a givenness presupposition. A context in which constituent C is deaccented is the following:

(4) Did you say: D́ or Ć?
No! (Á and B́) or C

The effects of givenness and contrast can be captured by the notion of RELATIVE GIVENNESS. The main result is that givenness marking always involves marking a constituent as given *relative to its sister*, in way that is made explicit using alternative semantics. This insight helps resolve several problems with earlier accounts and makes several new predictions.

The prosodic side of givenness marking surprisingly shows the same asymmetry observed in the case of the functor-argument asymmetry, and suggest that again the principles of PROSODIC MATCHING and PROSODIC SUBORDINATION are at work. The functor-argument asymmetry and RELATIVE GIVENNESS merely affect which principle to apply at each recursive step.

The result is an simple recursive system that makes decisions on prosody just based on looking at two sisters at a time—a version of the transformational cycle from SPE.

Part I

Prosodic Matching

Chapter 2

Prosody and Recursion in Coordinate Structures

Abstract

Based on evidence from the grammar of coordinate structures in English, this chapter proposes a compositional theory of prosody. A prosodic representation is assigned to complex expressions by recursively combining the prosodies of its parts, analogous to the compositionality assumed in semantics. A single mapping principle, recursively applied, captures an intricate set of metrical patterns. The system cyclically derives prosodic representations, integrating insights from the semantics, combinatorics, and syntax of coordinate structures. It is essentially an implementation of the 'transformational cycle' proposed in SPE. The new proposal, however, attempts to capture not only prominence relations—as the algorithm in SPE did—but also prosodic grouping and the strength of prosodic boundaries.

2.1 Prosodic Compositionality

Grammar is recursive. It combines basic elements to create complex expressions, and each output of the grammar can in turn be used as a building block in a new and even more complex expression. One necessary ingredient of a recursive grammar is a compositional semantics. It enables the grammar to assign a denotation to each output, by combining the denotations of the assembled parts. A second precondition is that the grammar must be able to specify how to *pronounce* a new output.

2.1.1 Coordinate Structures

Coordinate structures provide a simple illustration of the recursiveness of grammar. Two or more constituents can be combined to form a new constituent by connective functors such as 'and' and 'or'. Each conjunct can itself be a coordinate structure.

(1) a. Lysander or Demetrius.

b. Hermia and (Lysander or Demetrius).

c. Helena or (Hermia and (Lysander or Demetrius)).

Different bracketings of coordinate structures can have different truth-conditions. In orthography, parentheses can be used in order to avoid ambiguity. Consider the following two statements, which differ only in the bracketing of the expression to the right to the equivalence symbol. While (2a) is a valid tautology, (2b) is false.[1]

(2) a. $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$ "...p and q or p and r."

b. $p \wedge (q \vee r) \equiv ((p \wedge q) \vee p) \wedge r$ "...p and q or p and r."

A standard assumption in generative grammar is that the denotation of complex constituents is determined compositionally: the denotation of complex expressions is determined based on the denotation of its parts. Compositionality, first introduced by Frege, makes it possible to recursively build up new meaningful expressions drawing on a set of basic atoms. An additional assumption, made i.a. in Heim and Kratzer

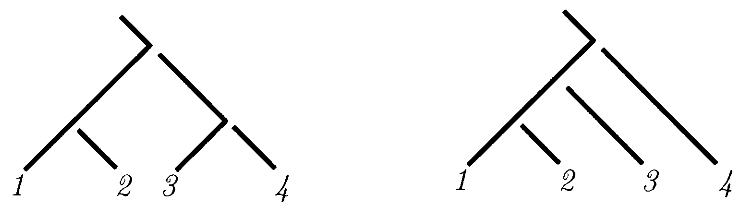
[1]Thanks to Michael Kenstowicz for suggesting logical formulas as an illustration.

(1998, 29), is that compositionality is local. The denotation of any node in the syntactic structure is determined by combining the denotations of its daughters.[2]

The truth conditions computed by this compositional system tell us that stating (2a) is true and (2b) is false, although they consist of the same string of words. The difference in meaning is due to the difference in syntactic structure:[3]

(3) Syntactic Bracketing

a. $(p \wedge q) \vee (p \wedge r)$ b. $((p \wedge q) \vee p) \wedge r$



Along with assigning a meaning, the grammar also assigns a prosodic structure to its outputs. Just as the composition of the meaning of an expression reflects the recursive internal structure of the linguistic object, so does the prosody. Prosody disambiguates spoken renditions of (2a) and (2b).

Prosodic bracketing is reflected among other things by lengthening effects. These lengthening effects in coordinate structures have been investigated instrumentally e.g. in Lehiste (1973). Similar bracketing asymmetries were also used in the study on algebraic formulas in Lehiste et al. (1976). Prosodically induced lengthening effects were also investigated in Cooper and Paccia-Cooper (1980) and many subsequent studies, which are discussed in detail in chapter 4. Lengthening effects are illustrated in the following oscillograms. Consider in particular the length of 'q' and the length

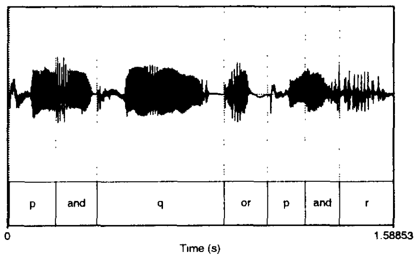
[2]This strict locality assumption is called 'context-free semantics' in Dowty (2003). The treatment of variable binding using assignment functions in Heim and Kratzer (1998) departs from strict locality. Jacobson (1999) proposes a notion of 'direct compositionality' that adheres to strict locality.

[3]The expression ((p and q) or p) seems redundant since logically it is equivalent to 'p' alone. But in natural language, disjuncts can be interpreted exhaustively, so the expression can mean ((p and q) or just p). This possibility of an exhaustive reading of the disjuncts is, I believe, responsible for the fact that disjunction can have an exclusive and an inclusive reading, which was shown by (Hurford, 1974). While I cannot motivate this claim at this point, I will stick with the example simply for the sake of illustration.

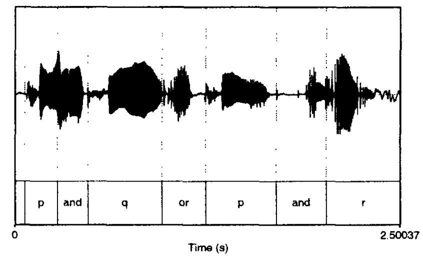
of the second ‘p’:

(4) Boundary Strength and Lengthening Effects

a. $(p \wedge q) \vee (p \wedge r)$



b. $((p \wedge q) \vee p) \wedge r$



The lengthening reflects the prosodic grouping of the expression, which in turn reflects the syntactic/semantic grouping. One key difference between the two structures is that in (2a), the boundary after ‘q’ is intuitively stronger than the boundary after the second ‘p’; the converse is true for (2b), where the boundary after ‘q’ is intuitively weaker than that after the second ‘p’.

I will indicate prosodic boundaries with the pipe symbol ‘| ’. The strength of a boundary can be conceived of as a rank of the boundary on a discrete strength scale of boundary strengths. The boundary rank is encoded in the number of pipes (‘| ’, ‘|| ’, ‘||| ’...). The boundary ranks reported in this chapter are based on native speakers’ intuitions. For the basic coordination cases they were also experimentally tested, as reported in chapter 5. The two structures above can then be represented as follows:[4]

(5) a. $|| \; p \; | \; \wedge \; q \; || \; \vee p \; | \; \wedge \; r \; ||$

b. $||| \; p \; | \; \wedge \; q \; || \; \vee p \; ||| \; \wedge \; r \; |||$

The strength of the boundaries directly reflect the syntactic constituency. Adjacent material separated from each other by weak boundaries and separated from the surroundings by stronger boundaries form constituent.

A similar notation is in fact used in Chomsky et al. (1957), where boundary symbols are hierarchically organized, and the organization is reflected by an index that

[4]In this chapter, I will not address the strength of the boundary between the connector and the following conjunct. This will be discussed in chapter 5.

reflects boundary rank. The representation for our example would be, where lower numbers reflect stronger boundaries:

(6) Syntactic Structure in Chomsky et al. (1957)

a. $\#_0\ p\ \#_2\ \wedge\ q\#_1\ \vee p\ \#_2 \wedge\ r\#_0$

b. $\#_0\ p\ \#_3\ \wedge\ q\#_2\ \vee p\ \#_1 \wedge\ r\#_0$

The phonetic lengthening effects illustrated in (4) can be made sense of on the assumption that they are a reflection of the boundaries in (5) and that stronger boundaries induce a higher degree of lengthening of the constituent preceding it compared to weaker boundaries.

The boundary ranks reported here reflect that the conjuncts/disjuncts connected by 'and' or 'or' are prosodically on a par, i.e. the perceived boundaries between the conjuncts/disjuncts are of equal strength. I will take this as evidence they form prosodic constituents of the same type. They will be presented as forming feet at the same level in metrical grid. A prosodic domain that contains several units that are prosodically 'on a par' in this sense will be henceforth called 'prosodically flat'.

In (5a), there are two main disjuncts, and 'p and q' is prosodically on a par with 'p and r'. Within each disjunct, there are two conjuncts separated by weaker boundaries, which again are on a par with respect to each other. In (5b), there are two main conjuncts, and the first, 'p and q or p', is prosodically on a par with the second, 'and r'.

An obvious question is whether or not the boundaries have to represented somehow in the prosodic structure or not. If the boundaries simply reflect syntactic branching it would seem redundant to code them in a metrical representation. A different way of looking at metrical representations however is that they are a *different way of representing syntactic information*—a way that is legible at the phonological interface. Furthermore, will see evidence that prosody does not just encode tree structure, it also encodes information about the *derivational history* of a syntactic derivation, and the metrical representation will reflects information about *syntax* and *time*.

The idea that I want to pursue is that the overall prosody of an expression is computed based on decisions made locally, just as the overall denotation of an expression is computed by a stepwise combination of the denotations of its parts. Prosodic compositionality goes back to the transformational cycle of early generative approaches to phonology (Chomsky et al., 1957, Chomsky and Halle, 1968).

According to prosodic compositionality, the prosodic representation of (7a) is a proper subpart of the prosodic representation of (7b):

(7) a. (p and q) or p

b. $\underline{\text{((p and q) or p)}}$ and r

This is not to say that the substring underlined in (7b) is *phonetically* identical to (7a). The two structures are identical at a more abstract level, just as the [k]s in spoken renditions of 'cup' and 'cat' are not phonetically identical, but are usually assumed to share an identical piece of information in their representation, the featural representation of [k]. Surface phonetic differences in the realization come about as a result of how phonological structure is implemented.

The main concern in this chapter is how the grammar assigns these relative boundary ranks, and prosodic structure more generally. The implementation of prosodic structure is discussed in chapter 4; experimental evidence is reported in chapter 5.

2.1.2 Outline

This chapter looks at the fragment of English consisting only of proper names and the functors 'and' and 'or'. This fragment provides a simple mini-grammar in which we can establish some of the basics of the mapping from syntax to prosody. The argument proceeds in three steps.

What is the Generalization about Prosodic Boundaries?

First, we need to identify the generalization about the circumstances under which coordinate structures receive a 'flat' prosody, in which each conjunct is prosodically on a par, and under what circumstances they are prosodically more hierarchically organized. Consider the following three structures:

(8) a. $A \vee B \vee C$ 'Flat' Prosody: A | ∨ B | ∨ C

b. $A \vee (B \wedge C)$ 'Articulated' Prosody: A || ∨ B | ∧ C

c. $(A \vee B) \wedge C$ 'Articulated' Prosody: A | ∨ B || ∧ C

Why does (8a) have a flat prosody, but (8b,c) have articulated prosodies? The proposal advanced in section (2.2) is that the reason is the same as in logical notation: (9a), but not (9b,c) would count as a well-formed and complete formula without parentheses.

(9) a. $A \vee B \vee C$ Associative

b. $A \vee B \wedge C$ *Not* Associative

Domains that have a flat prosody are domains in which the associative law holds. This is the case in (8a). When the associative law does not hold, then prosody reflects the syntactic bracketing, as in (8b,c). In those cases, constituency is reflected by boundary strength.

The prosodic generalization can be stated in terms of 'levels of embedding'. Levels of embedding are domains that are associative.[5]

(10) ASSOCIATIVE DOMAINS

Levels of embedding (to be called 'Cycles') are associative domains.

Levels of embedding can be nested. Operators within lower levels of embedding take narrow scope relative to operators in higher levels of embedding, since they are syntactically more embedded in the structure. The notion of 'level of embedding' lets us state the generalization about prosody:

(11) SCOPALLY DETERMINED BOUNDARY RANK:

Iff Boundary Rank at a given level of embedding is n, the rank of the boundaries between constituents of the next higher level is n+1.

[5] A level of embedding with two elements trivially satisfies the associative law since no rebracketing is possible.

Constituents within a single level of embedding are set off by prosodic boundaries of equal rank; constituents within a lower level of embedding are set off by weaker boundaries than constituents within a higher level of embedding.

This hypothesis makes predictions for the prosody of coordinate structures. These will be discussed in detail in this chapter. The generalization directly relates prosody and scope, and captures the generalizations about prosody in coordinate structures.

What is the Syntax of Coordinate Structures?

What is the syntax of coordinate structures? This question is addressed in section 2.3. I will present evidence from combinatorics and from c-command tests that a binary-branching syntax combined with a cyclic view of syntactic derivations has just the right properties: It provides the correct combinatorial power and allows us to maintain the standard notion of c-command.

The cyclic approach to syntax exploits the property of recursiveness of language to define steps in the derivation. Grammar assembles structures in a workspace. Upon completion, a workspace is 'spelled out', which means that it is phonologically and semantically interpreted. These work-spaces form processing units in the derivation, and are called 'cycles'. The outputs of the cycles are the outputs of the grammar.

Any output of the grammar, i.e. the output of any cycle, can re-enter a new cycle as an atom, contributing the semantic and prosodic interpretation assigned at the earlier cycle. This is simply the generally assumed recursiveness of language, and has entered recent minimalist theorizing under the keyword 'multiple spell-out' (Uriagareka, 1999). The key for any cyclic theory of syntax is to specify which domains constitute a cycle. This question is at the heart of much of the current research in the minimalist theorizing within the framework of 'Phase Theory', following Chomsky (2001).

The new proposal here is that cycles correspond to the 'levels of embedding' defined above: they are associative domains. In other words, whenever three elements are combined in such a way that the associative law does not hold, then they must be assembled in two separate cycles. Only associative domains are built in a single cycle.

(12) Generalization about Syntactic Cycles

Each Syntactic Cycle is an Associative Domain.

It is the interface to semantics then that determines which nodes are cyclic nodes in syntax and which nodes are not. Nodes become cycles and are sent of to be interpreted in semantics and phonology in order to obtain a certain semantic interpretation. This chapter focuses on coordinate structures, but the approach to cycles in terms of associativity will be extended to other domains in chapter 3. We will see in chapter 6 that there is at least one additional factor that can render a syntactic node into a cyclic node. This will be a case where the interface to phonology is at play.

Evidence standard tests for c-command discussed in section 2.3 suggests that for associative domains that are assembled in a single cycle, grammar imposes right-branching structures. This, I claim, is due to the fact that within a single cycle, grammar generally can only construct right-branching structures. The following principle captures this property of grammar, which I assume to be a language-universal:

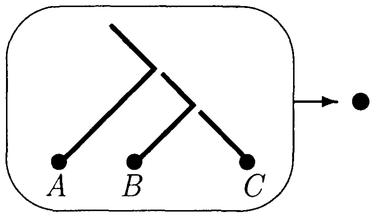
(13) Right-Branching Conjecture

Each cycle consists of a right-branching structure.

This is an adapted version of the principle proposed in Haider (1993, 2000), who proposed that 'shell-structures' are always right-branching. A similar principle was proposed in Phillips (1996, 19), who presents evidence both from syntactic constituency tests and from parsing preferences observed in the psycho-linguistic literature. The system proposed here shares with Phillips' system the property that structures in language are right-branching unless interpretation motivates a left-branching structure.

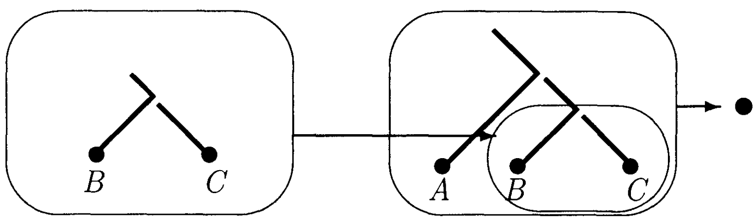
The structures in (8) can be built in one cycle. In the illustrations, each oval constitutes a cycle (also called 'workspace'). A cycle counts as an 'atom' for any higher cycle, and I sometimes represent them by filled circles to simplify the representation.

(14) (8a), Derived in One Cycle

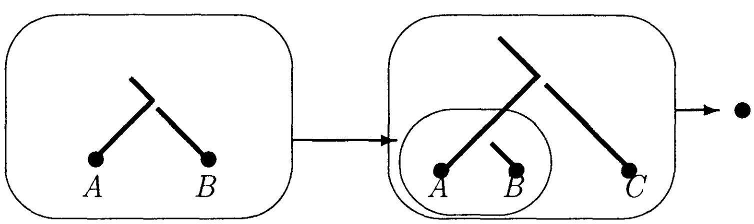


The structures (8b,c), on the other hand, have to be built in two cycles. Although (8a,b) are both right-branching, they differ in derivational history, and (b) consists of two separate cycles.[6]

(15) a. A or (B and C) (8b): Derived in Two Cycles



b. (A or B) and C (8c): Derived in Two Cycles



The cyclic theory based on binary branching has the right combinatorial properties and captures generalizations about c-command.

[6]Dresher (1994) observes that both types of right-branching structures are treated in the same way in the accent annotations in the Masoretic Bible. He convincingly argues that the accents reflect prosodic structure, and they faithfully distinguish right-branching from left-branching. The very notational system used makes it impossible to distinguish the two different right-branching structures. There are two possibilities: (i) the transcribers were forced by the notation they used to ignore the distinction, although it is present in the prosody; or (ii), the distinction is not prosodically realized, and the accent system faithfully transcribes the how the text is chanted. The latter possibility would be a problem for the present approach. It would imply either that Tiberian Hebrew is different in how it reflects prosody, or that the theory presented here is incorrect.

How does the Syntactic Cycle Interact with Phonology?

The final part of the argument shows how the syntactic derivation can be used to derive the correct prosodic structure. The correct prosody can be derived on the assumption that the phonological cycle applies upon completion of a syntactic cycle, following Bresnan (1971) and more recently Chomsky (2001). Once completed, the structure assembled within a cycle is 'spelled-out' and is mapped to a single prosodic foot in the metrical grid.

(16) Prosodic Matching

a. Concatenate

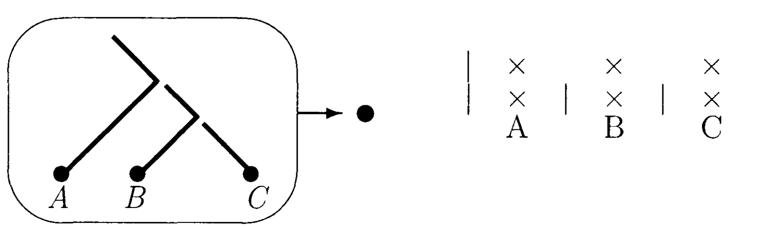
Concatenate the prosodic representation of the elements in the domain aligning their top lines and filling the columns where necessary.

b. Project

Create a new top-line grid line n by projecting all grid marks on line n-1, and mapping them into a single foot on line n.

'Concatenate' assures that constituents that are combined in a cycle start out on an equal footing in their prosodic representation. 'Project' derives a 'flat' prosody for elements assembled in a single cycle:

(17) A or B or C (8a): Derived in One Cycle



The pipes in the prosodic representation mark foot boundaries in the prosodic notation used here, the metrical grid. The height of columns of pipes is interpreted as the boundary rank of a juncture. The structure in (17) is 'flat' in that all elements are separated by boundaries of the same rank (here: a column of height 1). The metrical grid was proposed in Liberman (1975) to encode information about the timing of an utterance, and Selkirk (1984) explicitly relates the grid to phonetic correlates such as lengthening and pauses. I will discuss in detail how the representation is to be interpreted and how it relates to alternative representations in section 2.4.

When elements are combined in such a way that the associative law does *not* hold, then they have to be assembled in separate cycles. The output of a first cycle enters a new cycle as an atomic element. Since after each cycle spell-out maps the output to a prosodic foot in new grid line, the prosody reflects the way the structures are put together by syntax, deriving the following grids for (8b,c) respectively:

(18) Metrical Grid

a. A or (B and C) (8a)

```
| × | ×   × |
| × | × | × |
  A   B   C
```

b. (A or B) and C (8b)

```
| ×   × | × |
| × | × | × |
  A   B   C
```

This accounts for the fact that a more articulate prosody is derived for building (8b,c), but not for (8a). The crucial difference between the structures is that only (8a) is an associative domain and is therefore assembled in a single cycle.

While there are only two different syntactic trees involved in the three structures in (8), there are three different syntactic derivations. The prosody keeps track of the derivational history, and can be seen as another way of representing syntactic information. In addition to the syntactic relations represented in the tree, prosody reflects how it was assembled: prosody includes information about time.

The theory of syntax—phonology mapping presented here exploits the recursiveness of language to capture prosodic differences. Prosody reflects syntactic bracketing encoded in tree structures, but also the derivational history of how these tree structures were assembled. There are exactly as many prosodic structures as there are syntactic derivations, and the syntactic representation using binary branching trees allows us to maintain the standard notion of c-command in syntax. A single mapping principle (16) is sufficient to derive the intricate prosodic pattern of coordinate structures, and relates it directly to interpretative properties such as associativity and scope.

The interaction between cycles and prosody has the effect that a prosodically articulated structure is derived under the same circumstances that orthographic brackets are placed in the notation of logical formula: when the associative law does not hold. The prediction is then that structures are prosodically 'flat' unless interpretation motivates a more articulate prosodic structure.

2.2 Boundaries and Associativity

This section presents evidence that prosodic structure encodes the different bracketings of coordinate structures by assigning boundaries of different strength. Grammar differentiates boundaries of different strength only in cases in which the associative law does not hold.

2.2.1 Inducing Bracketing

Utterances consisting only of proper names and connectors can occur as fragment answers to questions:

(19) Who went to the forest?

 a. Lysander and Hermia.

 b. Lysander or Hermia.

The fragment answer (19a) can be seen as a shorthand for one of the following expressions, which seem to be identical in their truth conditions. In fact, fragment answers are often analyzed as being derived from full sentences by ellipsis (Merchant, 2003).[7]

[7]Some authors (e.g. Chomsky (1957, 36) andHarris (1957, 318ff)) have argued that sentences such as (20a) are derived from such sentential coordinate structures as (20b) by coordination reduction, an operation that deletes identical information in the two conjuncts. The appeal of this approach is that it offers the prospect to reduce all instances of 'and' to a Boolean connector that relates propositions with truth values. The following example illustrates that coordination reduction cannot be the right theory though:

(20) a. Lysander and Hermia went to the forest.

b. Lysander went to the forest and Hermia went to the forest.

The precise bracketing between the conjunctions with more than two elements often does not seem to affect the truth conditions. The associative law holds for coordination between propositions. The following three answers seem to have the same truth conditions in the context given:[8]

(21) Who went to the forest?

a. Lysander | and Demetrius | and Hermia.

b. Lysander | and Demetrius || and Hermia.

c. Lysander || and Demetrius | and Hermia.

The most natural answer in this context would be (21a). The more articulated prosodies in (21b,c) seem gratuitous, unless there is something in the context that motivates grouping two of the three individuals together. Structures (21b,c) introduce a presupposition that requires a certain grouping to be salient. Thus, if the prosodically articulated answer in (21c) is used, the listener would surely assume that the speaker has some reason to group Demetrius and Hermia together, possibly because they went to the forest together, or because they were grouped together in the conservation before.

---

(1) a. Lysander and Hermia gathered in the forest.

b. * Lysander gathered in the forest and Hermia gathered in the forest.

This does not necessarily imply that coordination is not always Boolean coordination. Schein (1997) and Winter (2001) both propose theories that only use Boolean 'and', and yet they do not analyze coordinate structures always via coordination reduction.

[8]I will not be concerned here whether a coordination should be treated as the projection of the connector or a projection of the first conjunct as in Munn (1993). I will assume however, that coordination in general returns an element of the same semantic type, and thus can apply recursively to connect more than 2 elements.

Consider the following context, in which 'Hermia' is contrasted with 'Demetrius and Helena'. This context favors (21c):[9]

(22) Lysander and Hermia?

No. Lysander || and Demetrius| and Helena.

The nature of the presupposition and how it affects prosody is discussed in chapter 7. In the following, I will look at cases where prosody affects truth conditions of expressions. Adverbs such as 'together' and 'alone' interact with prosodic grouping in ways that affect truth conditions. The answer in (23a) is compatible with Lysander going on his own, but (23b) cannot have this reading.[10]

(23) a. Lysander || and Demetrius | and Hermia together.

b. Lysander | and Demetrius || and Hermia together.

The adverb 'both' can also be used to induce bracketing (cf. Winter (2001)). 'Both' presupposes that there are two conjuncts, and not more. Now consider:[11]

(24) a. Both Lysander || and Demetrius | and Hermia.

b. Both Lysander | and Demetrius || and Hermia.

The flat prosody is impossible with more than two conjuncts and 'both':

(25) * Both Lysander, | [and] Demetrius, | and Hermia.

---

[9]It also seems to be possible to use the flat prosody in the answer, and contrast 'Lysander and Hermia' with 'Lysander, Demetrius, and Helena'. The point is that information structure *can* motivate prosodic bracketing.

[10]And the additional sub-grouping of Lysander and Demetrius requires that there is some motivation for grouping those two together to the exclusion of Hermia.

[11]Chomsky (1965, 196) alternates 'and' with 'as well as' in order to illustrate embedded coordination:

(1) John, as well as Mary and her child.

'as well as'. I am not sure whether 'as well as' necessarily requires an articulated prosody when combined with 'and' or not.

The interaction of 'both' with prosodic bracketing can be made sense of if (24a,b) involve a coordination distributing over two propositions, one of which includes a plural subject consisting of the two conjuncts that are grouped together more closely. This type of collective reading may also be involved when using adverbs such as 'together'. Collective readings are analyzed in Winter (2001) as involving a covert collective operator that creates a plurality (but see Schwarzschild (1996) for an alternative view). Consider the following example (cf. Winter (2001, 31), also Schein (1993)):

(26) a. Lysander and Demetrius weigh exactly 200 lbs.

b. Lysander weighs exactly 200 lbs and Demetrius weighs exactly 200 lbs.

Sentence (26a) has a collective reading that (26b) lacks, which is that Lysander and Demetrius together weigh 200 lbs. Both sentences have a 'distributive reading', in which Lysander and Demetrius each weigh 200 lbs. In cases that involve a coordination with a collective interpretation the bracketing of coordinate structures can play a crucial role.

(27) a. Lysander | [and] Demetrius | and Hermia weigh 200 lbs.

b. Lysander | and Demetrius || and Hermia ||| weigh 200 lbs.

c. Lysander || and Demetrius | and Hermia ||| weigh 200 lbs.

Each sentence has a distributive reading in which each of the three weighs 200 lbs. Example (27b) has an additional reading, where Lysander and Demetrius together weigh 200 lbs and Hermia does so alone. This collective reading that groups Lysander and Demetrius together is absent in (27c), which in turn has a reading in which Demetrius and Hermia together weigh 200 lbs and Lysander weighs this much alone.

Consider also the following example. The answer in (28a) seems inappropriate, while (28b) and (28c) differ with respect to who will have to share an apple:

(28) Two apples were given out, but I don't know to who. Who was given an apple?

a. # Lysander | and Demetrius | and Hermia.

b. Lysander | and Demetrius || and Hermia.

c. Lysander || and Demetrius | and Hermia.

Hoeksema (1983) observes the following case, where again a collective reading is involved, licensing a certain reading of the reciprocal:

(29) (Blücher and Wellington) and Napoleon fought against each other near Waterloo.

It is not necessary to make the subset of individuals into a constituent in order to form an antecedent for a reciprocal interpretation. The following sentence can have the same reading as (29):

(30) The generals fought against each other.

But it is hard if not impossible to give the following sentence the reading intended in (29):

(31) Blücher and (Wellington and Napoleon) fought against each other near Waterloo.

Another way to force the bracketing is to alternate 'and' and 'or'. With alternating functors it seems that only the articulated two bracketings are possible.[12]. This observation was also made in Min (1996):

(32) a. ?? Lysander | and Helena | or Demetrius.

b. Lysander | and Helena || or Demetrius.

c. Lysander || and Helena | or Demetrius.

[12]It may be possible to pronounce at least (32) with a pretty flat prosody, and in fact several renditions of the recordings collected in the experiment reported on in chapter 5 sound like this happened. Maybe lower bracketings can be 'washed out', at least in fast speech. But a careful pronunciation would certainly involve either of the two articulated prosodies, depending on the intended meaning.

While in the earlier examples, it was some overt or covert scope-taking element that seemed to interact with the bracketing in coordination structures, in the example in (32) it is the scope of 'and' and 'or' with respect to each other that is at stake. It seems that in each case some covert or overt scope-taking element takes scope over some conjuncts but not others. The scope of sentential adverbs in coordinate structures also interacts with prosodic grouping:

(33) Who went to the forest?

a. Lysander || and probably | Demetrius || and Hermia.

b. Lysander ||| and probably || Demetrius | and Hermia.

In example (33a), only Demetrius' going to the forest is uncertain, whereas in (33b) the whereabouts of both Demetrius and Hermia are not certain. When 'probably' takes scope over a single conjunct, then the conjuncts can be separated by boundaries of equal rank, as in (33a); but when 'probably' takes scope over the two last conjuncts, these two are grouped together and separated from the first conjunct with a stronger boundary (33b).[13]

Other sources for sub-groupings in conjunctions exist. Consider the difference in phrasing in the following examples (cf. Gleitman, 1965, 277):

(34) What's on the menu?

a. Ham | and tongue | and tuna fish | and pastrami.

b. Ham | and eggs || and tuna fish || and pastrami.

Here food items that form single dishes are grouped together. Items can also be coordinated to create complex names, here names of linguistic publications (Bob Ladd, p.c.):

[13]Of course, it is possible for 'probably' to take scope over a single conjunct even when there is a stronger boundary after the first conjunct. The difference in prosody from (33) is then that the boundary following 'probably' is weaker than the boundary separating the two last conjuncts:

(1) Lysander ||| and probably | Demetrius || and Hermia.

(35) What did you read?

Chomsky | and Halle || and Kiparsky | and Kiparsky.

What is the commonality among all cases in which prosodic boundaries of different ranks are induced? There is a a simple formal property that all cases that show a more articulated prosody share, which is that they are not *associative*.

2.2.2 The Associative Law and Prosodic Domains

A constituent consisting of three elements x, y, and z is associative if the meaning[14] do not change under either order of composition (The parentheses '()' stand for functional application, as defined in Heim and Kratzer (1998)):

(36) ASSOCIATIVE LAW

$([[x]])[[yz]] = [[xy]]([[z]])$

Structures with a flat prosody are associative, in other words the bracketing does not carry any important information.[15]

Consider an analogy to the orthographic convention on placing parentheses in logical notation. Formulas (37a,b) are well-formed as they are, while formula (37c) seems incomplete:

(37) a. $A \vee B \vee C$

b. $A \wedge B \wedge C$

c. $A \vee B \wedge C$

Parentheses are redundant in (37a,b), since in cases where 'and' and 'or', respectively, are iterated the associative law holds; but it does not hold when 'and' and 'or' alternate, as in (37c), hence the bracketing becomes crucial.

The generalization for the coordinate structures that emerged in the previous section is that in cases when $n > 2$ elements are combined, a prosodically articulated

[14] Relevant are at least the truth conditions and presuppositions, possibly also implicatures.

[15] In math, functions that are associative are sometimes called 'flat', e.g. 'flat' is the attribute used in *Mathematica* to mark a function as associative.

structure is obligatory exactly when the associative law *does not* hold, be it due to the presence of overt or covert additional scope-taking operators or because we alternated the functors 'and' and 'or', which take scope with respect to each other.

The correlation between 'flat' prosody and associativity, I want to claim, is not accidental. I propose the following hypothesis:

(38) THE ASSOCIATIVE-PROSODIC-DOMAIN-HYPOTHESIS

The associative law holds in the composition of constituents iff they are on a par prosodically.

The hypothesis is descriptively correct for the domain of evidence that we have looked at so far. It might seem that the solution is very specific to the domain under investigation: coordinate structures. In chapter 3 it is shown that it applies to other domains as well.

Let us call a sequence of elements, combined in such a way that the associative law holds, a single level of embedding, or a 'cycle'.

(10) ASSOCIATIVE DOMAINS

Levels of embedding (to be called 'Cycles') are associative domains.

In cases where grammar combines more than two elements, it makes a decision whether to map them to a single level of embedding or whether to map them to two different levels of embedding.

Depending on the grammatical framework one assumes, the syntactic definition of level of embedding will have to be adjusted accordingly. I will give a derivational interpretation in terms of 'cycles' in section (2.3).

When elements are combined such that associativity does not hold, they take scope relative to each other. Depending on the bracketing, in the following structures 'and' has either wider scope than 'or' or more narrow scope.

(39) a. A || or B | and C *or > and*

b. A | or B || and C *or < and*

Within each level of embedding, the conjuncts are separated from each other by boundaries of *equal rank.* If any conjunct is itself complex and contains further nested bracketings, the boundaries within that conjunct have to be of *lower ranks*, and the operators within this domain take narrow scope relative to the operators at higher levels of embedding. The syntactic grouping imposes a ranking of boundaries. The proposed generalization about the grouping can be summarized as follows:

(11) SCOPALLY DETERMINED BOUNDARY RANK (**SBR**):
If Boundary Rank at a given level of embedding is n, the rank of the boundaries between constituents of the next higher level is n+1.

A level of embedding can be then identified both based on its prosodic and interpretative properties. It forms a prosodic domain and is set off with stronger boundaries from material outside of this domain; semantically, it forms a domain in which the associative law holds.

If each conjunction and disjunction is binary, as I have been tacitly assuming, then we need to refer to associativity in order to explain why sometimes prosody in coordinate structures is flat, and sometimes it is more articulated. There is an alternative to the binary view of connectors, however. What if 'and' and 'or' are functors that are able to take infinitely many arguments? In this case, structures with 'flat' prosody can be analyzed as projections of a single 'and' or 'or'. We could then capture the differences in coordinate structure by stipulating that each single coordinate head defines a 'level of embedding' but can connect more than 2 elements in a prosodically 'flat' way—without reference to associativity. I will give some arguments against this view in the appendix in (2.5.1).

More generally speaking, the alternative to the the approach in terms of associativity is to simply stipulate which domains are 'cycles' or 'phases'—e.g. in Chomsky (2001), a list of phasal heads is given that are proposed to be phasal.[16] It is certainly

[16]According to Chomsky (2001), Nissenbaum (2000), it should be the *complement* of a phasal head that forms a cycle, so we would have to stipulate a head that takes each coordinate phrase as its complement, or show that every element that takes coordinate phrases as a complement happens to be phasal. It could of course be that sometimes coordinate phrases are *not* a separate cycle.

possible that it is necessary to stipulate a list of phases, and that there is no deeper generalization behind the list of cyclic domains. Using the associative law in the definition of a 'cycle' or 'phase' strikes me however as a more interesting possibility and the evidence from coordinate structures discussed here and from other domains discussed in chapter 3 suggests that it has some promise.

2.3 The Syntactic Cycle

The second step in the argument is to establish how the syntax of coordinate structures works. An important question is how many different bracketings there are in a coordinate structure with n conjuncts, and what representations of syntax are apt to encode them in an insightful way. A closer look at the syntax paves the way to link the observed prosody in a straightforward way to the syntactic derivation, exploiting the concept of the syntactic 'cycle'.

2.3.1 The Combinatorics of Coordination

A theory of syntactic representation should give us just the right number of bracketings for phonological and semantic interpretation. How many bracketings are there in coordinate structures, and what theories of syntactic representations have the right combinatorial power?

We can use the methods to induce bracketing described in section 2.2 to establish how many bracketings are possible. Based on the findings so far, it seems that we have to distinguish at least three bracketings for the case of n = 3 conjuncts:[17]

(40) a. Lysander | and Helena | and Demetrius.

b. Lysander | and Helena || or Demetrius.

c. Lysander || and Helena | or Demetrius.

[17]This is not to say that there are only three 'readings' of these coordinate structures. But all the strategies of forcing further bracketing discussed seemed to converge on one of these three prosodic options.

There is sufficient confusion in the literature on coordination to warrant a closer look at the combinatorics involved. Gleitman (1965, 276) gives 3 different bracketings of structures involving n = 4 conjuncts all connected by 'and'; Langendoen (1998) claims that there are 7. How many bracketings are there for n-many conjuncts in general?[18]

If all coordinate structures were binary, and if there are only binary branching trees (as argued i.a. in Kayne (1983)), then it would be unexpected that there are *three* different syntactic structures with 3 conjuncts. There are only two binary branching trees that one can construct for three elements under these assumptions:

---

[18]The bracketing of the functors themselves is irrelevant and not used for syntactic disambiguation. There might be a contrast between cliticizing to the right or to the left, but this, as far as I can tell, does not correspond to a syntactic/semantic distinction:

(1) a. Lysander and Hermia.
    b. Lysander'n Hermia.

In the following I will group the connectors with the following conjunct. There are several arguments in favor of this grouping. Ross observed the following contrast (Ross, 1967, 90–91):

(2) a. John left. And he didn't even say good-bye.
    b. * John left and. He didn't even say good-bye.

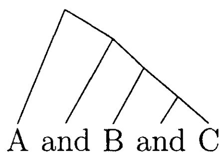
When a conjunct is extraposed, it is the second conjunct that extraposes, and the connector has to extrapose as well Munn (15 1993, , attributed to Gert Webelhuth):

(3) a. John bought a book yesterday, and a newspaper.
    b. * John bought a newspaper yesterday a book and.
    c. * John bought a newspaper and yesterday, a book.

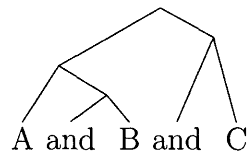
Since the connectors always group with the conjunct following them, we can ignore them when we count the syntactic bracketings of conjuncts in coordination structures.

(41) Two Binary Branching Trees, Assuming Binary Functors

a. Right-Branching



b. Left-Branching



The combinatorial question of how many binary branching trees there are over n linearly ordered elements is known as 'Catalan's Problem'.[19] The number of binary branching trees over n linearly ordered items is the Catalan number of (n-1). If coordination is always binary and attaches to the right, then the coordinators can be ignored in assessing the number of possible coordination with n conjuncts. The Catalan number for n = 3 is 2. The recurrence relation for the Catalan number is given below:

(42) Catalan number

Catalan Number of (n-1) is the number of binary branching trees over n linearly ordered items.

$$C_{n+1} = \frac{2(2n+1)}{n+2} C_n$$

Catalan Sequence: for n = 1, 2, 3, 4, 5..., C_n is 1 1 2 5 14 ...

While the Catalan number of n = 3 equals 2, the number of bracketings in coordinate structures is (at least) 3, as was shown above.[20]

[19]Named after the Belgian mathematician Eugène Charles Catalan (1814-1894). For references to proofs to the solution to Catalan's problem see Sloane and Plouffe (1995) or Sloane's archive of integer sequences at http://www.research.att.com/ñjas/sequences/Seis.html. The Catalan sequence is Sloane's A000108.

[20]Geoff Pullum claims in a language log post that there are 3 bracketings for 3 nominals: "This corresponds to the fact there are three bracketings for lifestyle consultation center: [lifestyle consultation center], [[lifestyle consultation] center], and [lifestyle [consultation center]]". http://itre.cis.upenn.edu/ myl/languagelog/archives/000160.html. Thanks to Chris Potts for pointing this out to me.

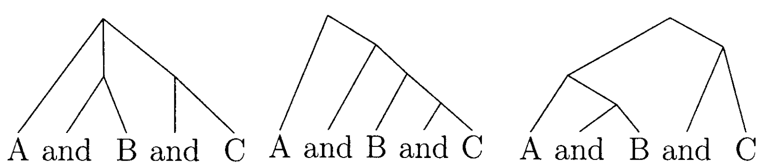
This is true at least if all connectors are identical, since if we alternate and/or the flat structure is ruled out. Only the two articulated prosodies are needed. Only for this case does the Catalan sequence gives us the correct result. Church and Patil (1982) show that this number of bracketings is the one generated by a phrase structure rule that assumes each step of coordination to be binary:

(43) $A \rightarrow A\ Co\ A$

One way to get to the right number of structures is to abandon the assumption of binary branchingness:

(44) Trees with Unbounded Branching

a. Flat b. Right-Branching Left-Branching



This is combinatorially equivalent to positing phrase structure rules that allow any number of arguments (Chomsky and Schützenberger (1963)):[21]

(45) $A \rightarrow A\ Co\ A\ (CoA)*$,

where $(Co\ A)*$ is zero or more occurrences of $Co\ A$

The number of bracketings for n conjuncts is then the number of trees over linear strings allowing arbitrary branching. This number is called the Super-Catalan number of n (see Stanley (1997) for discussion and references to proofs).[22]

These numbers are also called 'little Schröder numbers' and 'Plutarch numbers'.[23] A definition in terms of a recurrence function is given below:

[21]See also Gazdar et al. (1985).

[22]The Super-Catalan Sequence is Sloane's A001003.

[23]Stanley (1997) gives an incisive introduction into the combinatorics of this sequence and the history of its discovery, from which the following information is distilled:

Plutarch (ca. 50 A. D.–120 A. D.) was a Greek biographer. In his *Moralia*, Plutarch states the following: "Chrysippus says that the number of compound propositions that can be made from only ten simple propositions exceeds a million. (Hipparchus, to be sure, refuted this by showing that on the affirmative side there are 103049 compound statements, and on the negative side 310952.)"

(46) Super-Catalan Number

Super Catalan Number of n is the number of unbounded branching trees over n linearly ordered items.

$$S_n = \frac{3(2n-3)S_{n-1}(n-3)S_{n-2}}{n}$$

Super-Catalan Sequence: For n = 1, 2, 3, 4, 5..., S_n is 1 1 3 11 45...

The possible bracketings of coordinations with n = 3 and n = 4 are summarized schematically below:

(47) a. n = 3

xxx, x(xx), (xx)x

b. n = 4

xxxx, xx(xx), x(xx)x, (xx)xx, x(xxx), (xxx)x, (xx)(xx), x(x(xx)), x((xx)x), ((xx)x)x, (x(xx))x

In the previous section several ways of inducing prosodic bracketings were presented. We can use the bracketings in (47) and try to replicate them with prosodic bracketing in coordinate structure. We can induce all of the bracketings in (47) by keeping the functors within a level of embedding constant.

Changing the functor from 'and' to 'or' or vice versa will induce a hierarchical organization, as discussed above. It thus forces a prosodically articulated structure

---

103049 is the Super-Catalan of 10, and is the number of bracketings for n=10 conjuncts. If we only combine propositions, as Plutarch suggests, then the associative law would render bracketing irrelevant in cases where we iterate the same connector. If we allow ourselves the use of both and/or, then the number of complex statements is of course 2*103049, since for each bracketing with 'and' at the top level (of which there are Super-Catalan-many) there is one with 'or' at the top level, so Hipparchus may have been off by a factor of 2, or Plutarch quoted him wrongly.

See Stanley (1997) for a discussion of the actually fairly recent discovery that Plutarch was not just throwing out some random digits, in 1994, and Habsieger et al. (1998) for an interpretation of Plutarch's second number.

Ernst Schröder (1841–1902) was the first mathematician to specify a generating function.

precisely at the points where in orthography we write parentheses. In this way we can create structures corresponding to each of the ones listed in (47). This means that there are at least Super-Catalan-many bracketings in coordinate structures. To illustrate the point, consider first the case of n = 3.

The three prosodic bracketings of three elements in are illustrated in (48). (48a) is a disjunction with three alternatives; (48b) is a disjunction with two alternatives, the first if which is internally complex; (48c) is a disjunction with two alternatives, the second of which is internally complex. We alternate the functor exactly at the point that corresponds to parentheses in (47a):

(48) a. Lysander | or Hermia | or Helena.

b. Lysander | and Hermia || or Helena.

c. Lysander || or Hermia | and Helena.

Let's now turn to coordination structures with n = 4 proper names. The Super-Catalan number for n = 4 is 11. The simplest structure is a simple coordination or disjunction of four elements. The example given here is a disjunction of four alternatives:

(49) | Lysander | or Demetrius | or Hermia | or Helena. |

There are more complex structures to be considered, however. Consider the following three structures, all of which semantically involve a disjunction with three alternatives, one of which is internally complex:

(50) a. Lysander || or Demetrius or || Hermia | and Helena.

b. Lysander || or Demetrius | and Hermia || or Helena.

c. Lysander | and Demetrius || or Hermia || or Helena.

The following are structures with two alternatives, one or two of which are internally complex:

(51) a. Lysander || or Demetrius | and Hermia | and Helena.

b. Lysander | and Demetrius | and Hermia || or Helena.

c. Lysander | and Demetrius || or Hermia | and Helena.

There are four more structures which involve two alternatives. Just as in (51a,b), one of the two is internally complex. The difference to the structures in (51a,b) is that the complex structure involves another level of embedding:

(52) a. Lysander ||| or Demetrius || and Hermia | or Helena.

b. Lysander ||| or Demetrius | or Hermia || and Helena.

c. Lysander || and Demetrius | or Hermia ||| or Helena.

d. Lysander | or Demetrius || and Hermia ||| or Helena.

8

Structures with two levels of embedding are reported from corpora in Langendoen (1998, 240).[24]

(53) Combine grapefruit with bananas, strawberries and bananas, bananas and melon balls, raspberries or strawberries and melon balls, seedless white grapes and melon balls, or pineapple cubes and orange slices. (From the APHB corpus).

Langendoen (1998, 243) claims that two levels of embedding are impossible if the functors are all identical, and claims that in this case that there are only seven structures. This, however, seems to me to be incorrect.

Consider a situation in a diner, in which two guests have ordered breakfast. One guest ordered the special, which consists of bacon and eggs and a side of salad; the other guest ordered french toast. After a change of shift, a new waiter asks what was ordered:

[24] Langendoen (1998, 241) states as a fact that coordinate structures with n levels of embedding, with $n > 2$, are ungrammatical—this would mean that the depth of embedding of the structures in (52) is the maximal depth that natural language can generate—but he does not present any evidence in favor of this claim, apart from the intuition that a fully articulated right-branching structure involving only 'or' sounds odd out of context, which is hardly surprising.

(54) What should I take to tables 1 and 2 respectively?

((bacon and eggs) and salad)) and French toast

Again, this may not be the most practical way of communicating, but it seems to me that it is quite possible to use these answers, and that there is a grouping involved—least at the semantic/syntactic level. Whether or not the levels of embedding can all be prosodically differentiated is an empirical question. The question of how prosody differentiates structures with two levels of embedding is discussed in chapter 5.

The boundary ranks for the eleven structures containing four elements are summarized below:

(55) Scopally Determined Boundary Ranks

		A B C D		
1	A or B or C or D	1	1	1
2	A or (B and C) or D	2	1	2
3	A or B or (C and D)	2	2	1
4	(A and B) or C or D	1	2	2
5	(A and B) or (C or D)	1	2	1
6	(A and B and C) or D	1	1	2
7	A or (B and C and D)	2	1	1
8	((A or B) and C) or D)	1	2	3
9	(A and (B or C)) or D	2	1	3
10	A or ((B or C) and D)	3	1	2
11	A or (B and (C or D))	3	2	1

Natural language coordinate structures thus have at least Super-Catalan-many bracketings. I will assume from now on that there are *exactly* Super-Catalan-many bracketings.

The combinatorics presented in this section imposes a minimum on the number of structures that the grammar should provide for coordinate structures. A notation only allowing for binary branching trees, it seems, is not sufficient to distinguish enough coordinate structures.

A closer look at the syntax of coordinate structures offers an isomorphic alternative to flat trees that involves binary branching trees after all.

2.3.2 Syntactic Asymmetries in Coordination

In the exposition of the different possible coordinate structures over n elements in the preceding section we found that using unbounded branching introduces the right number of different tree structures, namely a number corresponding to the Super-Catalan number of n. One of the bracketings of n elements involves a 'flat' structure, i.e. a structure that semantically involves iteratively combining n elements with either 'and' or 'or'. This structure is prosodically 'flat' in that the conjuncts are separated by boundaries of equal rank and are thus prosodically on a par.

'Flat' unstructured representations have sometimes been taken as the null hypothesis for coordinate structures. Miller and Chomsky (1963, 196) take this to be the obviously correct approach: "Clearly, in the case of true coordination, by the very meaning of this term, no internal structure should be assigned at all within the sequence of coordinate items."

Unbounded trees were assumed in many subsequent generative approaches, e.g. Gleitman (1965) combines an arbitrary number of conjuncts under and n-ary branching single node and speaks of a 'main' and 'subordinate' conjunction in cases of nesting of coordinate structures that do not have a flat prosody.

This section presents evidence, however, that while the structures in question may be 'flat' in a prosodic sense, syntactically there is a clear asymmetry between the conjuncts. The source of evidence for the asymmetry is variable binding. Consider semantic binding first (Munn, 1993, 16):

(56) a. Everyone$_i$ and his$_i$ goat was at the party.

b. * His$_i$ goat and everyone$_i$ was at the party.

Semantic binding is possible from the first conjunct into the second but not vice-versa.[25] There is a clear asymmetry in coordination structures. But what is the

[25]Progovac (2003) argues that variable binding is licensed not by surface c-command but at LF after raising out the quantifier. The following example illustrates that this view is too simplistic:

(1) a. Every student$_i$, his$_i$ adviser, and the dean are invited to the reception.

b. * His$_i$ adviser, every student$_i$, and the dean are invited to the reception.

exact generalization about when semantic binding is possible?

Based on the examples so far, it could be that in coordination structures, linear precedence is sufficient for variable binding and for inducing condition-C effects. But more has to be said:[26]

(57) a. Every student$_i$ and (his$_i$ parents or his$_i$ adviser) have to be at the graduation.

b. * (Every student$_i$ and his$_i$ parents) or his$_i$ adviser have to be at the graduation.

This sentence does not have the reading indicated by the brackets in (b), in which it would be sufficient if all the advisers would show up, but no student. It seems that it is impossible to bind a variable from a more embedded conjunct into a conjunct of the next higher coordination or disjunction.

The left-right asymmetry persists even when the second connector takes scope over the first connector:

(58) * (His$_i$ adviser and his$_i$ parents) or every student$_i$ have to be at the graduation.

There is a simple generalization that captures the data—at least if we assume the binary right-branching representation of 'flat' coordinations:

The movement necessary out of the conjunction in (1) example would violate Ruys' generalization about movement out of conjunctions. The following examples illustrate further complications in the data pattern:

(2) a. ? The dean, every student$_i$, and his$_i$ adviser are invited to the reception

b. * Every student$_i$, the dean, and his$_i$ adviser are invited to the reception.

I am not sure how to account for the contrasts in judgements. Maybe Schein (2001)'s proposal that variable binding in coordinate structures involves 'telescoping' can account for them.

[26]It is not clear whether (a) involves genuine variable binding or involve e-type variables. Negative quantifiers could rule out an e-type reading, and it seems that they lead to ungrammaticality, at least with an articulated structure. The contrast between (57a,b) has to be accounted for in any case, and it suggests a distinctions in terms of c-command.

(59) C-command

A node A c-commands a node B, if and only if A's sister either

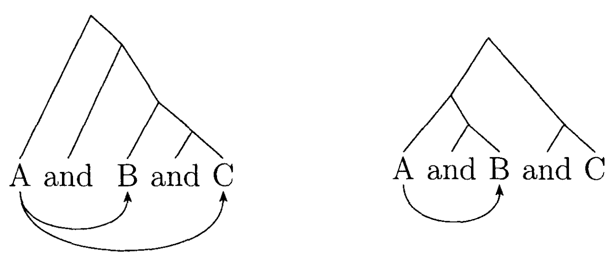
a. is B

b. or contains B.

This is the definition of c-command used in Adger (2003a, 117).[27] It correctly distinguishes left- and right-branching structures. Consider the following tree representations. All c-command relations between the conjuncts A, B, and C are represented by arrows:

(60) Left- *vs.* Right-Branching and C-Command

a. Right-Branching   b. Left-Branching



A c-commands C in the right-branching but not in the left-branching structure. In the binary-branching approach, even the coordination structures with a 'flat' prosody would have a right-branching tree. The structure in (57b) is ungrammatical because 'every student' does not c-command the variable. It is further embedded in a left-branch and the variable is not within the sister of the quantifier phrase.

Of course, the entire node (A and B) in the left-branching structure c-commands C, and consequently it can bind a variable within C:

(61) (Every student or visitor)$_i$ and his$_i$ parents are invited.

The important point is that A and B themselves do not c-command constituent C. Variable binding thus constitutes evidence that the structures that have a 'flat'

[27]It is close to the standard definition of c-command, at least assuming bare phrase structure Chomsky (1994); Reinhart (1976, 23) proposed a more elaborate definition, in part due to the possibility of intermediate projections.

prosody actually have the same c-command properties as prosodically articulated right-branching structures.

The observed pattern is compatible with an analysis that assumes binary branching, takes 'flat' coordinate structure to involve right-branching structures and adopts the standard notion of c-command. But unbounded branching gives us the correct combinatorial result for coordination structures with n elements, and thus seem necessary based on the combinatorial results. How can we square these conflicting results?

2.3.3 Binary Branching Revisited

In the following, I will present a notation using labeled binary branching trees that is isomorphic to the representation of flat trees. Since the two notations are isomorphic in their combinatorial power, we are free to choose based on some other consideration. In the interest of saving a unified definition of c-command without reference to linear precedence, I will choose the binary branching notation—but this is a choice of taste and not of empirical argument.

There is a translation of flat trees into binary branching trees. It is obvious that the set of binary branching trees is a subset of the set of arbitrarily branching trees. We can also give a translation of an arbitrarily branching tree into a binary tree (cf. Rogers (2003)). The idea is to use the linear precedence relation in the flat branching nodes (e.g. $A \prec B \prec C \prec D$) and convert it into a nested right-branching structure (i.a. [A [B [CD]] in this case). Each flat tree is assigned a unique binary branching tree in this way.

The translation of flat trees into binary trees also does not have to affect the way the composition of meaning proceeds.[28] In fact, the compositional interpretation of the translated structures does not differ at all from the way the meaning of the flat structures are composed if one follows Büring (2005, 40) in his definition of functional application for flat structures. Instead of using binary syntactic sisterhood, Büring

[28]Semantically, a function that takes n elements can be 'Schönfinkelized' or 'curried' into separate unitary functions (cf. Heim and Kratzer, 1998), so there is no reason to suspect that a translation of a flat tree into a binary branching tree would require a different theory of semantic composition.

uses the functor-argument asymmetry of adjacent constituents. For example, he composes flat VPs with two arguments by combining the predicate successively with its arguments one by one following linear order. The denotation of adjacent constituents are combined in a pairwise fashion. Function application and variable binding of flat structures then simply mimic the compositionality of a binary-branching structure. The translation to right-branching structures guarantees that it is always two uniquely defined sisters that compose and this voids the need to refer to adjacency.

The translation from flat trees to binary trees described here is not reversible: by converting flat branching nodes into right-branching we 'neutralize' the distinction between 'originally right-branching' and 'originally flat' nodes. What we need in order to show that we can represent all relevant structures with binary branching trees is an isomorphism that allows us to go from the arbitrarily branching trees to binary branching trees and back without loss of information. For this we have to annotate the binary-branching trees—eventually, we can abandon this annotation and replace it by a cyclic theory of syntax.

The piece of information that is lost is that all sisters in a flat branching tree are dominated by the same node. This missing information can be added to the binary trees by introducing labels to the tree notation. Borrowing from the notation tools of GB theory, we can label the nodes in the original flat tree 'XP-nodes'. The nodes newly introduced in the conversion from flat to right-branching remain unlabeled (we can refer to them as $\bar{X}$ nodes).

A property of the conversion function is that 'unlabeled nodes' or '$\bar{X}$' nodes are only assigned to right-branching nodes. We can call all items combined under a single XP-node (with intermediate $\bar{X}$-nodes) $\bar{X}$-shells or 'extended projections'. Larson (1988) indeed proposes extended projections to be right-branching, at least for VP-shells in English. Haider (1993, 2000) conjectures that this is a systematic property of all extended projections and states it as follows, essentially requiring a right-branching structure for extended projections (or shell-structures):[29]

[29] At least two syntactic approaches have in some way or other built into the system a preference of right-branching over left-branching structures: Kayne (1994) and Haider (1993). A detailed

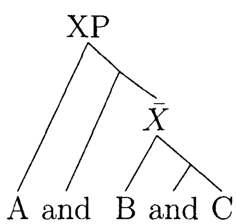
(62) Branchingness Conjecture (Haider, 2000)

For any two nodes *directly* attached to the *same* (extended) projection line, the node that precedes c-commands the node that follows.

Imagine that we would also allow left-branching $\bar{X}$ shells. Then for n = 3, if we allow an $\bar{X}$ label for any branching node (except the root) we get too many different structures, e.g. for n = 3 we get 4 different structures instead of 3:

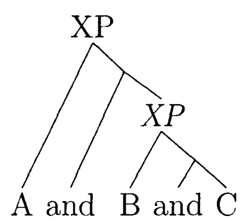
(63) Right-Branching

a. Flat prosody



'A and B and C'

b. Hierarchical Prosody

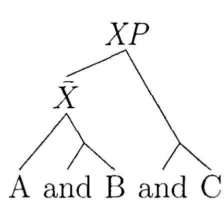


'A and (B and C)'

(64) Left-Branching

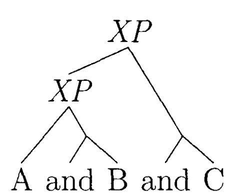
a. Flat Prosody

(Non-existant?)



?

b. 'Hierarchical Prosody



'(A and B) and C'

If we stipulate right-branching for $\bar{X}$ shells, labeled trees provide a combinatorial system consisting only of binary branching trees that map one-to-one to arbitrarily branching trees and thus a system that encodes Super-Catalan-many bracketings. The proposed mapping function has several useful properties. Once we convert the

consideration of coordinate structures within those two systems would go beyond the scope of this chapter.

flat structures according to the mapping function, there are exactly 2 right-branching structures for n=3 conjuncts, one prosodically 'flat' and one prosodically articulated:

(65) a. Lysander | or Demetrius | or Hermia.

b. Lysander || or Demetrius | and Hermia.

On the other hand, there is only a single left-branching structure, and this structure is prosodically fully articulated:

(66) Lysander | or Demetrius || or Hermia.

The right-branchingness of $\bar{X}$-shells in the binary notation leads to the generalization that right-branching structures can be prosodically flat, but left-branching structures are necessarily prosodically articulated.[30]

(67) Left-Branching Structures are Prosodically Articulated.

There are then two isomorphic ways to achieve the correct combinatorial power:

(68) Two ways get Super-Catalan-many structures

a. Unbounded Branching trees.

b. XP vs. $\bar{X}$ distinction in label of nodes + Right-Branching of Shells

They are notational variants of each other, at least with respect to the number of phrase markers/derivations they provide for a tree with n terminal nodes. The decision between them should be guided by the heuristic of which of those options allows for the simplest statement of relevant generalizations. The important lesson to learn from the combinatorics of coordination is that whichever of the two options in (68) we employ, we should pick *exactly one* of them, since otherwise the combinatorial possibilities explode.

Labeling nodes as XP-nodes or $\bar{X}$-nodes allowed us to use binary branching trees, and to maintain the standard notion of c-command—at the price of having to stipulate that shell-structures are right branching. The next section suggests a more insightful derivational view.

[30]Possessors in English seem to be a counterexample. See chapter 3 for discussion and evidence that they are underlyingly right-branching.

2.3.4 The Syntactic Cycle

This section presents a derivational view of syntax that will allow us to get rid of the labels (XP *vs.* $\bar{X}$) and has the desirable property that it allows us to exploit something that any generative theory assumes in order to define what the notion of a level of embedding is: the recursiveness of language.

The derivational theory distinguishes nodes that undergo a spell-out, and thus form a separate work-space in the cyclic building up of structure, from nodes that do not, and are part of a bigger structure constructed in a single work-space. Instead of using labels, the nodes we labeled XP-nodes earlier to constitute cycles, while those that are $\bar{X}$-nodes do not.

This allows for a simpler statement of the right-branching conjecture:

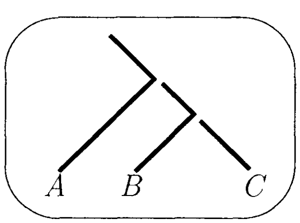
(13) Right-Branching Conjecture
Each cycle consists of a right-branching structure.

This is similar to the principle 'Branch Right' proposed in Phillips (1996), which favors right-branching structures unless it is not compatible with the semantic interpretation of an expression. The main difference is a technical one, namely that I am assuming a cyclic theory of syntax and stipulate right-branchingness as an necessary property of every syntactic cycle.

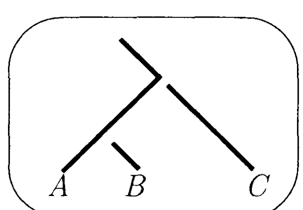
The idea is that a right-branching structure is the *only* structure the grammar can assemble given a numeration of elements. I will call the domain in which structures are assembled a cycle (workspace/phase/level of embedding). Within any workspace, only (69a) can be generated, (69b) is impossible:

(69) Structure within a Cycle

a. Possible:



b. Impossible:



In the absence of a motivation for either a left- or right-branching bracketing, grammar imposes a right-branching analysis. This is reflected in the preference for right-branching structures, which were developed into the asymmetric syntactic theories in Haider (1993), Kayne (1994), and is related to the 'cascade' structure discussed in Pesetsky (1995).

It is still possible to assemble left-branching structures—just not within a single cycle. The additional assumption that I make use of is that of recursion:

(70) Recursion

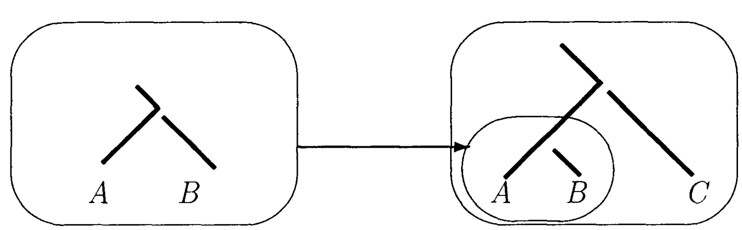
Each output of a cycle can enter a new cycle as an atomic expression.

This assumption is implicit in all syntactic theories I know of. What is not shared by all theories is that the derivational steps in the recursion matter in capturing grammatical phenomena. I will take this derivational view of syntax, as in (Uriagareka, 1999, Johnson, 2002).[31]

Left-branching structures require the combination of more than one work-space, while right-branching structures can be created in one go. Consider the following derivation:

(71) Creating a Left-Branching Structure

a. First Cycle b. Second Cycle:



[31]Johnson (2002) uses a notion 'Re-numeration', which is close to the way I will use the cycle, in order to explain adjunct island effects and also focus projection effects.

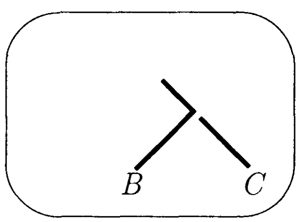
This is the *only* way of deriving a left-branching structure. A right-branching structure might be derived either in one workspace, or in two steps. There are thus 3 ways of assembling structures with 3 elements. In fact, for n elements, there are exactly Super-Catalan of (n-1) many ways of building structure. We can simply identify right-branching structures built in a single cycle with the flat nodes of the arbitrarily branching trees or with the shell-structures of the labeled tree representation.

The cycle has also been claimed to have an interpretative property, namely that the elements combined in it are combined in such a way that the associative law holds. In other words, we do not need labels to identify the cyclic nodes, it suffices to look at interpretative properties.

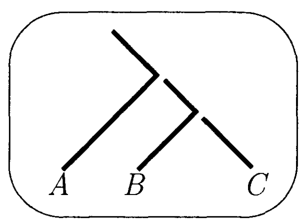
This means that the bracketing of the material within a cycle is immaterial. This property enables us to build up the structure in a cyclic way either top-down or bottom up. Derivations are conventionally assumed to work bottom-up:

(72) Bottom Up

a. First Merge



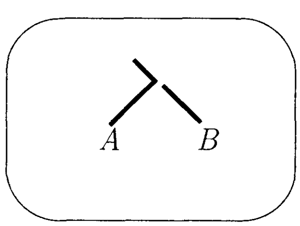
b. Second Merge



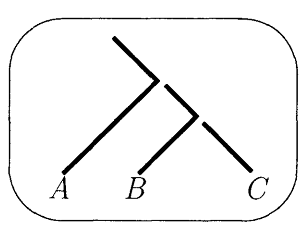
But this not the only possibility. Phillips (1996, 2003) proposed to build up structure left-to-right, in such a way that each step of Merge destroys phrase structure built up earlier:

(73) Top-down

a. First Merge



b. Second Merge



There is some evidence that a left-to-right, top-down derivation is possible in prosodically 'flat' coordinate structures. Frazier et al. (2000, 360) report results from an experiment in which subjects were asked to read coordinate structures, varying the number of conjuncts and whether or not all or only the last coordinator was pronounced.

(74) a. John and Fred.

b. John, (and) Bill and Fred.

c. John, (and) Bill, (and) Fred, and Harry.

There was no effect on reading time due to the number of conjuncts (or the presence/absence of connectors). This suggests that material pertaining to a single cycle can be assembled from left to right by subsequently adding new constituents, rather than assembling the structure bottom up. Otherwise, the processing cost of the ever increasing number of upcoming constituents should affect the realization of the preceding constituents. This would be a useful property of a unit for language processing in production and comprehension.

For the structure built within a single cycle, it is immaterial which of the options (72) or (73) we choose. If cycles are associative domains, then they are robust with respect to directionality of composition.

In the following I will assume then with Phillips (1996, 2003) that the order of the derivation is left-to-right, and the order in which elements are spelled-out into a new cycle is directly reflected in their linear order.

The last and final step of the argument is to show that the cyclic approach to syntax promoted here gives an elegant account of how to derive the prosodic structure.

2.4 Prosody and the Cycle

The last step in the argument is to show how the cyclic syntactic derivation can be put to use to derive a metrical grid that encodes the correct prosodic structure.

2.4.1 Prosodic Representation

Pierrehumbert (1999) defines prosody as the "grouping and relative prominence of the elements making up the speech signal". An adequate representation of prosodic structure should reflect intuitions about grouping, should transparently encode prosodically relevant information such as boundary strength and accent placement, and should encode prominence relations. Boundary strength is not transparently encoded in the syntactic tree structures usually employed in generative syntax:

(75) Syntactic Bracketing and Boundary Strength

a. A or (B and C) b. (A or B) and C

I will therefore use the notation of the *metrical grid*, a notion first proposed by Liberman (1975). The metrical grid reflects the timing structure of an utterance, and was originally proposed in analogy to musical notation. The particular notation I will use here is a version of the bracketed metrical grid (cf. Halle and Vergnaud, 1987). I illustrate the notation used here:

(76) Metrical Grid

a. A or (B and C) b. (A or B) and C

```
a.                          b.
| × | ×   × |               | ×   × | × |
| × | × | × |               | × | × | × |
  A   B   C                   A   B   C
```

The notation is to be read as follows: The unoriented parentheses ('pipes') demarcate feet at each level in the grid. Foot boundaries at higher boundaries necessarily correlate with foot boundary at lower grid lines. And a foot can contain multiple grid marks, i.e. feet are unheaded.

The notation diverges from the one employed in Idsardi (1992) and Halle and Idsardi (1995) in using foot boundaries that are not directional (left/right parentheses). Grids with undirected brackets have been proposed in Hall (2000) and Reiss (2003).[32] A further difference is that I am using matching parentheses, while Idsardi (1992) proposed unmatched parentheses—the difference is again not crucial for the data-set I am discussing here.

The representation encodes boundary strength in a straightforward way: feet at higher grid lines have stronger boundaries. The boundary rank can now be read off as the height of the column of pipes.

(77) BOUNDARY RANK

The rank of a prosodic boundary is the number of the highest grid line containing a pipe at that boundary.

How do boundaries induce the lengthening effects that at the heart of speakers' intuition about boundary strengths? One possibility is that the boundaries represented by the 'pipes' actually have a direct gestural correlate, which affects the timing of the realization of the segmental score. Such a boundary gesture was proposed in Byrd and Saltzman (1998) and Byrd and Saltzman (2003), the idea being that this prosodic gesture effects a slowing down immediately preceding the boundary. This effect recedes after the boundary, and hence there is a speed-up. I discuss durational evidence for boundaries in more detail in chapter 4 and chapter 5.

A second issue is the question of how does the representation encode the location of pitch accents in a structure? I will assume that the following generalization holds about the implementation of the metrical structure:

(78) ACCENT PLACEMENT

All and only top-line grid marks receive pitch accents. All other material is either unaccented or the accents are realized within a reduced pitch range.

In coordinate structures, in general each conjunct receives an accent, which is reflected in the notation here by projecting top line grid marks within each conjunct to the

[32] Hall (2000), Reiss (2003) present evidence that they are sufficient to account for variation in stress systems. Since I am not concerned here whether or not the notation is powerful enough to capture variation in word stress systems, I will not discuss the issue in any detail.

top line of the entire structure.

The grid notation in (76) encodes information about grouping that is already present in the tree-structure in (75). In fact, it seems that it encodes the same information, albeit in a different way. This redundancy is not problematic as long as we conceive of the grid notation and the syntactic tree notation as two ways of looking at the same information. We can think of the metrical grid as *a way of drawing a syntactic tree and the history of how it was assembled.*

The version of the grid I employ here departs from earlier literature (Prince, 1983, Halle and Vergnaud, 1987, Idsardi, 1992) in a number of ways. One major difference is that it does not represent nuclear stress, i.e. it does not encode the syllable that is perceived by native speakers as the most prominent. Just as the originally proposed grid in Liberman (1975), Libermann and Prince (1977), it encodes the major 'beats', but does not single one of them out as being the main stress by projecting it to a higher line than any other beat in the string.

How then does the grid representation encode prominence relations? I assume that the notion of 'nuclear stress' can be derived from the present notation without further addition or projection to a higher line. Newman (1946) and Truckenbrodt (1995) observe that in English, main prominence is perceived on the last accented element within a constituent.

(79) NEWMAN'S GENERALIZATION (Newman, 1946, 176)
"When no expressive accents disturb a sequence of heavy stresses, the last heavy stress in an intonational unit takes the nuclear heavy stress."

Coordination structures in English generally have main prominence within the last conjunct, but all conjuncts receive an accent. This is completely independent of the internal bracketing of the coordinate structure. In the following example, within the subject, the main prominence is perceived on 'Helena', no matter which of the bracketing of the coordinate structure is chosen:

(80) a. Lysander and (Hermia or Helena) are in the forest.
b. (Lysander and Hermia) or Helena are in the forest.

In SPE, for every phrase marker the nuclear stress rule (NSR) identifies exactly on vowel as most prominent. The vowel that native speakers perceive as most prominent in a given structure counts as its nuclear stress. The representation of stress in SPE is that of a phonological feature that can have any positive integer number as its value. The nuclear stress of an utterance is the unique vowel that has the value 1. Within every sub-constituent, there is exactly one vowel that has the lowest number. We can call this the vowel carrying nuclear stress within that sub-constituent.

A different notion of 'nuclear stress' or 'nuclear accent' is used in the intonation literature. Newman (1946, 174) singles out a heavy stresses as 'nuclear' if it 'acts as the nucleus of an intonational unit'. Within any intonational unit it is always the last heavy stress that counts as the nuclear stress. Newman (1946, 176) suggests that in coordinate structures ('enumerations') each conjunct receives a nuclear accent. The intuition that each conjunct is on a par will be captured here by the fact that the boundaries that separate them are identical. For the course of this section, I will stick to the SPE notion of nuclear stress, and I will return to Newman's idea about nuclear stress in the next section.

The reason the notation in (76) already encodes where nuclear stress falls within each constituent is that it is always the *last accented syllable.* Under the assumption that it is all and only elements that project to the top grid line that get linked to a pitch accent, we can capture the generalization as follows:

(81) Nuclear Stress Generalization

Within each foot, nuclear stress is perceived on the last of those grid marks that project highest.

Since each conjunct projects to the top line and receives an accent, the examples so far are not a particularly interesting illustration of Newman's generalization. Each conjunct in a coordinate structure projects, so nuclear stress is always final. In fact, the notation seems redundant, since we could just as well have used the following notation:

(82) Flat Notation

a. Right-Branching

|| × || × | × ||

A B C

b. Left-Branching

|| × | × || × ||

A B C

But under certain circumstances some conjuncts do not receive an accent:

(83) a. Demetrius and Hermia?

b. No, Lysander and Hermia.

In a natural rendition of (83b) the material following the first conjunct is deaccented or at least heavily pitch reduced. It is *prosodically subordinated.* It is this type of prosodic subordination which is responsible for cases in which nuclear stress does not fall on the final element in a phrase. The conditions under which prosodic subordination applies and the generalization about nuclear stress will be discussed in detail in chapter 6 and chapter 7. We can present prosodic subordination in the grid as follows:

(84) Subordination

| × |

| × | × |

Lysander Hermia

The 'nuclear' stress now falls on the syllable with main stress in 'Lysander', since it is the last unsubordinated element, i.e. the last element that projects to the top line. An important property of subordination is that the prosodic phrasing in the subordinated domain is maintained. Consider the following dialogue:

(85) Who will get married?

a. (Egeus and Helena) and (Lysander and Hermia)?

b. No! (Demetrius and Helena) and (Lysander and Hermia).

The grid notation of the response looks as follows:

(86) Phrasing in Post-nuclear domain

```
|      ×      |                               
|      ×      |    ×    |     ×          ×     |
|      ×      |    ×    |     ×     |    ×     |

   Demetrius     Helena    Lysander    Hermia
```

Nuclear stress of the entire answer is perceived on the main stress of 'Demetrius', since this syllable is the last (and only) one that projects a grid mark to the top line. For the same reason, nuclear stress within the first conjunct ('Demetrius and Helena') is perceived on that syllable.

The prosodic phrasing within the subordinated material is maintained, the break after 'Helena' is intuitively stronger than that after 'Lysander'. That post-nuclear phrasing still reflects differences in boundary rank was experimentally tested in Ishihara (2003) and Jaeger and Norcliffe (2005).

Finally, the nuclear stress generalization given here goes beyond Newman's generalization and applies to constituents that are entirely unaccented: The main prominence within the subordinated conjunct 'Lysander and Hermia' is perceived on 'Hermia', since this word contains the last syllable projecting to the highest line *within that constituent.*

The notion 'prominence' is thus not uniformly represented in the metrical representation: among a series of elements that project to the same grid-line, it is the last one that is perceived as most prominent, but elements that only project to lower lines are less prominent even when they are at the phrase end. The tacit assumption is that prominence is not a primitive of the theory but emerges in the task of asking prominence judgements. The benefit of the present representation is that it allows us to capture the observed prosodic grouping, as well as intuitions about prominence.

2.4.2 Prosody and the Cycle

In order to map syntactic structure to prosody, I propose a cyclic mapping to prosody. Just as in SPE, prosody is derived here using a 'transformational cycle'. I assume that the syntactic cycle corresponds to the phonological cycle. This hypothesis was first proposed in Bresnan (1971), and has recently gained more currency in the context of 'phase'-theory (Chomsky, 2001), i.a. Adger (2001), Marvin (2002), Arregi (2002), Wagner (2002a), Adger (2003b), Dobashi (2003), Ishihara (2003), Kahnemuyipour (2003, 2004) and Selkirk and Kratzer (2005).

The idea that I want to propose here is very simple: Elements are combined by Merge throughout the cycle. When an associative domain is completed, the cycle is 'spelled out'. Part of 'spelling-out' is to map the content of a cycle to a prosodic unit in the phonological representation. In particular, a new top line to the grid is added, and each top line grid mark within any of the elements of the cycle is projected up to this new grid line.

(16) PROSODIC MATCHING

a. Concatenate

   Concatenate the prosodic representation of the elements in the domain aligning their top lines and filling the columns where necessary.

b. Project

   Create a new top-line grid line n by projecting all grid marks on line n-1, and mapping them into a single foot on line n.

In essence, this principle simply maps the output of the cycle to a single foot in metrical structure. 'Concatenate' plays a similar role in the present theory as the 'stress equalization principle' in Halle and Vergnaud (1987). It assures that constituents that are combined in a cycle start out on an equal footing in their prosodic representation.

The elements of a cycle can be terminal nodes. I assume for simplicity that they come in as vocabulary items with some metrical structure associated with them in their phonological representation.[33]; but they may also consist of the output of other cycles an be internally complex.

In order to apply this principle, no knowledge is needed about constituency within the cycle, about the nature of the objects in it (which is a functor, which is an argument), or about the information structure in the context. All of these factors

[33] I will not discuss the cyclic foot structure below the word in this dissertation. We can think of the cycle relevant for within-word phonology as syntactic cycles as well, as proposed e.g. in (Marvin, 2002).

affect prosody, of course. But their effect is indirect, and is due to how these factors interact with syntactic recursion. This is why much of this thesis is dedicated to unravel properties of the syntactic derivation, although its ultimate goal is to account for the generation of the appropriate metrical structures.

Each conjunct in a coordination structure, I assume, has already undergone a cycle, and is thus spelled-out before in a separate cycle. I will not discuss the functors at this point but treat coordination structures as if they consisted just of the conjoined elements.

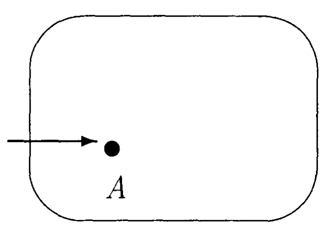
I will not discuss how linear order comes about. I will assume that along with assigning prosodic structure, at each cycle also the linear order within the cyclic domain is fixed, following Fox and Pesetsky (2005), Ko (2005), Sabbagh (to appear) and Takahashi (2005).

Merging a constituent in a cycle that already comes with a prosodic representation goes along with concatenating this element to the existing prosodic representation.[34]

(87) Building a Right-Branching structure in one single cycle:

A | and B | and C

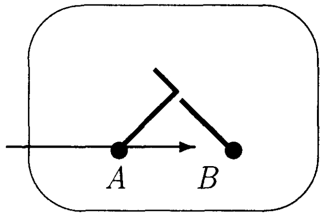
a. Start Point　　　　Prosodic Representation



| × |
A

b. First Merge

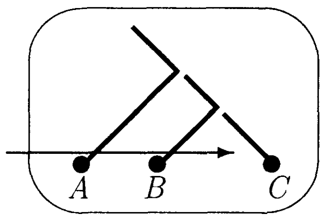
[34] I am presenting the derivation assuming with Phillips (1996, 2003) a top-down derivation. Given the discussion in section 2.3 above, it is not relevant whether a cycle is assembled bottom-up or top-down, and given the way the prosody is derived here it would not affect the outcome.



| × | × |

A B

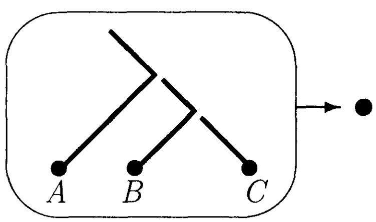
c. Second Merge



| × | × | × |

A B C

d. Spell-Out



| × × × |
| × | × | × |

A B C

Each of the terminal nodes in this tree could be terminal nodes[35] or themselves be cycles. In the coordinate structures discussed here they might always be separate cycles. Once spelled out, a cycle is opaque and the information within it is no longer accessible. *A filled circle stands in for a cycle that has already been spelled out.*

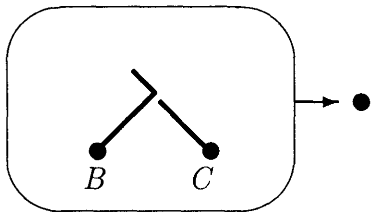
A right-branching structure with three terminal nodes can also be derived in two separate cycles. This becomes necessary when the associative law does not hold, as in the following expression:

(88) A and (B or C)

[35]For simplicity we can assume that lexical vocabulary items come in with some metrical structure. I will not go into the syntactic cycle below the word in this thesis. One way of thinking about word internal cycles is to simply assume that they are also syntactic cycles (as in Marvin (2002)).

One cycle combines the sub-constituent (B and C).

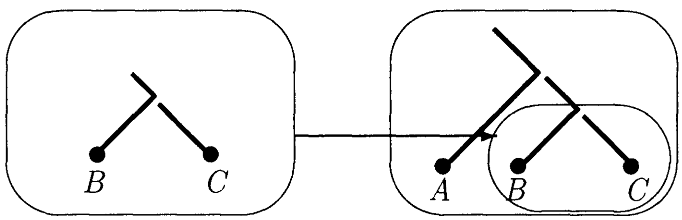
(89) First Cycle, Spelled-Out




```
| ×   × |
| × | × |
  B   C
```

The output of the first cycle is inserted into a second cycle:

(90) Second Cycle



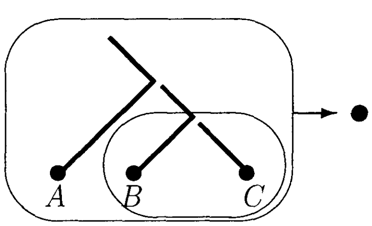
Before projecting the a foot on a new top grid line, the metrical elements corresponding to the elements in the cycle are concatenated. The grid representation of the concatenated grids looks as follows:

(91) Concatenating the Grids

```
| × | ×   × |
    | × | × |
  A   B   C
```

When the second cycle is spelled out, a new grid line is projected. I will also 'fill up' the grid column that is unfilled, simply for notational reasons—the relations in the grid are relative, and constituents of different complexity can be matched to the same grid level.

(92) Spell-Out of Second Cycle

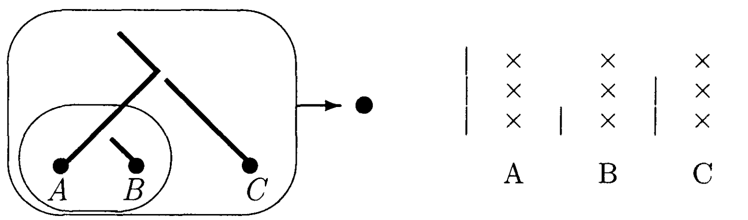



```
| ×   ×   × |
| × | ×   × |
| × | × | × |
  A   B   C
```

Two different grids are derived, although the same binary tree is involved. The grid keeps track of the difference in derivation.

A left-branching structure necessarily is created via two cycles, as observed above. The grid that is derived looks as follows:

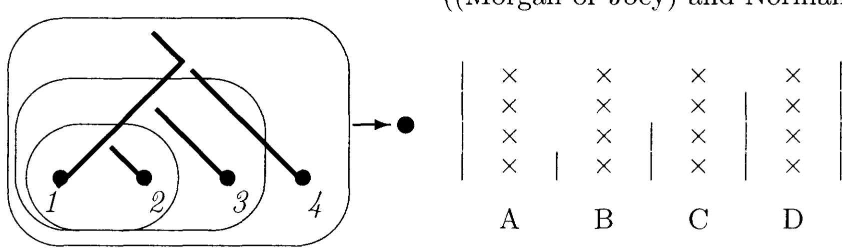
(93) Left-Branching Structure



Three different grid structures are derived, for the three different derivations. Similarly for the case of n=4, the metrical grid distinguishes 11 different derivations for the 5 different binary trees. Consider first a fully left-branching tree, which necessarily goes through 3 cycles:

(94) Left-Branching: 1 Derivation

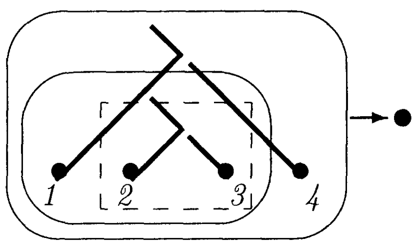
((Morgan or Joey) and Norman) or Ronny



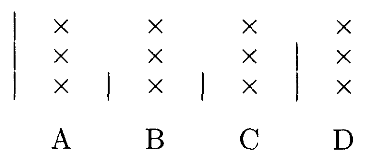
There are three binary branching tree involving 4 elements that have 1 right-branching node and one left-branching node. The elements on the right branch of a right-branching node can be either a separate cycle or be assembled together with the other material in the tree—which one is the right output depends on whether or not the associative law holds; the elements under the left branch on a left-branching node necessarily form a cycle together.

In the tree below, it is the first three elements that are on the left branch, and thus always form a separate cycle. Elements 2 and 3 may or may not form an additional separate cycle (a node that could be a separate cycle or be part of a bigger cycle, i.e. a right-branching node, is marked by the a dotted rectangle, which reflects the fact there could be an extra cycle—depending on the content of the terminal nodes and whether or not the associative law holds within the cycle that contains it):

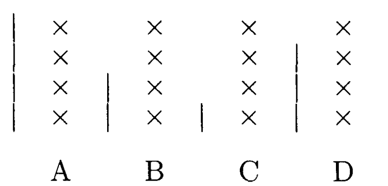
(95) One Right-Branching node: 2 Derivations



i. (Morgan and Joey and Norman) or Ronny

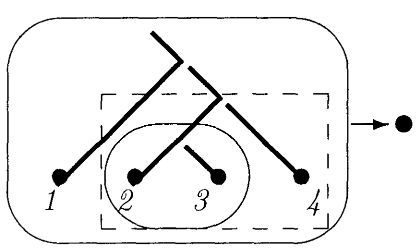


ii. (Morgan and (Joey or Norman)) or Ronny

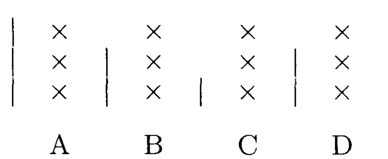


In the second structure, the middle two elements are on the left branch of the left-branching node and thus form a cycle. The last three structures may or may not form another separate cycle:

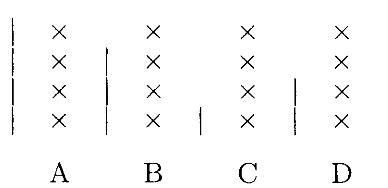
(96) One Right-Branching node: 2 Derivations



i. Morgan or (Joey and Norman) or Ronny

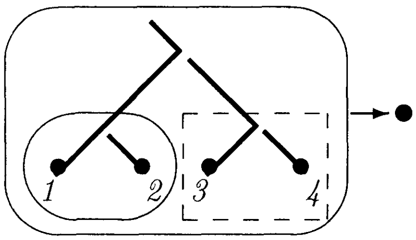


ii. Morgan or ((Joey or Norman) and Ronny)

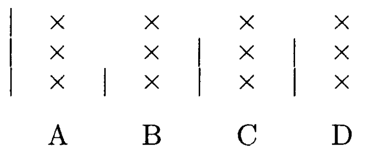


In the following tree structure, the first two elements are on the left branch of a left-branching node, and the last two elements can form an additional cycle:

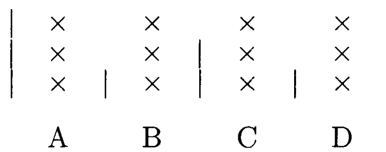
(97) One Right-Branching node: 2 Derivations



i. (Morgan and Joey) or Norman or Ronny

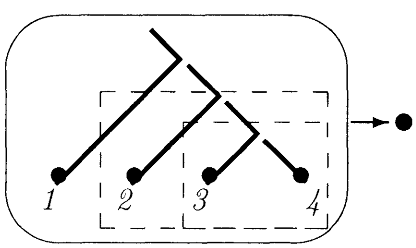


ii. (Morgan and Joey) or (Norman and Ronny)

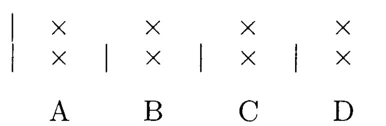


Finally, the fully right-branching tree corresponds to 4 derivations, since each right branching node may or may not contain a separate cycle.

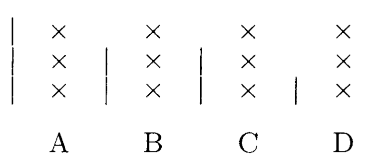
(98) Two right-branching nodes, 2 x 2 = 4 derivations



i. Morgan or Joey or Norman or Ronny



ii. Morgan or Joey or (Norman and Ronny)



iii. Morgan or (Joey and Norman and Ronny)

```
| ×   ×   ×   × |
| × | ×   ×   × |
| × | × | × | × |

  A   B   C   D
```

iv. Morgan or (Joey and (Norman or Ronny))

```
| ×   ×   ×   × |
| × | ×   ×   × |
| × | × | ×   × |
| × | × | × | × |

  A   B   C   D
```

The cyclic approach derives representations that reflect speaker's intuitions about grouping in coordinate structures. There are 11 different derivations, and correspondingly there are 11 different grids. The only mapping principle necessary is the one that the domain of a spell-out is mapped to a new foot in the grid (16). The grid comes about as a result of the way syntactic derivations interact with this mapping principle. The metrical representation reflects not just the tree structure, but also how it was assembled.

Summary

In the fragment of coordinate structures discussed here, prosodic phrasing can be derived from the way syntactic derivations work by the simple assumption that the output of cycles gets mapped to a foot in prosodic representations. Prosodic boundary ranks reflect syntax in a direct fashion. There is a one-to-one mapping between syntactic derivations and grid representations. Instead of viewing the grid as a separate representation, completely independent of syntax, we can see it as another way of representing syntactic information.

This is as expected according to the program for a metrical theory originally outlined in Liberman (1975, 258): "Thus the most basic assumptions of our theory depend on the idea that the phonological component is not so much a destruction of structure, which maps a complex tree onto a simple serial ordering of segments, as a transmutation of structure, which maps a structure suitable for operations in one domain (syntax and semantics) onto a structure suitable for use in another domain (the motor control of articulatory gestures, and its perceptual analogue)."

While it is useful to think of the complete grid as an output of the grammar in order to state phonological generalizations such as phrasing and accent placement, it is not obvious whether or not the complete grid of a derivation is ever represented at a single point it time in the mind of a speaker. It is conceivable that less information is kept after spell-out, e.g. it might be that the only information kept is the relative rank of the current foot boundary relative to the starting point.

The presentation here assumes the minimalist framework, the presented recursive theory could be implemented indifferent way. For example, the findings lend themselves easily to be implemented in tree-adjoining grammar. The recursive steps necessary for this limited data set could be taken to be substitution operation, and the restriction of right-branching could be implemented as a restriction elementary trees in TAG. I will not explore the compatibility/incompatibility of the findings with alternative theories of syntax.

The discussion was entirely based on evidence from the domain of coordinate structure. Of course, it might be that the generalizations about carefully articulated coordinate structures are a special case and do not tell us anything about the interaction between syntax and prosody more generally. In chapter 3 argues, however, evidence is presented that the analysis extends to other domains the grammar.

The proposal in this chapter 2 captures some aspects of the prosody of coordinate structures, namely those that relate to relative boundary ranks. Consider the following constraint on the prosody of coordinate structures proposed in Taglicht (1998, 192):

(99) *Coordination Constraint (CC)*

For any coordinating node X, if any two daughter nodes of X are separated by an IP boundary, all the daughter nodes of X must be separated by IP boundaries.

This constraint is captured by the metrical representation derived here: it cannot be the case that some boundaries between conjuncts at a given level of embedding in a coordination structure are intonational phrases and others are not, since this would entail that they are not separated by boundaries of equal rank. The prediction made here is even stronger: the boundaries separating the conjuncts must always be the of equal rank.

A question left open is how exactly the impressionistic boundary ranks are implemented phonologically and reflected in the phonetics of coordinate structures. For example, the lines in the metrical grid derived here are not labeled with respect to the intonational categories (intonational phrase, phonological phrase, etc.) which are standardly assumed in prosodic phonology (Selkirk, 1978, 1980, 1981, Nespor and Vogel, 1982, 1986). The representation derived here is an *unlabeled* grid, following the proposals in Liberman (1975), Selkirk (1984), Halle and Vergnaud (1987) and other works.

The theory does not make any predictions about which types of boundaries are actually used. In chapter 4 phonetic and phonological evidence is reviewed and an argument is presented that it is a desirable result that the syntax—phonology mapping only derives a relational prosodification, and does not provide prosodic labels.

The proposed generalization about coordinate structures is recursive in nature, and invokes *relative* boundary ranks. The mapping of the relative boundary ranks to phonological categories is a matter of implementation, and depending on speech rate and the place of a phrase marker in bigger syntactic context there is substantial variation as to how exactly the relative ranks are realized.

The alternative to this recursive and relational view would be a theory of syntax-to-phonology mapping that ties certain prosodic boundaries to certain prosodic categories (such as the intonational phrase). However, the data discussed in this paper involved recursion of the *same* syntactic category—Determiner Phrases. In other words, if the different ranks are realized with different prosodic categories, hence the source of these prosodic categories cannot be differences in the syntactic category of the objects that are coordinated.[36]

[36]Barry Schein points out that it would be conceivable to propose that the different prosodic

A theory that relates prosodic domains of a certain category with certain syntactic categories runs into problems in face of the data pattern observed in this chapter. A common way of implementing the idea of designated syntactic categories for prosodic domains is edge-marking or edge-alignment (Selkirk, 1986, Chen, 1987, Hale and Selkrik, 1987, Selkirk, 1995b, Truckenbrodt, 1995, 1999). Here an instantiation from Selkirk (1996, 444):

(100) The edge-based theory of the syntax-prosody interface

Right/Left edge of $\alpha \rightarrow$ edge of β,

---

breaks in coordinate structures reflect the size of elided material, instead of being a reflex of recursive nesting. The idea is that in (1b), the first conjunct—'Lysander'—is really an entire sentence 'Lysander is in the forest', or at least some bigger structure compared to (1a), while the second conjunct consists of two smaller conjuncts that only contain the material of the two DPs (and their event structure in Schein's theory of event semantics):

(1) a. Lysander | and Demetrius | and Hermia went to the forest.

b. Lysander || and Demetrius | and Hermia went to the forest.

The strength of a boundary is correlated then with the size of the elided material. In this theory, it would indeed be possible to relate the strength of boundaries to differences in syntactic category (e.g. DP-coordination vs. CP-coordination).

It is entirely possible that some strong boundaries are a reflex of more elided structure. But it is not clear that this alternative hypothesis can explain all the data reported in this chapter. For example, it seems that even a strong boundary can be used in cases where only DPs are being coordinated (1a) should not have allow for a collective reading just like (2b):

(2) a. Lysander | | and Demetrius | or Hermia gathered in the forest.

b. * Lysander gathered in the forest and Demetrius and Hermia gathered in the forest.

At least it is clear from the contrast with (2b) that the elided constituent is an entire IP. It is also not clear how such an alternative could cover the entire range of data regarding multiple levels of embedding, since it is unclear how to make the necessary distinctions when there is an additional level of embedding:

(3) Lysander ||| and (Demetrius || (or Hermia and Helena).

α is a syntactic category, β is a prosodic category.

The edge of certain syntactic types of constituents, e.g. XP, are mapped to the edge of a certain type of prosodic constituent, e.g . phonological phrases ('Φ').

Each conjunct in a coordinate structure involving names should be mapped to the prosodic category corresponding to XPs (maybe by XP-alignment). Say the designated category is the phonological phrase, Φ. But this would predict that the phrasing is always flat:

(101) * (Lysander)$_{\Phi}$ (and Hermia)$_{\Phi}$ (or Demetrius.)$_{\Phi}$

The hierarchical prosody that is actually observed constitutes 'XP-within-XP' effects. This point will be discussed in more detail in chapter 4. The relational view of syntax-to-prosody mapping is further confirmed by experimental evidence. The experiment and some evidence bearing on how the relational grid is implemented in on-line speech production is presented in chapter 5.

2.5 Appendix

2.5.1 Binary or Infinitary Connectors?

In our binary notation, we used the two labels XP vs. $\bar{X}$ to distinguish nodes that form their own cycle from nodes that are part of a bigger cycle. The proposed generalization was that nodes that are combined in such a way that the associative law holds can be computed in a single cycle.

But there is an alternative to this approach, which assigns the labels XP vs. $\bar{X}$ in a different way. We could instead treat the XP-nodes as projections of heads, which define their own cyclic domain. This has been proposed for 'flat' coordination in English, which were argued to constitute &Ps (also called Boolean Phrases or ConjPs), headed by a Boolean head (Munn, 1987, Collins, 1988, Zoerner, 1995). In the following, I will give evidence against this alternative.[37]

[37] An updated version of this account could be to posit &Ps and stipulate that they are spell-out

An observation that is often taken (cf. Progovac, 2003) to speak in favor of the &P approach is that in cases of iterated coordinated and disjunction, not all coordinators have to be pronounced:

(102) a. Lysander, Demetrius, and Hermia.

b. Lysander, Demetrius, or Hermia.

This occurrence of only a single 'and'/'or' is unsurprising under the &P-approach, since there is only one head involved in the first place.

Of course, why the functor surfaces where it does is not obvious under this approach. Also, every connector in a flat structure *can* be pronounced, and there is no reason to assume that the two types of constructions differ radically in phrase structure or complexity. E.g., Frazier et al. (2000, 360) report results from an experiment comparing coordinate structures with elided and unelided connectors, and found no difference in reading times. Under the &P-story, the instances of 'and'/'or' in addition to the final one would have to be analyzed as some kind of coordinate concord.[38]

Another fact that could be taken to point to an &P analysis is that adverbials cannot take scope over two conjuncts unless they are linked by an overt connector. The argument would be that sentential adverbs cannot be placed inside of &Ps, just as they cannot usually be placed inside of some other XPs such as DPs.

(103) a. Maybe Lysander, probably Demetrius, and definitely Hermia.

b. * Maybe (Lysander Demetrius), and definitely Hermia.

Before looking more closely at the premise of those arguments, it should be noted that these data are not necessarily in conflict with the binary approach. Faced with (102), the approach that takes all connectors to be binary can invoke some notion of

domains or 'phases'. We would then have to make some stipulation that coordinate structures with arbitrarily many arguments are 'spelled-out' as a unit, maybe by stipulating some higher head which constitutes a 'phase' and triggers the spell-out of its complement.

[38]According to Zoerner (1995, 27), the optional pronunciation of the connectors may be used to indicate some degree of emphasis.

phonological deletion or backwards 'gapping' of the connectors. It is not inconceivable that this alternative could be made to work.

The observation in (103), for example, can then be explained by requiring that in order for gapping of 'and'/'or' , the connectors *have to follow boundaries of equal rank*. We already observed above that adverbials induce prosodic boundaries around the conjuncts that they take scope over, so the connectors in (103) could not be gapped for prosodic reasons under this approach:

(104) Maybe (Lysander | and Demetrius), || and definitely Hermia.

The boundary preceding the last 'and' is stronger than the boundary preceding the first 'and'. The arguments for the &P-approach over the binary approach are thus not conclusive.

The premise of the argument based on omitted connectors is that all the elements combined under a single &P projection are united by the same functional head, be it 'and' or 'or'.

However, it is not clear that this premise is correct. The next section argues that it is incorrect. The section after that proposes an alternative account for structures with 'omitted' connectors, involving an unpronounced set-union operator '*', which voids the need for infinitary functors.

Omitted Connectors are Different Connectors

The analysis within the &P approach requires conjunctive heads that can take any number of arguments. Suspiciously, coordination and conjunction seem to be the only functors that allow infinitary arguments.

Predicates in general do not allow this, e.g., there is no verb in English that has this property. An example would be a version of 'precede' or 'is-an-ancestor-of' that would take an arbitrary number of arguments and denote a completely ordered chain or a completely ordered line of ancestors. The following sentence means that A, B, and C each precede D, but not that the order is necessarily $A \prec B \prec C \prec D$:

(105) A, B, C precede D.

There are of course ways of expressing a complete ordering, but they involve coordination, and in ordered to obtain a completely order chain of precedence meta-linguistic reference to linear order with the adverbial 'in that order' is necessary:

(106) A, B, C and D, in that order, precede one another.

It seems that (105) involves a coordination structure without any overt coordinator at all, and maybe it does not sound quite complete without any overt 'and'. But certainly the zero connector involved is not the predicate 'precede'.

There is more evidence that there are zero connectors that figure into a structure similar to (102), but are clearly not derived by combining n arguments by the same operation.

The functor 'through' is an example. 'Through' cannot be used with items that are adjacent on a list (107a), and yet (107b) is grammatical (Assuming that in the context the relevant domain is the alphabet, and A directly precedes B). The zero-functor involved cannot be 'through' itself.

(107) a. * A through B.

b. A, B, C through Z.

If it is not the functor 'through' that connects the other arguments, there must be some covert connector that does. Another similar case is 'but', which is used instead of 'and' in contexts where the two adjacent conjuncts differ in polarity (Horn, 1996).

(108) Who went to the forest?

Lysander, Demetrius, Helena, but not Hermia.

Also, there are structures that show zero connectors without any overt functor at all, called asyndetic coordination (as opposed to syndetic coordination):[39]

(109) With bracelets of thy hair, rings, gauds, conceits,
knacks, trifles, nosegays, sweetmeats

[39] Barry Schein points out that this asyndetic coordination could be the deliberate rhetorical use of an incomplete, interrupted coordination. This could apply to all cases of asyndetic coordination in English. I am not sure how to rule out this possibility.

—messengers of strong prevailment in unhardened youth.

(A Midsummer Night's Dream, I.1 36–38)

Another argument is related to prosodically conditioned allomorphy in coordinate structures. Zoerner (1995, 28) observes that if all coordinators are realized, they have to be identical in form. The judgements presuppose a flat prosody:

(110) a. Robin | and Kim | and Terry.

b. Robin'n | Tim'n | Terry.

c. * Robin | and Kim'n | Terry.

d. * Robin'n | Kim and | Terry.

The prosodic allomorph of 'and' is used to encode the strength of the following boundary. Since in 'flat' structures the conjuncts are separated by boundaries of equal rank, the allomorphs have to be identical. Relating the allomorph of the connectors to the strength of the following boundary can be further motivated by the fact that it is the lowest level of embedding that takes the prosodically weaker allomorph, and not vice-versa:

(111) a. ?? Robin and | Kim'n || Terry.

b. ?? Robin'n || Kim | and Terry.

The allomorphs of 'and' are thus used in a way that reflects the grammatically assigned boundary rank. The only exception to this generalization about the distribution of allomorphs of 'and'—phonological identity at boundaries of equal rank—would be the case were the last functor is realized and all the others remain unpronounced.

Finally, coordinate structure with overt connectors are not identical in their grammatical properties to those with omitted connectors. To illustrate this, an excursion into NPI-licensing is necessary.

NPI-Licensing

Consider the following example:[40]

[40]Progovac (2003) erroneously claims that NPI-licensing is not possible in coordinate structures. One example used as evidence involves the conjunction 'and', which is an intervener for NPI-licensing,

(112) Did they find anything suspicious?

a. No weapons or any drugs or any other sign of a crime.

b. * Any weapons, or no drugs, or any other sign of a crime.

NPIs are licensed from the first disjunct into later disjuncts, but not vice-versa, as expected under the right-branching account for prosodically flat coordinate structures outlined here.

There is an interesting contrast, however, depending on whether 'or' is pronounced: While a negative quantifier in the first disjunct can license NPIs in later

---

as discussed in the context of example (1).

(1) * He chased nobody and any dogs.

The connector to use to test for NPI-licensing is 'or', which is not an intervener. Using 'or', however, does not improve the example much according to Progovac (2003)—but this example might be odd for a number of reasons. A confounding factor is that it is marked to use a negative quantifier in direct object position, instead of using sentential negation. A further problem might be that 'nobody' instead of 'no human being' is used to contrast with 'dog'. NPI-licensing is fine in the following case:

(2) No cat or any dog was chased.

There are also many naturally occurring examples (from Google, and then tested on native speakers) with NPI licensing in coordinate structures in direct object position:

(3) a. I could get no help, or any response from the on-line customer service.

b. ... and enclose no money or any checks at this time.

I started my own print magazine

in this town and got no help or any replies from anybody in this town.

The following example illustrates that 'nobody' can also license NPIs (again a naturally occurring example found with Google, then tested on native speakers):

(4) He warned that nobody or any amount of intimidation would stop the community from pursuing its cultural heritage.

disjuncts (113a), is impossible to repeat the negative quantifier instead of using an indefinite (113b).

(113) a. No weapons, or any drugs, or any money was found at the site of the crime.

b. * No weapons, or no drugs, or no money was found at the site of the crime.

The facts dramatically change when we drop 'or'. It turns out that NPI-licensing is only possible into a disjunct that is headed by an *overt* 'or':[41]

(114) a. ?* No weapons, any drugs, or any money was found at the site of the crime.

b. No weapons, no drugs, or any money was found at the site of the crime.

The pattern observed between the first two conjuncts in (114a,b) is identical to that observed in the case of coordination:[42]

(115) a. * No weapons and any drugs were found at the site of the crime.

b. No weapons and no drugs were found at the site of the crime.

The explanation for the ungrammaticality of (115a) offered in Linebarger (1980) is that 'and' is an intervener and blocks NPI-licensing. Another way of looking at it is to say that 'and' is a PPI. If this is the correct explanation, (114) raises the following puzzle: if 'or' is not an intervener, why is the zero version of 'or'?

[41] Or alternatively the negative quantifier can be altogether omitted—the use of 'any' is of course also not obligatory in the last disjunct either.

[42] Schein (2001, 12) reports some examples in which an NPI can be licensed within a coordination:

(1) No wolf mother and any of her cubs remain close for very long.

(1) is at least much better than (115). This seems to be related to the fact that the negative quantifier quantifies over all possible mother/cub pairs, rather than the set of the mother and all of the cubs.

In sum, when looking at conjunctions/disjunctions, it is doubtful whether the zero-connectors are really the equivalent of the overt connectors. Thus the premise for the argument for n-ary conjunction and disjunction is dubious at best.

A List Operator

Another way to look at the 'dropped' connectors is that the connector that relates the conjuncts in the absence of an overt 'and'/'or' is *neither* of those two, but instead is a set-union or list operator '*' which combines elements into a single set.[43] The operator '*' connects two arguments by set union and is involved in the creation of lists:

(116) [[*]] (A)(B) = {A, B}

We can decompose 'and' into two parts: a universal quantifier, and a set-union connector '*'; similarly, we can decompose 'or' into an existential quantifier, and the connector '*'. This was suggested to me by Marta Abrusan. For a decomposition of disjunction into an existential quantifier and a set union operation see Alonso-Ovalle (2005), who proposes that the set that disjunctions create can be bound by various different operators, in analogy to the case of indefinites in the approach of Heim (1982).

In order to get the right meaning for both 'and' and 'or', we have to be able to let different quantifiers take scope over the conjuncts:

(117) a. A and B: $\forall x.x \in A * B$

b. A or B: $\exists x.x \in A * B$

The connectors 'and' and 'or' are allomorphs of '*'. They they agree with a c-commanding quantifier. The precise conditions on when 'and', 'or' and $\emptyset$ and whatever other possible agreement forms there may be (e.g.'nor') are intricate and I cannot discuss this issue at this point.

[43]At least we follow Schwarzschild (1996) in allowing individuals to be represented either as such or as sets containing themselves we can treat '*' as simple set union.

(118) a. A, B, and C = (A*B * C): $\forall x. x \in A * B * C$

b. A, B, or C = (A*B * C): $\exists x. x \in A * B * C$

The operator '*' itself is associative and binary. It creates sets. These sets are given in the form of lists, with a right-branching structure and an iteration of binary '*'. Lists without overt connectors in general can combine with predicates. The meaning in those cases is just as what one would expect if they are combined by the list operator '*'.

The following expression states that each member of the list precedes 'Z' and expresses a partial ordering, but it does not denote a precedence chain with a complete ordering:[44]

(119) A, B, C, F, E, D precede Z.

The reason why (119) might in general be dispreferred compared to an alternative with an overt 'and' and 'or' might be that it leaves open how to quantify over the set created by the list:

(120) bracelets of thy hair, rings, gauds, conceits, knacks, trifles, nosegays, sweetmeats.

The approach using '*' voids the the necessity of functors with infinitely many arguments and voids the need to commit us to treat zero-connectors and and/or as identical, both of which are problematic assumptions.

The account using '*' *explains* why there are only a few items that appear to be able to take any number of arguments: they are all pronunciations of '*' that agree with different operators that can quantify over the set. The inventory of operators is a

[44]'etc.' could be seen as a function that takes a single argument which be a list. The following would then be another instance of a list without overt connective:

(1) A, B, C, etc.

Zoerner (1995) offers an alternative interpretation, and interprets 'etc.' as a pronoun standing in for '[and X]'.

closed set, and probably very small. Cross-linguistically, there is quite some variation in how connectors can be pronounced, but the inventory does not seem to be involve open class items (Haspelmath, 2004).

2.5.2 A Pragmatic Principle 'Avoid Ambiguity'?

Sequences of 2+ elements in which the associative law does not hold are ambiguous. Inserting prosodic boundaries to signal the bracketing disambiguates these structures. Grice (2005) proposes the principle 'Avoid Ambiguity' as one of his maxims of manner. Is prosodic disambiguation due to a pragmatic principle?

The definition of level of embedding in terms of the associative law is formulated irrespective of whether or not in a particular utterance context there is any danger of an ambiguity. The reason is that under the pragmatic approach one might expect that sufficient contextual information would void the need to realize prosodic phrasing. This would mean that (38) is too weak, since it excludes the possibility of having a prosodically flat structure in cases of non-associative combinations, where information from the context disambiguates the structure. But evidence from the literature of prosodic disambiguation shows that disambiguating prosodic cues are placed *regardless of whether or not the context can be used to disambiguate a structure.*

In natural language it is less common that each conceivably possible bracketing is equally likely, as is often the case in logical formulas. Orthography illustrates that it is possible to get rid of prosodic information and still communicate—omitting disambiguating parentheses in logical formulas can easily make the content unrecoverable. In orthography, prosodic information is at least minimized and only sporadically encoded with punctuation.[45]

[45]In fact, according to (Nunberg, 1990), punctuation obeys its own textual grammar and cannot be conceived of as a deteriorated version of spoken prosody (certain correlations notwithstanding, as reflected by the term 'comma-intonation' often used in the context of the intonational realization of certain modifiers, parentheticals, and appositives (e.g. Potts (2005)). See Nunberg (1990, 39) for the use of punctuation to disambiguate coordination structures.

Orthography thus shows that prosodic structure *is* not necessary to be able to communicate, and that at least in writing we convey meaning without an explicit prosodic structure. The question is whether in speech, given a context that would disambiguate a structure, prosody is still produced by the speaker.

Several authors have suggested that prosody might only be used in cases where a speaker is a aware of an ambiguity and intends to disambiguate a structure Lehiste (e.g. 1973, footnote 1). Some authors proposed an 'audience design' model of prosodic disambiguation, in which speakers use prosodic cues for disambiguation in order to help the adressee in decoding the message (see Krajic and Brennan, 2005, for a summary). Some experiments report that prosodic cues are only inserted when an ambiguity exists (Allbritton et al., 1996, Snedeker and Trueswell, 2003), which could speak in favor of a pragmatic account.

But experiments that create situations in which speakers actually interact with an addressee—in other words, experiments that try to create a natural communicative situation—show that prosodic cues are placed independently of whether or not context helps to disambiguate structure and irrespective of the needs of the addressee in the particular situation (Schafer et al., 2000, Krajic and Brennan, 2005).[46]

In the experiment on coordinate structures reported in chapter 5, ample context was provided, but prosodic phrasing was realized nonetheless. It was not a very natural context, and the subjects knew their interlocutor did not have to understand the message based on their utterance.

Therefore, I will assume in the following that the generalization about prosody and associativity is a grammatical generalization that holds true independent of context. The prosodic boundaries discussed here are a result of the way grammar assembles structure, and a reflex of the syntactic derivation. A cycle is induced if the logical properties of the elements involved is such that the associative law does not hold—

[46]Krajic and Brennan (2005) suggest that subjects in the experiment in Snedeker and Trueswell (2003) have fallen into 'routine, uninformative prosody'. The results of the production experiment in chapter 5 suggest that prosodic structure can be 'washed out' in such a way that sub-distinctions between weaker boundaries are not realized.

whether or not a speaker is aware of an ambiguity.

The view that takes prosody to be shaped by syntax is taken in much of the literature on the phonetics of prosody (Klatt, 1975, Cooper and Paccia-Cooper, 1980, Gee and Grosjean, 1983, Ferreira, 1988), and also the standard assumption in the literature on phrasal phonology in generative grammar (Selkirk, 1984, Nespor and Vogel, 1986, Inkelas and Zec, 1990, Cinque, 1993, Truckenbrodt, 1995).

Chapter 3

Prosody, Scope, and Associativity

Abstract

In chapter 2, a cyclic algorithm is presented to derive the prosody of coordinate structures. In domains in which the associative law holds, a flat prosody is assigned; otherwise the prosody reflects the syntactic/semantic bracketing. This chapter presents evidence that the results obtained for coordinate structures carry over to other domains of the grammar.

3.1 Prosody in Coordinate Structures

Coordinate structures sometimes receive a 'flat' prosody, in which each conjunct is prosodically on a par, and sometimes a more articulated prosody, that reveals the syntactic bracketing as either left- or right-branching:

(1) a. $A \vee B \vee C$ 'Flat' Prosody: A | ∨ B | ∨ C

b. $A \vee (B \wedge C)$ 'Articulated' Prosody: A || ∨ B | ∧ C

c. $(A \vee B) \wedge C$ 'Articulated' Prosody: A | ∨ B || ∧ C

In chapter 2, a cyclic mapping between syntax and phonology is presented to account for the prosody of coordinate structures.

The goal of this chapter is to show that the generalizations about prosody in coordinate structures carry over to other domains of the grammar. The following section summarizes the relevant properties of the theory of the syntax—phonology mapping proposed in chapter 2. Section 3.2 discusses associative domains other than coordinate structures, and shows they have a 'flat' prosody; section 3.3 shows evidence from domains that are not associative and in which scope crucially matters; and finally section 3.4 illustrates that some mismatches between prosody and syntax reported in the literature are only apparent and upon closer scrutiny give more evidence for the theory of syntax—prosody mapping developed in chapter 2.

3.2 Associative Domains

All and only structures with 'flat' prosodic structures are predicted to be associative domains, that is domains composed in a single cycle. The associative law holds when 2+ elements are combined in such a way that changing the bracketing would not change the meaning:

(2) Associative Law

$(\llbracket x \rrbracket)\llbracket y\ z \rrbracket = \llbracket x\ y \rrbracket\ (\llbracket z \rrbracket)$

The example discussed in chapter 2 are coordinate structures. This section presents additional examples of domains that are associative. Just as expected, they have a 'flat' prosody and are right-branching.

3.2.1 Predicate Sequences

Sequences of predicates in English appear to be right branching and yet they have a flat prosody. In the following example, the complement of each predicate is exactly everything that follows it:

(3) She wanted | to try | to begin | to plan | to move.

Each predicate can (but not need not be) accented. One factor in which predicate actually is pronounced with an accent is rhythm, as will be discussed in chapter 6. The prosodic grouping is similar to a coordination of predicates:

(4) to want, to try, to begin, to plan, and to move.

The predicate sequence receives a 'flat' prosody, i.e. the boundaries between the predicates are perceived as being equal in strength. Given the flat prosody, the expectation is that the domain is associative. But how can we interpret the predicates when the bracketing is changed?

A simple way to think about how predicates can be interpreted given a different bracketing is to think about the rebracketing in analogy to movement. E.g., the VP containing the lower two predicates can rightward move, λ-abstraction makes the the re-bracketed structure interpretable:

(5) (λx. wanted to try to begin x) ([$_{VP}$ to plan to move]).

This VP-movement leaves a remnant expression, and predicate abstractions turns it into a one-place predicate. The denotation of the final expression not changed by this restructuring, since the denotation of the moved VP semantically reconstructs to the complement position of 'begin' due to the λ-abstract. In other words, the law of association holds:

(6) $[\![A(BC)]\!] = [\![(AB)C]\!]$

In principle, this renders any predicate order interpretable:

(7) (λx. wanted x to begin to move) ($[_{VP}$ to try).

But of course this is not a grammatical word order for the sentence:

(8) *She wanted to begin to move to try.

The reason (8) is not grammatical is the movement step is ruled out by the constraints on rightward movement (Sabbagh, to appear).

That remnant sequences of predicates can be interpreted as one-place predicates can be motivated based on a number of grammatical phenomena. It is useful in the analysis of right-node raising. Consider first a simple case of coordination of predicates:[1]

(9) She wrote | and defended || her dissertation.

The prosody of this sentence suggests that the constituent 'her dissertation' attaches outside of the coordinate structure 'wrote and defended'. The boundary preceding the direct object is stronger than the prosodic boundaries within the conjunction, and is thus analogous to that of a left-branching coordinate structure:[2]

[1] It is not clear whether the boundary right before the right-node raised constituent has to be stronger than the one separating the conjuncts, but there is a clear intuition about a strong boundary.

[2] This is the case at least when the right-node raised constituent is accented. Hartmann (2001) observes cases were the right node constituent clearly is not preceded by a strong boundary. I assume that this is only the case when an unaccented constituent is right-node raised. This could be a pronoun or a head:

(1) a. John sáw | and Mary mét him

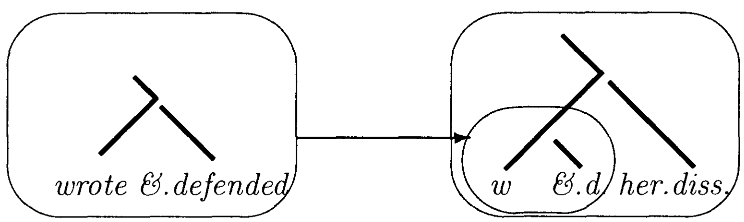
b. John had his glásses | and Mary had her lénses replaced

Hartmann (2001), Selkirk (2002), and Selkirk (to appearb), on the other hand, argue that even right-node raised constituents phrase with the material in the second conjunct. The predictions of this view with respect to durational effects as their are induced by boundary strength clearly differ then from the proposal here, and can be easily experimentally tested. This is an interesting question for future work.

(10)   (A | and B) || or C

If the prosody in (9) is derived in the same way as that in (10), the coordinate structure 'wrote and defended' in (9) must constitute its own cycle before combining with its argument. The two functions are coordinated to create a conjoined function, which in the next cycle takes the direct object as its argument.

(11)   Two Cycles



The first cycle is semantically interpretable: the remnant predicates form a complex predicate by λ-abstraction. The prosody furthermore suggests that in combining the coordinate structure with the direct object, the associative law does not hold, since the prosody clearly marks the structure as left-branching. Indeed, with a different bracketing, the meaning changes:

(12)   She wrote || and defended | her dissertation.

This example involves a coordination of two VPs, one involving a predicate 'wrote' with an implicit argument and the other involving the transitive verb 'defend' and its object. Clearly, (9) and (12) differ in meaning.

Now consider the following example, an instance of right-node raising, analyzed as rightward movement in (Ross, 1967, Sabbagh, to appear):

(13)   She wanted to begin | and then decided to postpone || her dissertation.

The prosody of this sentence suggests that the constituent 'her dissertation' attaches outside of the coordinate structure 'wanted to begin but then decided to postpone'. Again, the coordinate structure should constitute its own cycle. But this would imply that two apparent non-constituents are coordinated, namely 'wanted to begin' and 'then decided to postpone'.

In order for the coordinate structure to be semantically interpretable, the predicates have to be assigned a meaning, despite the fact that in each conjunct the direct object is missing. The idea is that they become one-place predicates by virtue of λ-abtraction:

(14) ((λx. wanted to begin x) and then (λx. decided to postpone x))

The entire structure in (14) constitutes a cycle and is combined with the object 'her dissertation' in a second cycle. Various different bracketings are in principle possible:

(15) a. She wanted | to try | to begin | to write.

b. She (wanted | to try | to begin) || to write.

c. She (wanted | to try) || (to begin | to write).

d. She (wanted) || (to try | (to begin | to write).

Just as in the coordinate structures discussed in chapter 2, the expectation is that they are only felicitous when the extra bracketing is motivated by information structure:

(16) She wanted to try to begin to write?
No, she wanted to try to begin || to procrastinate.

A sequence of predicates in which each predicate is the argument of the preceding one forms an associative domain, and is usually assigned a flat prosody. Flat prosodies are also observed in the domain of adjectives, which also can be viewed as sequences of functors:[3]

[3]But not all functor sequences are such that the emerging prosody is flat. Consider the difference between the following two structures:

(1) a. a (very | likely | false) || story. (a story that is very likely false)

b. a (very | likely) || false || story. (a story that is very likely but false)

When there is a left-branching node, the prosody of predicate sequences in (1) is more articulated.

(17) A dark | old | forest.

The combination of function application and λ-abstraction renders sequences of functors into associative domains. Just as in associative domains in coordinate structures, they receive a 'flat' prosody in which the elements are separated by prosodic boundaries of equal strength.[4]

3.2.2 Additional Arguments

Sentences can contain a variety of nominal arguments, receiving different thematic roles. Some of them are arguments of the main verb of the sentence, others are introduced by other heads; they are 'additional' arguments that are not part of the argument list of the main predicate. Consider the following sentence:

[4]An alternative view to the one involving variables and λ-abstraction is the one taken in categorial grammar. Steedman (1985, et. seq.) proposes that any sequence of predicates can be composed in more than one way. He uses the notion 'functional composition', an operation of meaning composition used in categorial grammar in addition to 'functional application'. In particular, Steedman uses a rule of forward function composition (Steedman (2001, 40), Steedman (2004)).

(1) Forward composition (simplified)
$X/Y : f \;\; Y/Z : g \Rightarrow X/Z : \lambda x.f(gx)$

Function composition, from now on abbreviated by the symbol '$\circ$', composes two functions f and g into a new function $f \circ g$. The definition makes sure that function composition does not change the meaning of an expression:

(2) $f \circ g(x) \equiv f(g(x))$

Function application assigns a meaning to a sequence of three elements A, B, and C if the meaning of C is in the semantic domain of B and B(C) is in the domain of A. Function composition allows the assignment a meaning to a rebracketed structure. This rebracketing does not alter the truth conditions of the outcome, i.e. , the associative law holds. Another domain in which forward composition is used in categorial grammar are predicate reorderings (cf. Steedman, 1985, 2000), to be discussed in more detail in chapter 6. The analysis of right-node raising in categorial grammar also involves forward composition (Steedman, 1985, et.seq.).

As observed in Sabbagh (to appear), the approach in terms of 'forward composition' in categorial grammar shares many properties with rightward movement accounts.

(18) Lysander baked a cake for Hermia.

Many theories treat the direct object as an argument of the verb, and the benefactive argument as being introduced by a different functional head. Part of the reason for making this difference is the intuition that a benefactive can be added to just about any sentence that involves an agent, while the thematic role of the direct object is closely tied to the meaning of a verb of creation such as 'bake'. Another reason for treating the benefactive as different is the fact that it is an optional argument.

The status of the subject in (18) is more controversial. Some theories of argument structure treat agents just like the direct object as an argument of the verb (e.g. Bierwisch (1983), Grimshaw (1990)); but there are also theories that treat agentive subjects as being introduced by a separate functional head, analogous to the benefactive.

One piece of evidence that subjects are not arguments of the main predicate is that the thematic role of the subject is not fixed by the verb, but depends on the combination of the verb and the direct object it combines with (Marantz, 1984). Schein (1993) presents a semantic argument for this view and argues that subjects are the argument of a separate event predicate. More arguments that agentive subjects are indeed introduced by a functional head other than the main predicate are discussed in Kratzer (1996) and Pylkkänen (2002).

The alternative to viewing an agentive argument as an argument of the main predicate is to assume an event predicate that takes it as an argument. This event predicate relates an event argument and a nominal argument, and was pursued within the 'Neo-Davidsonian' approach (Parsons, 1990) to argument structure. According to Kratzer (1996) and Pylkkänen (2002), agentive subjects are arguments of an inflectional *voice* head. Kratzer (forthcoming) furthermore presents evidence that an analogous analysis of the theme argument is not warranted, and direct objects should be treated as arguments of the verb.

I follow Larson (2005) in assuming that *voice* and any other Davidsonian predicates are tied together by existential closure, which binds the event variable in each of the predicates.[5] The event predicates are then combined by coordination.[6] The meaning of (18) is the following:

(19) $\exists e.[voice(e)(Lysander)\ \&\ bake(e)(cake)\ \&\ BEN(e)(Hermia)]$

A property of this analysis is that the combination of the event predicates obeys the associative law, just as in any other coordination or list structure:

(20) [[(voice(e)(Lysander) & bake(e)(cake)) & BEN(e)(Hermia)]] = [[voice(e)(Lysander) & (bake(e)(cake) & BEN(e)(Hermia))]]

The expectation is then that the elements of the list of additional arguments should be set off by prosodic boundaries that are on a par. This is compatible with speakers' intuitions about the prosody of these structures. Experimental tests, however, would be necessary in order to investigate this question further.

(21) Lysander | baked a cake | for Hermia.

If indeed the separate event predicates form an associative domain, the right branching conjecture predicts that they should form a right-branching structure. C-command evidence indeed shows that there is left-to-right c-command in sentences with several additional arguments:

(22) a. Every guest baked a cake for his host.

b. He reedited every movie for its main actor.

Pylkkänen (2002) identifies a number of arguments that are *not* introduced by a predicate that relate an event and an individual. Instead, they are introduced by a head that directly relates two nominal arguments. A case in point are the two

[5] The alternative would be to tie the predicates together by 'event identification' as in Kratzer (1996) or predicate modification Heim and Kratzer (1998), Pylkkänen (2002), which for the purposes here would not make a difference.

[6] I assume that the unpronounced list operator '*' is in fact involved here, discussed in chapter 2. Whether or not the coordinator is represented in syntax is not so clear: to be consistant with the discussion in chapter 2, I would actually have to claim that right-branching structures that lack functors are simply interpreted as lists conjoined with the set union operator '*'.

arguments in the double object construction in English, and low applicatives more generally.

Prosodically, they are not predicted to be on a par with the other arguments, but should form a domain of their own. I do not have the time and space to discuss the prosody of ditransitives, but refer the reader to Seidl (2000) and McGinnis (2002) for discussion of the prosody of different types of applicative constructions.

3.2.3 VP-Final Adverbs

VP-final adverbials seem to modify the meaning of entire VPs. Below is an example with 4 VP-final adverbials (Taglicht, 1984, 67):

(23) She saw him once, in 1939, outside the Albert Hall, after a concert.

One common analysis of VP-adverbials is that they take scope over the VP material preceding them, leading to a 'right-ascending' structure (cf. Andrews (1983), Ernst (2001)). Under the right-ascending view, VP-adverbials can be conceived of as functors taking VP-meanings as their arguments and returning an element of the same type.

If this right-ascending analysis were correct, then the structure would be left-branching. The prosody is then predicted to look as follows, at least if syntax-prosody mapping works analogous to coordinate structures. Left-branching structures necessarily show a fully articulated prosody:

(24) She saw him once, | in 1939, || outside the Albert Hall, ||| after a concert.

This does not match the intuitions that speakers have about the prosodic phrasing of this type of construction. The intuition about phrasing is that the prosody in the VP is 'flat', and the adverbs are separated by prosodic boundaries of equal rank.

How would the syntax have to look to obtain this prosody, if the theory proposed here is correct? It would have to be a right-branching, i.e. *right-descending* structure, and the domain should be associative, i.e. the bracketing should not matter in terms of the derived meaning. In other words, the structure should be parallel to an iterated coordinate structure.

(25) A, B, C, and D.

The parallel may not be too far-fetched. Taglicht (1984, 67) in fact *does* analyze sequences of VP-adverbials as coordinate structures. Similarly, Larson (2005) proposes to connect VP-adverbs with the VP by coordination, and argues that semantically, adverbs of this class are really event predicates. Just as in the case of additional arguments, they take an event variable rather than the entire VP as their argument. In fact, Davidson (1967) originally introduced event predicates precisely to account for modifiers. The denotation of a VP with several sentence adverbials looks as follows then:

(26) a. kissed Hermia in the forest for an hour.

b. $\exists$ e [kiss(L, H, e) & in-the-forest(e) & for-an-hour (e)].

Event predicates are added successively to the VP-structure. Larson (2005) takes them to be part of a coordinate structure in the scope of existential closure, which eventually ties them together to predicate over the same event. Existential closure is assumed to apply to the entire VP following Diesing (1992).[7]

Since the event predicates form an associative domain, they can combine within a single cycle. This correctly captures the fact that their prosody is flat. Based on the right-branching conjecture, the prediction is then that the structure of the VP should be right-branching.

Larson (2005) argues for a a right-descending analysis of VP structure. One piece of evidence for a right-descending analysis, as originally proposed in Larson

[7]An argument that the class of VP adverbs discussed here are not scope-taking comes from Phillips (2003), who observes in cases where three adverbs are stacked at the right periphery of the VP, the first two can take either scope.

(1) Sue kissed him many times intentionally in front of his boss.

Larson (2005) argues that the reason why the last predicate appears to take widest scope is due to the way the structure is split into a restrictor and a nuclear scope, and relates to focus and information structure. I will not discuss this problem here.

(1988), comes from NPI-licensing, which requires c-command. NPIs are licensed in VP-modifiers by the direct object and by preceding VP-modifiers:

(27) a. Lysander kissed nobody in any forest at any time.

b. John spoke rarely during any of our meetings. (Larson, 2005)

Variable binding also provides evidence for a right-descending structure:[8]

(28) a. Sue spoke to every child on his or her birthday.

b. John spoke during every session for as long as it lasted.

Under the theory of syntax-prosody mapping proposed here, the prosody of sentence-final adverbials is just as expected based on their syntactic structure, at least if the Larson's approach is correct.

3.2.4 Secondary Predicates

A second type of predicate can be analyzed analogously to VP-final adverbials: depictive predicates. Consider first subject-oriented secondary depictive predicates. They are prosodically similar to VP-Adverbs in that they do not induce strong prosodic boundaries in a sequence of predicates, and in that the order can vary.

(29) a. Hermia was dancing, completely drunk, without any fear, unaware of the abyss.

b. Hermia was dancing, unaware of the abyss, without any fear, completely drunk.

Not every order is equally unmarked. This is just as in coordinate structures and lists, where the order of conjuncts is preferred where it matches the order of the occurrence of events, of which event causes which, or other relations or scales that are used to connect the conjuncts in a meaningful way:

[8]More evidence from variable- and anaphor-binding is discussed in detail in Pesetsky (1995), Phillips (1996). Pesetsky (1995) also argues for a co-present second, right-ascending structure. I will not discuss the issue here, but refer the reader the discussion in Phillips (2003) and Larson (2005).

(30) a. open Monday, Tuesday and Friday.

b. big, bigger, biggest.

Pylkkänen (2002, 27) observes that depictive predicates also share semantic properties with VP-Adverbs. Similar to VP-adverbs, the state described by a depictive predicate must hold during the event described by the verb, i.e. they can be seen as predicates that attribute a property to an event. That secondary predicates indeed have to be eventive, Pylkkänen argues, is evidenced by the fact that individual-level predicates sound odd in depictive predications (as observed in Geuder (2000)):

(31) *? Hermia was dancing tall.

In Pylkkänen's analysis, depictives differ from VP-adverbs only in that in addition to the event argument, they also have an unsaturated individual argument of type *e*. The event argument, just as in the case of VP-adverbs, is bound by existential closure; the e-type individual variable is bound by a c-commanding nominal argument.

If secondary predicates form a list structure with the VP and form an associative domain, just as VP-adverbs do, then a right-branching right-descending structure is expected. Indeed, subject-oriented secondary predicates are c-commanded by the direct object, as is evidenced by variable binding, suggesting that they are indeed put together with the preceding VP by the default right-branching structure expected in a cycle:

(32) Mary Poppins rewarded every child convinced of his worthiness.

Under this analysis, the structure of the VP can be assumed to essentially that of coordination of predicates, which contain argument variables that are bound by c-commanding elements.

For secondary predicates, just as for VP-adverbs, the bracketing between multiple secondary brackets is immaterial—the associative law holds. Just as expected, a flat prosody emerges; and as expected based on the Right-Branching-Cycle conjecture, grammar imposes a right-branching structure on the domain.

Discussion

According to the argument in chapter 2 the domains that receive a 'flat' prosody in coordinate structures are precisely those in which the associative law holds. This section discussed four other cases of domains with a 'flat' prosody, and argued that the associative law also holds within these domains. The results from coordination carry over to other domains of the grammar.

In the present theory, the explanation for the 'flat' prosody between multiple constituents relates to the way they combine with the rest of the structure. If they are added in such a way that the associative law holds, they can be combined in one cycle, hence the resulting prosody is flat.

3.3 Prosody and Scope

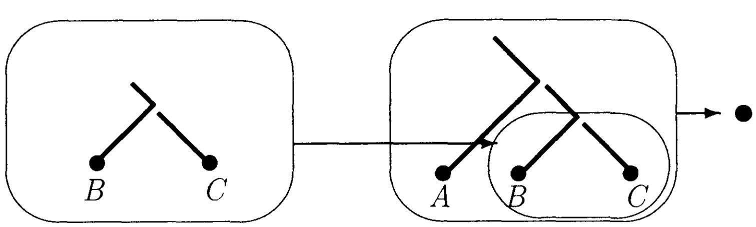
This section presents examples of ambiguities that parallel the pattern observed for coordinate structures in which the associative law does *not* hold. When elements are combined such that associativity does not hold, they take scope relative to each other. Depending on the bracketing, in the following structures 'and' has either wider scope than 'or' or narrower scope.

(33) a. A or (B and C)

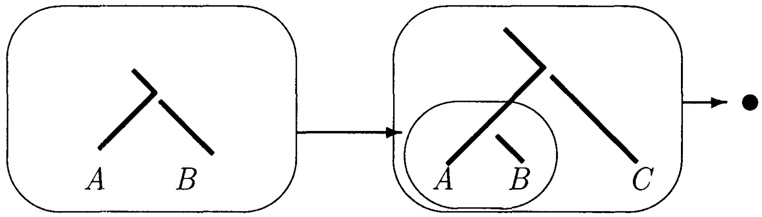
b. (A or B) and C

The break separating the constituent in the higher level of embedding is stronger than any boundary within the lower level of embedding.

(34) a. A || B | C:



b. b. A | B || C:



In other words, within each level of embedding, the conjuncts are separated from each other by boundaries of *equal rank*. If any conjunct is itself complex and contains further nested bracketings, the boundaries within that conjunct have to be of *lower ranks*. The syntactic grouping imposes a ranking of boundaries. The following generalization is derived:

(35) Scopally Determined Boundary Rank (**SBR**):
If Boundary Rank at a given level of embedding is n, the rank of the boundaries between constituents of the next higher level is n+1.

This has an effect on the relative scope of constituents: *An element within the lower level of embedding takes narrow scope relative to elements at the higher level of embedding.*

In the following I discuss evidence that the relation between prosody and scope observed in coordinate structure is also valid in other parts of the grammar.

3.3.1 Quantifier Scope

Taglicht (1984, 129) discusses an ambiguity in the scope of negative quantifiers.[9] I mark the material that negation takes scope over by underlining.

(36) a. I advised him to do nothing illegal, didn't I? *not* $<$ *advise*
I told him not to do anything illegal, didn't I?

b. I advised him to do nothing illegal, did I? *not* $>$ *advise*
I didn't tell to do something illegal, did I?

The tag-questions are just added in order to disambiguate the preceding ambiguous sentence. This section discusses evidence that prosody can (but does not have to) disambiguate these structures as well. The nature of the interaction between prosody and scope in those cases exactly parallels the case of simple coordinate structures.

Blaszczak and Gärtner (2005) observe that if a prosodic boundary intervenes between the quantifier and the higher predicate, the wide scope reading for the quantifier becomes impossible. The example Blaszczak and Gärtner (2005) give is the following:

(37) She has requested || that they read not a single linguistics book.
**not* $>$ *requested*; *not* $<$ *requested*

According to Blaszczak and Gärtner (2005) and Kayne (1998) wide scope is possible in embedded sentences, at least when subjunctive tense is used. The point can also be illustrated by an infinitival case:

[9] Taglicht (1984) also discusses similar facts for the scope of 'only'. According to Taglicht, the following ambiguity holds:

(1) a. They advised him to learn only Spanish. *only* $<$ *advise*

b. They advised him to learn only Spanish. *only* $>$ *advise*

I will only discuss negative quanitifiers here. The ambiguity of negative quantifiers is also discussed in Bolinger (1977, 48). Bolinger notes that taking wide scope is easier with negative quantifiers that he calls 'formulaic', such as 'no such thing', and 'no place at all'. These seem to be maximal elements on a scale, and possibly the fact that taking wide scope becomes easier is a reflex of the fact that negative quantifiers of this sort need to be focused and evoke alternatives.

(38) I will advise him || to do nothing illegal.

Another restriction on scope taking is that the quantifier has to be right-peripheral in order to take wide scope. Kayne (1998, 142) observes that if a negative quantifier is followed by a particle, it cannot take wide scope, but reports that it can when it follows the particle:

(39) a. I will force you to turn no one down.
*$not > force$; $not < force$

b. I will force you to turn down no one.
$not > force$; $not < force$

The analysis that Blaszczak and Gärtner (2005) proopose is that there is a restriction on scope-taking that makes reference to linear order and prosody:

(40) Condition on Extended Scope Taking (CEST)
Extending the scope of a negative quantifier $Q_{\neg}$ over a region σ requires σ to be linearly and prosodically continuous.

The relation between scope and prosody observed in the case of negative quantifiers is similar to that observed with scope ambiguities in coordination structures. A strong boundary following the first conjunct prevents wide scope of the 'or' in the following example:

(41) A || and B | or C *$or > and$; $or < and$

The analysis of the of case of negative quantifiers could be analogous to the one in coordinate structures if wide-scope taking involves a rebracketing:[10]

(42) (advised him to do) nothing illegal

[10]The scope ambiguities and their interaction with syntax and prosody discussed here suggest that scope-taking involves *overt* rebracketing. According to Kayne (1998), *all* scope-taking movement is overt, and covert movement does not exist. I will discuss a problem for this approach at the end of this section.

The prosody should then be parallel to a left-branching coordinate structure involving two cycles, with a boundary separating preceding the negative quantifier that is stronger than any boundary *within* the material it takes scope over.

(43) A | and B || or C

In order to derive the bracketing in (42), a first cycle would have to compose 'advised him to do', and spell it out. A second cycle composes the output of the first cycle with the negative quantifier.

Within the first cycle, the complement of 'to do' is not yet present, but the higher predicate 'advise' that in turn takes the infinitival phrase as its complement is. Blaszczak and Gärtner (2005) and Steedman (2001, 2004) use forward composition, as defined in (1), to analyze the data.

The alternative to use movement and λ-abstraction to explain how the meaning can be compositionally determined after rebracketing. Under either approach, the solution uses the same mechanism independently used for the various bracketings of predicate sequences discussed above.

In (38), 'advise him || to' cannot be a constituent to the exclusion of 'not a single linguistics book'—if it were, the prosodic break preceding the negative quantifier would have to be stronger than the one following the predicate it takes scope over, not weaker, as in the following structure:

(44) I will advise him | to do || nothing illegal.

The prosody in (38) is such that the 'advise him to do' cannot be a syntactic constituent. A similar rationale applies to the linear order restriction illustrated in (39). In order to take wide-scope, 'no one' should out-scope the underlined material, which would have to be composed in a first cycle. But this material does not form a syntactic constituent.

(45) I <u>will force you to turn</u> no one <u>down</u>.

A single cycle cannot comprise linearly discontinuous material, just as it cannot comprise a domain whose prosody reveals that it cannot form a well-formed level of embedding obeying the generalization in (35).

According to Blaszczak and Gärtner (2005), CEST is the restriction that prevents a wide-scope reading. In (38) it does so because a prosodic break intervenes, and in (39) it does so because the linear order condition is not met.

Even if the generalization in CEST holds descriptively, it may not be necessary to treat it as an active constraint in the grammar. An argument in favor of the view that CEST is not the driving factor comes from more complex examples. Blaszczak and Gärtner (2005) claim that a wide-scope reading in the following example is ruled out because there is a prosodic boundary following the relative clause. The CEST predicts that *any* intervening prosodic break should block a wide-scope reading.

(46) She requested that the students who finish first | read not a single linguistics book.

If the prosody reflects syntactic structure in the way proposed in chapter 2, the generalization should be a different one, namely that the break preceding the quantifier should be *stronger than any intervening prosodic break.* The reason is that all the intervening material must be in a deeper level of embedding in order for wide scope to be possible. This, according to the SBR, automatically entails that the prosodic break separating the quantifier from the VP must be stronger.

The following examples involve a relative clause intervening between the higher predicate and the quantifier, and yet wide scope is possible, at least if the prosodic boundaries in the underlined domain are not stronger than the boundary that sets off the negative quantifier. This goes against the prediction of Blaszczak and Gärtner (2005). An example involving a finite embedded clause is the following:[11]

[11] The example in (46) may be not constructed in a way that facilitates the wide scope reading. But the prediction is that if the prosodic breaks indicate the right bracketing, then the reading should be possible:

(1) She requested that the students who finish first read no additional linguistics book, did she?

(47) She expected the students who failed the mid-term to take none of the advanced tutorials, did she?

The tag-question should force the intended reading, and in fact it should also force a prosody that does not have a strong break after 'expected'. The same holds for the following example:

(48) She advised the new students that came in today to do nothing illegal, did she?

The prosodic fact that there can be no strong boundary between the predicate and the quantifier that takes scope over it can be related to the way syntax is mapped to prosody. The analysis presented here is similar to the approach to scope in categorial grammar Steedman (2004), and in particular to the proposal in Blaszczak and Gärtner (2005). It is different from Blaszczak and Gärtner (2005) in that it tries to derive the facts that CEST captures from syntactic bracketing and the way it maps to prosody instead of stating a constraint on interpretation that refers to prosody. Different predictions are made for examples like (47).

The discussion here unites the interaction between scope and prosody in the case of negative quantifiers with the facts observed on the coordinate fragment in chapter 2.[12]

The data discussed here suggests that scope *can* be taken by overt rightward movement (or alternatively by CCG-restructuring), which is directly reflected in prosody. This is not to say that scope *must* be taken by overt movement.

Fox and Nissenbaum (2000) give an argument that points to the possibility of covert QR in cases of negative quantifiers. They show that extraposing an adjunct from the direct object across a purpose clause forces a wide scope reading of the negative quantifier:

[12]Blaszczak and Gärtner (2005) also discuss the antisymmetric approach to the scope ambiguities in Kayne (1998), and propose a revision of this approach that is compatible with the data. I will not consider either proposal here, but refer the reader to Blaszczak and Gärtner (2005) for discussion.

(49) a. John must miss no assignment that is required by his math teacher in order to stay in school.

b. # John must miss no assignment in order to stay in school that is required by his math teacher.

c. John must hand in no assignment in order to stay in school that is required by his math teacher.

Sentence (c) shows that wide scope is possible under extraposition: it has a reading in which John is said not to have the obligation of handing in any assignment required by the math teacher. Sentence (b) shows that wide-scope is in fact obligatory. This sentence is infelicitous under a wide scope reading, but would be felicitous under a narrow scope reading as illustrated in (a); however, this reading is not available here.

The point is that 'no assignment' does not appear to attach higher than 'must' in this sentence, since it is followed by the purpose clause, which is presumably inside of the VP.

The analysis that Fox and Nissenbaum (2000) give this sentence involves covert QR of the negative quantifier and late merger of the adjunct. Covert QR might not be reflected in the prosody of a cycle. It would be interesting to investigate the prosody of these constructions and see how they square with the facts reported here.

3.3.2 Attachment Ambiguities

Taking scope interacts with prosodic boundary strength. Operators within lower cycles take narrow scope relative to operators within higher cycles. The discussion above showed that depending on whether or not disjunction takes wide scope over coordination, the boundary preceding 'or C' is either weaker or stronger than the boundary preceding 'and B':

(50) a. A | and B || or C and < or

b. A || and B | or C and > or

This section discusses data from the literature of on attachment ambiguities, and illustrates that again regularities already observed for coordinate structures apply.

Lehiste (1973) provided evidence that prosody disambiguates attachment ambiguities that relate to adjectives and their relative scope with respect to coordinated noun phrases.

(51) a. the (old men) (and women) the old | men || and women

b. the old (men and women) the old || men | and women

The results showed that the duration of words increased directly preceding stronger boundaries. Other correlates of strong boundaries found in Lehiste (1973) were laryngealization and insertion of pauses. A perceptual experiment showed that listeners are very good at telling the syntactic bracketing based on the prosodic cues involved.

The prosodic contrast in (51) is parallel to the contrast in the difference in relative boundary rank in the following two coordinate structures:

(52) a. A | and B || or C.

b. A || and B | or C.

An mirror image attachment ambiguity was studied in Cooper and Paccia-Cooper (1980, 34ff). Here, a modifier *followed* the constituent it modifies.[13]

(53) a. Lieutenant Baker instructed (the troop with a handicap).

b. Lieutenant Baker (instructed the troop) with a handicap.

The experimental measure in this experiment was the duration of the word preceding the boundary in question. More specifically, the duration of the 'key segments' in each example were measured, in the example above the duration of the segments /tru/ in 'troop' (i.e. the duration of the words excluding the codas). The difference between the structures was found significant.

The boundaries expected by the current approach are compatible with these findings, and are summarized below:

(54) a. Lieutenant Baker || instructed || the troop | with a handicap.

[13] I assume that 'with a handicap' is not an object-oriented depictive predicate as those in (3.2.4), but it a restrictive modifier of the NP 'troop' within the DP.

b. Lieutenant Baker || instructed | the troop || with a handicap.

The boundary following 'troop' in (a) is weaker than that preceding it; the reverse it true in structure (b).

More recent studies (Price et al., 1991, Wightman et al., 1992, Carlson et al., 2001) have shown further evidence that boundaries come in different strengths, and that stronger boundaries induce stronger lengthening effects than weaker boundaries. The relation between syntactic bracketing and prosodic boundary strength is similar to that observed in coordinate structures. The phonology and phonetics of attachment ambiguities in coordinate structures are further investigated in chapter 5.

3.4 Bracketing Mismatches

The prosody in coordinate structures showed a close match between prosody and syntax. Looking at other structures, a number of cases emerge where syntax and prosody seem to diverge in their bracketing.

It is important to note that there is no pre-theoretical notion of a mismatch in bracketing between syntax and prosody. Diagnosing a mismatch requires a theory of syntax that provides criteria to establish syntactic constituency; a theory of prosody that provides criteria for prosodic constituency; and a theory of the general mapping function from syntax to prosody, which provides an expectation how syntactic bracketing should map to prosodic bracketing. A diversion from this expectation constitutes a bracketing mismatch.

Under the theory of syntax—phonology mapping proposed in chapter 2 a structure in which the associative law holds receives a right-branching syntactic structure and a 'flat' prosodic structure:

(55) A, | B, | C | and D.

It seems, though, that the right-branching syntactic structure makes (C and D) a syntactic constituent, but the flat prosody does not make (C and D) a prosodic constituent. Does this not mean that there *is* a mismatch between syntax and prosody in (55)?

A mismatch would only exist if the prosody did not encode the syntactic bracketing, i.e. either encode the wrong constituent structure or not encode bracketing at all. But it is a general property of the mapping proposed here that right-branching structures map to flat prosodies (unless the associative law does not hold). Each domain with a 'flat' prosody' is mapped to a right-branching structure composed in one cycle, and vice versa. It is thus possible to retrieve the syntactic bracketing of (55), and it would hence be misleading to call this type of case a mismatch.

This section discusses what appear to be true mismatches between syntax and prosody, given the assumptions I made about syntax and prosody and the proposed mapping function. Upon closer inspection, the syntactic structure turns out to be just as expected based on the prosody, and what appear to be counter-examples to the theory proposed here turn out to lend it further support.

3.4.1 Coordinate Structures and Extraposition

Across various languages, a mismatch between syntax and prosody has been observed in cases in which a predicate takes a coordinate structure as its complement. The predicate tends to phrase with the first conjunct:

(56) (Predicate A) (and B)

An example is the case of Tiberian Hebrew, discussed in Dresher (1994, 19), who reports that fixed expressions, such as 'good and evil' in (57a), are phrased together. Otherwise, the predicate frequently phrases with the first conjunct (57b).

(57) a. (yōḏʕē) (ṭoḇ wārāʕ)
knowers (of)good and.evil (Gen. 3.5)

b. (kabbēḏ ˀet-ˀāḇīḵā) (wəˀeṯ-ˀimmeḵā)
Honor ACC-your.father and.ACC-your.mother (Deut. 5.16)

Phrasing in Tiberian Hebrew is reflected by phonological processes such as spirantization, which applies to post-vocalic stops within a phonological phrase. The underlined consonants are the ones that have undergone spirantization. But phrasing is also more directly noted in the Masoretic texts by an intricate system of accents, which according to Dresher reflect the prosodic structure of the verses.

The same phrasing seems to be possible in English. While it is possible to realize a stronger boundary after the verb, it is certainly not necessary, and (b) seems more natural:

(58) a. She kissed || Lysander | and Demetrius.

b. She kissed | Lysander || and Demetrius.

One context in which (a) is attested is the following:

(59) She hugged Hermia, and she kissed || Lysander | and Demetrius.

Phrasing a first conjunct with a preceding predicate looks like a genuine bracketing paradox. One possible solution discussed in Dresher (1994) is that the asymmetry could be related to XP-edge marking. The idea is that only left or right edges of XPs induce prosodic brackets, and that a language that has right-edge marking would predict the phrasing observed:

(60) (kissed Lysander)$_{\Phi}$ (and Demetrius)$_{\Phi}$.

Of course, this can only account for the phrasing in (57b, 58b), but not the phrasing in (57a, 58a).

Consider now the following cases, where again the predicate phrases with the first conjunct; the second conjunct follows the prosodic constituent consisting of the first conjunct and the predicate following it.

(61) a. She has (some blúeprints to leave) (and a book about Frank Lloyd Wright.)

b. (The students arrived), (and the professors as well.)

It is impossible in this case to phrase the second conjunct with the head:

(62) a. * (some blúeprints) (to leave and a book about Frank Lloyd Wright.)

b. * (The student) (arrived, and the professors as well.)

The phrasing follows straightforwardly from the approach in chapter 2. Extraposition rebrackets the string yielding a structure parallel to the coordinate structure in (52) and thus gives the correct boundary ranks.

The examples in (61) do not just involve a *prosodic* bracketing issue. It seems as if the predicate occurs *inside* of the coordinate structure. There is thus a purely syntactic bracketing issue: the verb ends up in the first conjunct, although it seems to take the entire conjunct as its argument. Any approach must analyze the construction as a case of syntactic extraposition (whether or not one resolves the bracketing problem by rightward movement or some other theoretical tool is a separate issue).

The bracketing in (61) is compatible with the edge-alignment approach, at least if [some blueprints to leave] is followed by an XP-break—which is plausible, since it could be the edge of the VP. But in order for this to work, the approach *also* has to invoke extraposition.

The major flaw of the edge-alignment approach is that it does not link the bracketing paradox to extraposition, at least for the case where the head precedes the conjuncts in (58). There is no reason to invoke extraposition here, in fact, edge-alignment is designed to derive a *mismatching* prosodic phrasing and assumes that the syntactic phrasing is unchanged. This, however, turns out to be false. A look at agreement illustrates why. I will first look at cases in which the predicate *follows* the first conjunct.

There is independent evidence that a syntactic rebracketing, and not just a prosodic rebracketing, is at stake. In structures like (61b), it is only the first conjunct that agrees with the following predicate:

(63) a. A student attends, and a professor.

b. * A student attend, and a professor.

If and only if the second conjunct is extraposed is first conjunct agreement possible:

(64) a. * A student and a professor attends.

b. A student and a professor attend.

A look at similar examples in German shows the same pattern. In the extraposed case, the verb agrees only with the first conjunct, while when the second conjunct does not extrapose, the verb agrees with both conjuncts:

(65) a. dass beim Seminar der Professor zugegen war/*waren, und die
that at.the seminar the professor there was/were and the
eingeschriebenen Studenten.
registered students

b. dass beim Seminar der Professor und die eingeschriebenen
that at.the seminar the professor and the registered
Studenten zugegen *war/waren.
students there was/were

This may be a sign that the second conjunct is really an adjunct when it is extraposed (cf. Munn (1993) for an analysis of coordination as adjunction). The grammar of extraposition/late merger of adjuncts is discussed in detail in Fox and Nissenbaum (1999) and Fox (2002).

The question is now whether there is any evidence for extraposition in the cases where the predicate *precedes* the coordinate structure, i.e. in the cases in (58)? Extraposition would be string-vacuous here. But there are cases where adverbials intervene which reveal that indeed extraposition is possible in English (Munn, 1993):

(66) John bought a book yesterday, and a newspaper.

There is also evidence for Extraposition from agreement facts. First, observe that verbs taking post-verbal subjects can agree with the first conjunct or with both conjuncts.[14]

(67) Yesterday at the seminar, there was/were a professor and two students.

Just as in German, extraposition forces first conjunct agreement:

[14]A complication is that some speakers use singular agreement at least in the present tense even with plurals:

(1) There's lots of them.

I am using the past tense, since the speakers I asked rejected singular agreement in the past tense in sentences with plurals.

(68) There was/*were a professor yesterday at the seminar and two students.

If first-conjunct agreement is really a test for extraposition, the prediction for English is now that there should be a correlation between the prosodic bracketing and first conjunct agreement, as summarized in the following paradigm, a prediction that is not easy to test just on impressionistic data:

(69) a. In the seminar room there were || a teacher | and two students.

b. In the seminar room there was | a teacher || and two students.

I cannot test whether the extraposition analysis of the apparent bracketing paradox presented in (58) applies for the Tiberian Hebrew example in (57). A restriction on phrasing in Tiberian Hebrew illustrates another parallel to the case of extraposition. Dresher (1994, 19) notes that when the predicate follows the coordinate structure, it is impossible for it to phrase just with the second conjunct:

(70) a. (kī-ḥemˀā ūdb̲aš) (yōk̲ēl)
for-curds and.honey shall.eat

b. * (kī-ḥemˀā) (ūdb̲aš yōk̲ēl)

The reason for this restriction is that there is no leftward extraposition of the first conjunct. Extraposition of the second conjunct on the other hand would lead to a structure in which the verb follows the first conjunct and phrases with it, as in the structure of (61). I do not know whether this latter kind of extraposition is available in Tiberian Hebrew, however.

The question of why extraposition applies and under what circumstances is a complicated one and cannot be discussed here in detail. Still, a straightforward prediction emerges from the discussion here: whenever extraposition is forced, the prosodic bracketing should follow accordingly, phrasing the predicate with the first conjunct; conversely, whenever extraposition is blocked, the coordinate structure should phrase together and be set off from the predicate by a stronger boundary.

Consider the following example due to Halliday, cited in (Clark, 1971). Here, the second conjunct must be a shorthand for an entire second sentence 'and I saw Mary, too'; otherwise combining it with 'too' would not make sense. The focus particle 'too' takes a propositional argument and presupposes that another proposition of a similar shape is true as well. So the second conjunct must stand for an entire proposition, rather than just an individual:

(71) a. I saw John yesterday, and Mary too.

b. ?? I saw John and Mary too yesterday.

This suggests that what I have called extraposition really involves a structure that is also semantically different from the unextraposed structure. The same is true for the case where the predicate follows the first conjunct. Only in the extraposed version can the focus adverb 'too' be used:

(72) a. Two people have arrived: John has arrived, and Mary, too.

b. * Two people have arrived: John and Mary have arrived, too.

As predicted based on the fact that 'too' forces extraposition, the *prosody* obligatorily signals the expected syntactic bracketing:

(73) a. I saw | John, || and Mary, too.

b. * I saw || John | and Mary, too

The case of extraposition in coordinate structure serves to illustrate that mismatches between syntax and prosody may sometimes only be apparent. The syntactic bracketing in this case may be just as expected based on the prosodic evidence.

3.4.2 Relative Clauses

One of the key observations that shaped theorizing on the prosody/syntax mapping is that clausal elements are mapped to their own intonational domain, independent of the syntactic bracketing. Consider the case of relative clauses (Chomsky (1965, 13), SPE, 372):

(74) This is | the cat || that chased | the rat || that stole | the cheese.

The relative clause forms its own phonological domain and is separated from the head of the relative clause by a boundary that is stronger than the boundaries that in turn separates the head from the predicate that precedes it. According to SPE, the bracketing in syntax should be as follows:

(75) This is [the cat that chased [the rat that stole the cheese.]]

The solution proposed in Lieberman (1967, 120) and in SPE was that a readjustment that rebrackets (75) into a different structure in which the three clauses are treated as on a par: "The resulting structure appears then as a conjunction of elementary sentences (that is, sentences without embeddings). This allows us to say that intonation breaks precede every occurrence of the category S (sentence) in the surface structure, and that otherwise the ordinary rules prevail" (SPE, 372).

Subsequent approaches took up this 'readjustment' account, but often considered the 'readjustment' to be an effect of the mapping of syntactic structure to prosodic structure, which result in a mismatch. The boundaries separating the relative clause can e.g. be taken to be due to alignment constraints that force prosodic boundaries at the edge of certain syntactic constituents, in this case a clausal node. The different bracketing is then seen as a genuine mismatch between prosody and syntax.

It is conceivable though that the readjustment takes place in *syntax*. The nature of readjustment was left open in SPE, and a 'performance' explanation was considered a possibility, although this was not fleshed out in detail. In this section, I will give some arguments in favor of a syntactic approach—the main point that I want to make is that there is evidence for a *syntactic* bracketing that *matches* the prosody.

In a natural rendition of (75) the boundaries after the verbs are usually weaker than those preceding the relative clause. This is exemplified in (76a), which is the preferred phrasing compared to (76b):

(76) a. that chased | the rat || that stole the cheese.

b. that chased || the rat | that stole the cheese

The prosody in (76b) may be appropriate in a context that has narrow focus on the direct object, but may otherwise not be the preferred phrasing:

(77) Who did the cat chase?

the cat chased || the rat | that stole the cheese

The phrasing that corresponds to the syntactic bracketing that Aspects and SPE assumed is then possible, at least under certain circumstances. But how to derive the apparently mismatching bracketing?

One possibility is that just as in the case of coordination discussed above, extraposition is involved. Extraposition of relative clauses is certainly possible:

(78) I saw the cat yesterday that caught the rat on Monday that had stolen the cheese on Sunday.

Extraposition would render a bracketing that corresponds to the prosodic bracketing positing in SPE:

(79) [[This is the cat] [[that caught the rat] [[that stole the cheese]]]]

In order to test whether prosody really mirrors the syntax, it would be necessary to control whether or not extraposition takes place or not. Adverbs can be used to force extraposition as in (78), but how can we prevent it from happening?

Hulsey and Sauerland (2005) argue that extraposition is impossible if the head of the relative clause is an idiomchunk:

(80) a. Mary praised the headway that John made.

b. * Mary praised the headway last year that John made.

If extraposition is involved in rendering the prosody in (75), then we expect to see an effect of idiom chunks on prosody. Consider the following two constructions, with the prosody that groups the head with the following relative clause:

(81) a. This was entirely due || to the advantage that he took || of the headway that she had made before.

b. This was entirely due || to the extra funds that he took || from the surplus that she had made before.

The prediction is now that the 'mismatching' prosody discussed in SPE should be impossible when the head of the relative clause is an idiom chunk. This is confirmed by impressionistic data collected from native speakers, who sense a contrast in acceptability between the following two examples:

(82) a. ?* This was entirely due to the advantage || that he took of the headway || that she had made before.

b. This was entirely due to the extra funds || that he took from the surplus || that she had made before.

This contrast constitutes evidence for the claim that readjustment, the rebracketing observed in SPE and Aspects, requires syntactic extraposition. There is no bracketing paradox between syntax and prosody in this case.

The judgment is reminiscent of cases with focus within idioms, which are infelicitous since there is no alternative for a focused material that would make sense as a replacement given the idiomatic interpretation of the structure:

(83) She *KICKED* the bucket.

It seems that one factor involved in a 'matching' (i.e. extraposition) derivation of relative clauses is that the head has to be focused. This would capture why idiom chunks are not allowed as heads of extraposed relative clauses.[15]

[15]There is another property that might distinguish different types of relative clauses. Relative clauses derived by raising are often prosodically subordinated. This was observed i.a. in Bresnan (1971, 258):

(1) Mary liked the propósal that George left.

Nuclear stress falls on the head of the relative clause. But this, as was critically observed in Berman and Szamosi (1972) andn Lakoff (1972), is not always the case. Bresnan (1972, 337) illustrates the problem with the following two examples due to Stockwell (1972). When presented out of context, they differ in how they tend to be pronounced (of course, depending on context the stress could shift—How contexts influences prosody will be discussed in detail in chapter 2):

(2) a. Introduce me to the man you were tálking about.

b. I'll lend you that bóok I was talking about.

A comparison with German supports the extraposition-analysis of the English data. The OV word order in embedded clauses has the effect that extraposition is easy to diagnose—it is not string-vacuous as in English. Consider the following example, similar to the one discussed in SPE:

(84) Ich glaube dass dies die Katze ist, die die Ratte gejagt hat, die den
I believe that this the cat is that the rat chased, that the cheese
Käse gestohlen hat.
stolen has
'I think this the cat that chased the rat that stole the cheese.'

The relative clauses are extraposed, as is obvious considering the word order. The head is separated from its relative clause by the predicate. Without extraposition of the relative clauses, the structure would be a center-embedded structure, and end up unintelligible:[16]

(85) ?* Ich glaube dass dies die Katze die die Ratte die den Käse gestohlen hat gejagt hat ist.

According to Bresnan, there is also a semantic difference between the relative clauses: if the relative clause receives nuclear stress, then it is used to pick out one out of a set of alternatives defined by the head noun. In this case the head of the relative clause is a concealed partitive. In the sentence with nuclear stress on the head noun, however, there is no such partitive reading. I will present an potential explanation for the intuition reported by Bresnan in the discussion of effects of information structure in chapter 7.

The prosodic contrast between subordinated and non-subordinated relative clauses was related to raising *vs.* matching derivations of relative clauses in unpublished work by Karlos Arregi. Based on reconstruction effects, Arregi concludes that a raising analysis is applicable for relative clauses that are subordinated but not for those that contain the nuclear stress. If this analysis is correct, then extraposed relative clauses should not be prosodically subordinated, since extraposition forces a matching derivation. I have not tested this prediction.

[16]Center-embedding structures, used as an example for a performance restriction in Chomsky (1965), have been to shown to be hard to process in structures across languages and different constructions. The preference for extraposition of relative clauses my well have a motivation that relates to processing. This is in fact the explanation favored in SPE for the readjustment of relative clauses in (75). But this relation to processing factors does not mean that extraposition does not form part of syntax proper, and the syntactic and semantic effects of extraposition observed in Hulsey and Sauerland (2005) clearly show that extraposition is not just a phonological phenomenon.

Similar to the case of English, extraposition is impossible in German when the head of the relative clause is an idiom chunk.

(86) a. Peter war über den Bären den Maria ihm aufgebunden hatte verärgert.
Peter was about the bear that Maria him given had annoyed
'Peter was annoyed about the prank that Maria played on him.'

b. * Peter war über den Bären verärgert, den Maria ihm aufgebunden hatte.

The analysis of the extraposed relative proposed in Hulsey and Sauerland (2005) is that the relative clause is merged late as an adjunct, after the head of the relative clause has undergone covert movement. This is based on the analysis of extraposition of adjuncts in terms of 'later merger' discussed in Fox and Nissenbaum (1999) and Fox (2002).

The proposed solution to resolve the alleged mismatch between prosody and syntax in (75) is then exactly parallel as the solution for the case of the phrasing of the predicate with the first conjunct: syntactic extraposition is involved, and the prosodic bracketing matches the syntactic bracketing after all.

It may be that extraposed structures are sometimes preferred over non-extraposed structures for reasons that relate to parsability and maybe even pronounceablity, as discussed in SPE (p. 372) and also in Lieberman (1967, 120–121).

3.4.3 Possessor Constructions

One corollary of the way the cyclic system is set up is that all left-branching structures are prosodically fully articulated. This statement appears to be incorrect, however. Consider possessor-structures:

(87) John's brother's sister's dog's house.

This structure looks left-branching, but surely has a 'flat' prosody. The expected articulated prosody would look as follows, and is clearly very marked at best:

(88) John's brother's | sister's || dog's ||| house.

It is not clear though that the structure of possessors is underlyingly really left-branching.

(89) John's former house.

a. A former house that belongs to John (dispreferred).

b. A house that formerly belonged to John (preferred).

The NP in (89) has two readings, the more salient reading without further context being (b), the reading in which 'John' is in the scope of the adverb 'former', and the adverb modifies the NP 'John's house'. Larson and Cho (1999) argue that the left-branching structure gives wrong bracketing for this reading, and proposes that possessor constructions are underlyingly right-branching.

In Larson's analysis, possessor constructions have a structure comparable to the analysis Freeze (1992) gives to the verb 'have'. Freeze analyzes 'y have x' underlyingly as 'is x to y'. The locative preposition incorporates into the copula to create the complex auxiliary 'have'. Larson proposes that analogously, the possessor 's' is derived by incorporating a preposition into the determiner, so 's' is the equivalent of 'to.the':

(90) Movement analysis: the house to John $\rightarrow$ John to.the house

The underlying structure is compatible with building the complex possessor structure in a single cycle, since it is right-branching. The question of what triggers the surface reordering will be left open at this point. The facts from possessor constructions necessitate a looser notion of linearization of the constituents put together within a single cycle. For discussion of movement both preceding and following spell-out and a model of cyclic linearization see Fox and Pesetsky (2005), Ko (2005), and Sabbagh (to appear).[17]

[17]Another question that I will not address is the question of whether or not the associative law holds in these structures—which would be predicted based on their prosody.

Summary

This chapter discussed evidence that the findings from chapter 2 on the prosody of coordinate structures carry over to other domains of the grammar. Domains that are associative show a flat prosody, and in domains that are not associative the prosody reflects the syntactic bracketing that can be independently established looking at scope facts.

A look at apparent bracketing mismatches between prosody and syntax revealed further evidence for the mapping proposed in chapter 2. The prosody closely reflects the way structure is assembled during syntactic derivations.

This is not to say that there are *no* factors apart from syntax that shape prosody. Eurythmic factors e.g. can effect adjustments of the output grid structure. I will not explore these other factors here .

Chapter 4

Boundary Ranks in Phonetics and Phonology

Abstract

The theory of syntax-prosody mapping presented in chapter 2 derives metrical presentations that define prosodic feet at various lines in a metrical grid. The number of the grid line corresponds to the strength of the perceived boundary separating feet at a given line. This chapter reviews how well this representation reflects the phonetics and phonology of prosodic structure, and argues that it correctly captures that (i) prosodic structure reflects the expected boundary ranks, and (ii) that the mapping between syntax and phonology is recursive in nature and not tied to specific syntactic categories.

4.1 The Representation of Prosodic Structure

The prosodic facts reported in chapter 2 are based on intuitions on prosodic grouping based on boundary strength. They are used to motivate a relational metrical grid, first proposed in Liberman (1975) and developed in many subsequent studies (e.g. Prince (1983), Hayes (1984), Selkirk (1984). I am assuming an adapted version of the bracketed version of the metrical grid (cf. Halle and Vergnaud, 1987, Idsardi, 1992), which organizes the phonological representation of an expression recursively into feet.

(1) Relational Metrical Grid

```
| ×   ×   ×   × |      4
| × | ×   ×   × |      3
| × | × | ×   × |      2
| × | × | × | × |      1

  A   B   C   D
```

The grid encodes grouping by virtue of foot boundaries of different strength: The strength of a foot boundary at line i in the grid is perceived as stronger than the strength of a foot boundary at line i-1. This chapter looks at the phonetics and phonology of prosody and how it relates to these boundary ranks.

There are various quantitative effects that suggest that a recursive hierarchy of boundary ranks makes the correct predictions. One set of facts relates to durational effects: domain-initial strengthening and domain-final lengthening are both cumulative effects that correlate with boundary rank. A second effect regards pitch-accent scaling. While whether or not a pitch accent is down-stepped may be a categorical effect, the precise scaling of the pitch value requires a recursive representation that reflects nested layers of grouping, as expected based on the recursive metrical grid representation. These quantitative effects will be discussed in section 4.2.

The grid in (1) represents hierarchically organized feet. Feet and their boundaries fall into different categories by virtue of being on different grid lines. However, the theory does not categorize these feet into prosodic constituents with different labels, as is usually assumed in the literature on the 'prosodic hierarchy'. An example of a prosodic hierarchy is given in (2):

(2) Prosodic hierarchy (Selkirk, 1986, 384):

```
(------------------------------------------------------------)  Utt
(-------------------------------------)(---------------------)  IPh
(----------)(-------------------------)(---------------------)  PPh
(----------)(-------)(---------------)(------------)(-------)  PWd
(--)(------)(------)(--)(-----------)(-------)(--)(------)  Ft
(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)  Syl
```

While the theory is compatible with an implementation of the different boundary ranks that distinguishes prosodic domains of different types, the central departure from earlier work on the mapping between syntax and prosody is that the category of the prosodic domains cannot be directly related to syntactic categories such as $\bar{X}$ vs. XP. The issue and its implications are discussed in section 4.3.

The discussion of this chapter forms the basis for an experimental investigation of the relational view of metrical structure. The results will be discussed in chapter 5.

4.2 Quantitative Evidence

Prosodic structure has effects on the duration of segmental gestures. This section presents evidence that domain-initial strengthening and pre-boundary lengthening are cumulative effects, which correlate with the boundary ranks provided by the metrical structure proposed here.

4.2.1 Durational Effects and Boundary Strength Scaling

Durational effects were identified in Lehiste (1973) as the most reliable in disambiguating syntactic structures based on their bracketing. This section reviews evidence from pre-boundary lengthening, pauses, and domain-initial strengthening.

The coordinate structures in (3) were among the data set Lehiste (1973) investigated, and was one of the structures for which listeners scored best in identifying the correct bracketing in a listening task.

(3) a. ((Steve or Sam) and Bob) will come.

b. (Steve or (Sam and Bob)) will come.

Lehiste (1973, 117) found that the length of the underlined parts were longer if it contained a constituent boundary—that is if the higher-level bracket cuts into this constituent as in (3q). Sometimes, but not always, there was laryngealization and/or a pause at the constituent break. She also notes that 'Sam and' in the structure (3b) was 'drawled', i.e. reduced and less carefully pronounced.[1] These durational effects observed by Lehiste (1973) can be related directly to boundary ranks.

Boundary Ranks affect duration. The strength of a boundary is reflected by pre-boundary lengthening of the material preceding it. Klatt (1975) showed that segments are lengthened preceding boundaries, even in the absence of pauses. Pre-boundary lengthening affects mainly the material directly adjacent to a prosodic boundary, but also the main stress syllable according to Shattuck-Hufnagel (2005). It has been shown to correlate closely with the strength of the following boundary (Wightman et al., 1992, Price et al., 1991, Shattuck-Hufnagel and Turk, 1996).[2]

Byrd and Saltzman (1998), Cho and Keating (2001), and Byrd and Saltzman (2003) concluded that the degree of lengthening at particular boundaries depends on the strength of the boundaries, but that the difference is one of degree and not one of *type* of effect. In other words, durational effects give evidence for a relative notion of boundary strength but not of different types of boundaries.

The metrical grid assigned here provides a measure of the strength of boundaries: the boundary rank. The boundary ranks differentiate prosodic domains on the discrete scale defined by the metrical grid representation. The degree of domain-final lengthening in coordinate structure will be shown in chapter 5 to correlate with the

[1]In a follow-up study (Lehiste et al., 1976), the sentences were manipulated in order to verify whether segmental duration (as opposed to pauses and pitch) was one of the cues used by listeners in assigning a syntactic structure. Pitch was set constant at 100 Hertz for the sentences, and duration was manipulated. The results showed that durational cues indeed determined the reading listeners assigned to the structures.

[2]Some researchers (e.g. Martin (1971)) interpret pre-boundary lengthening as a 'slow down', and speculate that that this reflects the extra-time needed by the speaker to compose the next phrase. I will return to this point in chapter 5.

boundary ranks predicted by the metrical grids proposed here.

The boundary ranks assigned here are relative. A boundary is stronger or weaker than other boundaries depending on its rank. The metrical grid does not specify an absolute strength or category of a boundary. The theory is then able to account for quantitative phonetic effects that distinguish different prosodic boundary ranks—it does not distinguish different *qualitative* categories of prosodic boundaries.

There is phonetic evidence that boundaries are indeed scaled relative to each other. Carlson et al. (2001) show evidence that the *relative* boundary strength is used to disambiguate attachment ambiguities.

In chapter 5, a model is proposed in which *all* boundaries are phonetically realized by scaling the strength of a boundary relative to already produced boundaries earlier in the utterance.

(4) BOUNDARY STRENGTH SCALING

The strength of a boundary is scaled relative to d boundaries produced earlier. It must be realized such that it is perceived with the appropriate assigned boundary rank compared to boundaries of different or equal rank realized before.

Durational cues such as pre-boundary lengthening are used to achieve the correct boundary strength scaling. The very first constituent in an utterance sets the benchmark for the realization of later constituents. This relative scaling of boundary ranks is compatible with an unlabeled grid, and shows that indeed the organization of prosodic feet into a hierarchy may be relative and not categorical. Experimental evidence for this model is presented in chapter 5.

Another durational correlate of boundary strength are *pauses*. Grosjean and Collins (1979) asked subjects to read test sentences at various speech rates. They added the overall length of pauses produced between each pair of adjacent words as a measure of the grouping of words. These pause durations were then used to derive binary tree structures in the following way: the two words separated by the shortest pause duration were grouped together under one node, and then treated as one word afterwards; this procedure was continued until a complete structure was derived. Crucial in deriving this tree representation was the *relative* duration of pauses between words. The derived tree structure matched closely listener's intuitions about grouping of the sentences, which were independently elicited. Gee and Grosjean (1983) gave a prosodic interpretation to the derived tree structures, again based on relative degrees of pause duration, as opposed to categorical criteria that identify prosodic constituents of a particular kind.

Pre-Boundary lengthening and perceived pauses are closely related. Martin (1970, 1971) found that listeners report hearing pauses although there were no unfilled pauses in the signal. In those cases, there was evidence for syllable lengthening just before the perceived juncture. The syllable before the juncture was longer compared to the syllable following the juncture. Conversely, for those pauses that were present in the signal but not detected by listeners, syllables before the boundary were *not* lengthened—syllable lengthening was thus the dominant cue.

Finally, domain-initial strengthening is also a durational effect that reflects boundary strength. Fougeron and Keating (1997), Lavoie (2001), Cho (2002), Keating et al. (2003), Keating (to appear) discuss evidence that the phonetic realization of segments depends on their position in the prosodic structure. In particular, they argue that articulatory strengthening occurs at the beginning of prosodic domains, and that this strengthening quantitatively increases cumulatively with the strength of the prosodic boundary.

One type of experiment conducted in order to show this effect uses electropalatography to measure linguopalatal contact, which is a reflex of the degree of oral constriction. The examples investigated in Fougeron and Keating (1997) were arithmetic expressions, whose bracketing is of course very similar to the bracketing one finds in coordinate structures:

(5) a. 89 + 89 + 89 + 89 = a lot.

b. (89 + 89) * (89 + 89) = a lot.

c. 89 * (89 + 89 + 89) = a lot.

d. $(89 + 89 + 89) * 89 =$ a lot.

Subjects were asked to read statments in reiterant speech that were based on the bracketing modeled by the arithmetic formulae. The following statement is equivalent in bracketing to (5d):

(6) (nonono no nonono no nonono) no nonono equals a lot

The results of the experiment showed (i) that nasals are articulated more strongly at the beginning of domains compared to domain-medially, and (ii) that they are articulated more strongly at the beginning of higher-order prosodic constituents compared to lower order prosodic constituents.

Similar results were reported from an acoustic study of VOT in Korean in Jun (1993), based on which Jun suggests that there is hierarchy of strength which can be defined in relation to the position in the prosodic structure.

Fougeron and Keating (1997) propose that phonetic strengthening is cumulative and continuously increases with the strength of the boundary. It is not tied to particular prosodic domains (say the prosodic word or the intonational phrase). Cho and Keating (2001), Keating (to appear) explicitly argue that domain-initial strengthening is continuous, and does not lend support for a fixed category of a particular number of categorically distinct types of prosodic constituents.

Keating et al. (2003) illustrate evidence that domain-initial strengthening occurs cross-linguistically, independent of the particular prosodic system of the language. Thus languages that are not stress-based such as English, e.g. French, Korean, and Taiwanese, show very similar patterns of domain-initial strengthening.

Not only are segments strengthened at the beginning of prosodic domains, segments can also be inserted to strengthen the beginning of a domain. Dilley et al. (1996) show evidence that glottal stop insertion is more likely at stronger prosodic domain breaks compared to weaker boundaries.

Based on the evidence in the phonetics literature, we can conclude that domain-initial strengthening and pre-boundary lengthening are cumulative effects, which can be related to the boundary ranks the metrical grid structures derived here provide.

Boundaries of different strengths differ in the degree of the lengthening/strengthening they induce, but they are not differentiated in the type of effect they have.

The evidence from studies on domain final lengthening, pause realization, and domain-initial strengthening is compatible with a prosodic representation that only distinguishes relative boundary ranks, and not categorically distinguished prosodic constituents.

The experiment presented in chapter 5 provides more evidence that the phonetic implementation of prosodic structure is indeed relative, and that the strength of prosodic boundaries is scaled relative to earlier boundaries in the utterance.

4.2.2 Pitch Accent Scaling

The metrical structure does not only play a role in determining *where* accents are placed. It also influences *how* they are realized. Ladd (1988) and Fery and Truckenbrodt (2004) show evidence from pitch accent scaling in English and German respectively. The following type of coordination structure was investigated:

(7) a. A but (B and C)

b. (A and B) but C

The following are example sentences illustrating the data:

(8) a. Ryan has a lot more money, (but Warren is a stronger campaigner and Allen has more popular policies.

b. Allen is a stronger compaigner, and Ryan has more popular policies, but Warren has a lot more money.

The pitch accent scaling distinguishes the two types of structure. In structures of type (7a), conjunct C is downstepped relative to B, and B in turn is downstepped relative to A. The pitch level goes down from accent to accent. In structures of type (7b) on the other hand, C and B are at about the same level[3], but both are set to a lower

[3]In fact, C is scaled a little higher in pitch than B, a fact discussed in detail in Fery and Truckenbrodt (2004).

pitch compared to A. This contrast in pitch scaling requires a prosodic representation that encodes the syntactic difference between (7a) and (7b).

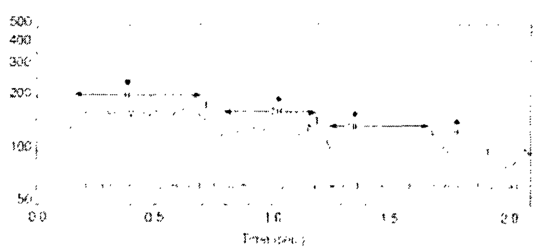
Ladd proposes to explain the difference in pitch scaling by a hierarchically metrical representation that allows a recursive nesting of intonational phrases. Within each level of coordinate structure, conjuncts are downstepped relative to the preceding conjunct. In structures of type (7b), conjunct C is downstepped relative to the first conjunct (A and B). This has the effect that the pitch of an accent in C is lower than the maximal pitch in (A and B), but it is not lower than the pitch in a preceding conjunct, in this case B.

In van den Berg et al. (1992), evidence is presented for scaling from coordinate structures involving place names, and show further evidence for Ladd's idea about how pitch scaling reflects hierarchical structure. The accent within each conjunct is downstepped relative to the preceding one ('en' is the Dutch for 'and'):

(9) Harlem, Rommeldam, Harderwijk en Den Helder.

The pitch scaling is illustrated in the following figure:

(10) Downstepping in 'Flat' Coordinate Structure



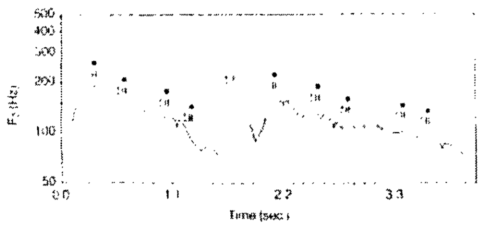
When a coordinate structure is further bracketed, then the pitch scaling reflects this bracketing:

(11) Downstepping in a Nested Structure

(Merel, Nora, Leo, Remy), en (Nelie, Mary, Leendert, Mona en Lorna).

The pitch accents on the conjuncts *within* each of the more embedded brackets is as in the unembedded structure in (9). Each accent is downstepped relative to a preceding one. But there is a further downstepping effect, that affects *the entire cycle* (Nelie, Mary, Leendert, Mona en Lona): It is downstepped in pitch range relative to the first conjunct (Merel, Nora, Leo, Remy).

(12) Pitch Scaling in Nested Coordinate Structure



The generalization about pitch scaling has to be recursively stated: downstepping exists at different levels of embedding. This is captured in van den Berg et al. (1992), who propose that accents are downstepped relative to preceding accents, and in addition, entire domains containing accents can be downstepped relative to preceding domains. How can we state the generalization based on the prosodic representations used here?

Consider the representation assigned here for structures with the bracketing in (7):

(76) Metrical Grid

a. A or (B and C)

```
| × | ×   × |
| × | × | × |
  A   B   C
```

b. (A or B) and C

```
| ×   × | × |
| × | × | × |
  A   B   C
```

In deciding how to scale the accent on C, it matters how deeply embedded C is. I define a notion of embeddedness of accents by introducing the concept of a 'minimal foot'.

(13) Minimal Foot

The minimal foot of a grid mark is the highest foot in the column in which the grid mark is adjacent to a preceding foot boundary.

I mark the minimal foot for each top line grid mark by the symbol ◇:

(14) Metrical Grid

a. A or (B and C)

```
| ◇ | ◇     × |
| × | × | ◇ |
  A   B   C
```

b. (A or B) and C

```
| ◇     × | ◇ |
| × | ◇ | × |
  A   B   C
```

I will assume with van den Berg et al. (1992) that each foot has a pitch register level. This is the pitch value to which the first H* that it contains is scaled, if there is one. The generalization about downstepping can now be stated as follows:

(15) Pitch Accent Scaling

The pitch level of a minimal foot is downstepped relative to the pitch level of the foot that precedes it on the same line.

H* pitch accents are realized at the pitch level of the current minimal foot. This has the effect that if H* pitch accents are placed, they are downstepped relative to the preceding pitch level of the closest minimal foot (marked by ◇) at the same grid line or a higher line.

Consider the minimal foot corresponding to the grid mark of C in the two structures in (76). The convention on pitch accent scaling states that C is scaled relative to the accent on B in (76a), but is scaled relative to the pitch level of the foot containing (A and B) in (76b). Since the pitch level of this foot is the same as the level to which the accent on A is scaled, this effectively means that C is downstepped relative to A, just as proposed in Ladd (1988).[4]

[4]How can we motivate the notion 'minimal foot'? The pitch register level of a prosodic domain is necessary in order to phonetically implement an accent. If the pitch scaling convention here is correct, then the information about how to realize the accents within the minimal foot phonetically is available once this foot is completed, since only the preceding minimal foot is necessary to scale its pitch level correctly. The amount of information that has to computed (or anticipated) before a constituent can be pronounced is the minimal foot, so maybe minimal feet are planning units for speech production.

This statement about pitch scaling also works for the unembedded structure in (9):

(16) Prosodic Structure for (9)

```
| ◇   ×   ×   × |
| × | ◇ | ◇ | ◇ |
  A   B   C   D
```

The pitch level in each conjunct is downstepped relative to the immediately preceding foot. This has the effect that H* placed on those conjuncts will be down-stepped in pitch level from left to right. The metrical structure assigned to the nested structure in (11) looks as follows:

(17) Prosodic Structure for (11)

```
| ◇   ×   ×   ×   ×   ×   ×   ×   × |
| ×   ×   ×   × | ◇   ×   ×   ×   × |
| × | ◇ | ◇ | ◇ | × | ◇ | ◇ | ◇ | ◇ |
  A   B   C   D   E   F   G   H   I
```

Pitch accent scaling has been used in Ladd (1988) to argue for a recursive nesting of prosodic categories with the same phonological 'label', the intonational phrase. The representation used here provides such a recursive notation, albeit one without any labels at all. More evidence for recursive nesting can be found in the recursive accent system of Tiberian Hebrew used in the Masoretic Bible (Dresher, 1994). Dresher argues that the accents annotate prosodic structure, and relate directly to regularities in segmental phonology and in the distribution of pausal forms.

The grid encodes the information necessary to state the generalization about pitch scaling observed in Ladd (1988) and van den Berg et al. (1992). More evidence for pitch scaling is discussed in Kubozono (1989), Kubozono (1992), and Fery and Truckenbrodt (2004). Further arguments in favor of a hierarchical representation of metrical structure and a model of how accents can be implemented based on it are discussed in Dilley (2005).

To summarize, there is quantitiative evidence that suggests that boundaries of different ranks are distinguished by gradient phonetic correlates in duration and pitch value. If all correlates of prosody domains were gradient, then a relational and recursive representation such as the metrical grid would be entirely sufficient to capture prosodic structure. The next section will review evidence for categorical distinctions between prosodic domains, and how thy can be made compatible with the recursive approach pursued here.

4.3 Qualitative Evidence

The phonological literature on prosody often assumes a 'prosodic hierarchy' of phonological domains. This was proposed in (Selkirk, 1978, 1980, 1981) and was extended in Nespor and Vogel (1982, 1986). The metrical grid derived in chapter 2 is purely relational, just as the metrical grid proposed in (Liberman, 1975, Libermann and Prince, 1977, Prince, 1983, Selkirk, 1984). This section discusses evidence for categorical distinctions between prosodic domains, and how they can be made compatible with the proposal in chapter 2.

The main point will be that the theory is compatible with categorical distinctions between prosodic domains, but the theory is not categorical with a theory that tries to tie specific prosodic labels to specific syntactic categories (e.g. $\bar{X}$ vs. XP). Instead, the mapping between syntax and phonology is relational, even if the implementation of the relevant ranks in phonology may distinguish prosodic domains with different labels.

4.3.1 Designated Categories

The prosodic hierarchy has been key in many subsequent studies of prosody (cf. Selkirk, 1986, Hayes, 1989, Inkelas and Zec, 1990, Selkirk, 1995b). It consists of a layered prosodic structure in which each line bears a different label. Different authors have made different proposals with respect to the exact number and label of the categories. The following is one illustration, repeated from (2):

(2) Prosodic hierarchy (Selkirk, 1986, 384):

```
(--------------------------------------------------------)   Utt
(----------------------------------)(--------------------)   IPh
(---------)(----------------------)(--------------------)   PPh
(---------)(-------)(--------------)(-----------)(-------)   PWd
(--)(------)(------)(--)(----------)(-------)(--)(------)   Ft
(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)   Syl
```

The different types of prosodic domains were originally identified by phonological rules that apply only within certain domains but not others. Work on the tonology of prosodic domains (Pierrehumbert, 1980, Beckman and Pierrehumbert, 1986) lead to a second set of criteria to identify types of prosodic domains, namely particular pitch accents and boundary tones. These criteria were developed into an annotation system to transcribe intonation, the ToBI-system (Silverman et al., 1992). The intonational criteria and the criteria based on phonological rule application are assumed to converge on the same categories, and several studies have argued that there is indeed such a correlation (i.a. Jun, 1998).

A central assumption in much of the literature on prosodic phonology is that a nesting of prosodic constituents bearing the same label is impossible (Selkirk, 1986). This restriction has been demoted to a violable rather than an inviolable constraint in the context of Optimality Theory in Selkirk (1995a, 443) and Truckenbrodt (1995) (C^i stands for a prosodic category with label i):

(18) *Nonrecursivity*: no C^i dominates C^j, j = i.

The metrical grid derived in chapter 2 on the other hand is purely relational, just as the metrical grid originally proposed in Liberman (1975) and Libermann and Prince (1977). Liberman (1975, 73) describes the idea behind a relational metrical grid as follows:

> "A metrical grid is a pattern which subdivides intervals of time, in the way that is familiar from musical notation; [...] The Grid, like time, is in principle infinite – we can pick any place in it to start, and have available

sequences forward and backward, or up and down in the hierarchy, as we choose."

The representation used here is a version of the bracketed grid (Halle and Vergnaud, 1987). Prosodic feet are ranked relative to each other by virtue of which line in the grid they are part of, but not by a difference in categorial label. The representation thus captures some but not all of the information encoded in the prosodic hierarchy.

(1) Relational Metrical Grid

```
| ×   ×   ×   × |
| × | ×   ×   × |
| × | × | ×   × |
| × | × | × | × |
  A   B   C   D
```

What the metrical structure used here fails to assign are *labels* of the grid-lines. Prosodic feet are different in status only by virtue of being on different lines in the grid.

Since Liberman's proposal, a number of studies have used an unlabeled recursive metrical structure to encode prosody, e.g. Selkirk (1984), Halle and Vergnaud (1987), Idsardi (1992), Cinque (1993), Arregi (2002) and Ishihara (2003). Selkirk (1984, 26ff) presents arguments in favor of reducing the prosodic hierarchy to the intonation phrase and the syllable, and represent prosody otherwise by a relational grid representation.

The representation used here is compatible with a mapping of the output of the algorithm to a fixed hierarchy of categorical prosodic categories. For example, it is possible that the syntax-to-phonology mapping principles give us the unlabeled grid, which is then mapped to a labeled grid in a second stage of phonological implementation. The prosodic hierarchy would then be the way the relative boundary distinctions provided by the syntax-phonology mapping are formally implemented by the phonological component.

What the theory presented here is *not* compatible with is that specific syntactic constituents are mapped to specific prosodic constituents. The way the mapping from syntax to phonology itself works only derives relative ranks of prosodic boundaries. The present theory is then incompatible with one of the standard assumptions in the theory of prosodic phonology. The assumption at the heart of this theory is summarized by Selkirk (1986, 385) as follows:

> The proposal is that the relation between syntactic structure and prosodic structure above the foot and below the intonational phrase is defined in terms of the *ends* of syntactic constituents of designated types. [...] The idea is that a unit of phonological structure, a derived domain, will have as its terminal string the stretch of the surface syntactic structure that is demarcated by the right or left ends of selected constituents, as in (19):
>
> (19) a. $_{\alpha}[$ ..., where ... contains no $_{\alpha}[$
>
> b. ... $]_{\alpha}$, where ... contains no $]_{\alpha}$
>
> The general claim I am making here is that α will be drawn from the set of categories defined in X-bar theory, and that α indicates only a level (or type) in the X-bar hierarchy, which is to say that the syntax-to-prosodic structure mapping is claimed to be cross-categorial in nature.

The crucial point in Selkirk (1986) is that there are certain designated syntactic constituents that end up providing the edges of prosodic domains with a certain label. This notion of designated syntactic categories for prosodic domains was expressed in Selkirk (1996, 444): as follows:

(19) The edge-based theory of the syntax-prosody interface
Right/Left edge of $\alpha \rightarrow$ edge of β,
α is a syntactic category, β is a prosodic category.

The particular list of prosodic domains and how they are 'syntactically grounded', i.e. which syntactic nodes correspond to them, was updated in Selkirk (to appeara).

Following Dresher (1994) and other works, the proposal in chapter 2 does *not* assume designated categories. The claim is that rather than specifying what type of syntactic constituents are mapped to a single phonological phrase only the relative rank is derived. This section reviews some of the evidence bearing on the issue.

4.3.2 Phonological Rule Application

In the theory of prosodic phonology, prosodic domains are labeled and fall into several different categories.[5] Some of the evidence used to motivate these prosodic categories is summarized below:

(20) Evidence for this type of approach according to Inkelas (1990)

a. Fewer Rule Domains than there are Rules. There is a small, fixed set of domains.

b. Exhaustive Parsing: Every utterance seems to be exhaustively parsed in those domains. Evidence: for each rule, every part of an utterance will be parsed into some domain where that rule applies.

c. Strict Layering: Each domain seems to be properly contained in a domain of the next higher level.

The mapping algorithm developed here also derives a strictly layered representation, but it fails to assign categories to the different grid lines. So far I have discussed evidence from phonetics that suggests that a relative unlabeled grid may actually provide sufficient information to account for the implementation. But evidence from phonological rules might eventually force us to distinguish prosodic categories of different types.

If it is true that certain phonological rules apply in prosodic domains that bear a particular label, then our representation has to be enriched to capture this fact. A second type of evidence for identifying particular types of prosodic constituents is by the presence/absence and type of boundary tone and presence/absence of a special nuclear accent (cf. Pierrehumbert (1980), Beckman and Pierrehumbert (1986), Pierrehumbert and Beckman (1988) and also the transcription system ToBI

[5] The precise number and label of categories varies between the papers.

Silverman et al. (1992)). There is a close correlation between domains of phrasal rule application and tonal domains (i.a. Jun, 1998), which supports the view that prosodic constituents that are special exist.

The representation used here is *compatible* with a mapping of the output of the algorithm to a fixed hierarchy of categorical prosodic categories. For example, one could claim that the mapping principles give us the unlabeled grid, which is then mapped to a labeled grid in a second stage. The prosodic hierarchy would then be the way the relative boundary distinctions provided by the syntax-phonology mapping are formally implemented by the phonological component.

What the theory presented here is *not* compatible with is that specific syntactic constituents are mapped to specific prosodic constituents. The way the mapping from syntax to phonology itself works only derives relative ranks of prosodic boundaries. The present theory is then incompatible with one of the standard assumptions in the theory of prosodic phonology. The idea is that rather than specifying what type of syntactic constituents are mapped to a single phonological phrase, only the relative rank is derived.

Let's assume that phonological criteria allow us to identify a prosodic domain that is the phonological phrase (Φ). The idea is now that in implementing the relational grid, the theory allows us to *choose* a line i in the grid that counts as the Φ-line. Within each foot at this line (and by implication all lower feet contained in it) the phonological rules associated with Φ will apply.

The segmental rules that pertain to the set of rules that apply in Φ can then be used to distinguish this domain from higher domains—but they no longer give information about the precise syntactic status of the elements they contain.

This theory makes predictions that are quite different from theories operating with designated syntactic nodes for specific prosodic domains. To illustrate this, I will use the segmental rule of flapping in English (thanks to Charles Reiss for suggesting this). Flapping can occur in coordinate structures:

(21) a caɾ or a rat?

First, it predicts that two constituents A, B that are usually mapped to a single prosodic domain P_i should be able to phrase separately if we make sure that either A or B contains a prosodic break itself. This prosodic break is predicted by the present theory to be prosodically weaker compared to the one separating A and B. The need to implement this difference in boundary rank should enable the grammar to map A and B to separate domains:

(22) a catˀ or (a raɾ and a hatˀ.)

Conversely, If a rule does *not* apply between two constituents, it should be possible to still enable the rule to apply by further embedding the structure in such a way that the boundary separating the two constituents has to be relatively weaker compared to some yet higher boundary.

(23) a doɡ and (a caɾ or (a raɾ and a hatˀ.)

Testing these predictions would necessitate running a controlled experiment, but certainly for the rule of flapping in English the hypothesis is at least plausible. Flapping, glottalization, and glottalling interact with prosodic structure, both in American and British English (cf. Kahn, 1976, Gussenhoven, 1986):

(24) a. American English Flapping and Glottalization:
(a caɾ or a ratˀ) and a hatˀ. *vs.* a catˀ(and a raɾ or a hatˀ.)

b. British English Glottaling:
(a cat or a raʔ) and a haʔ. *vs.* a caʔ(and a rat or a haʔ)

In the American English example flapping occurs between lower boundaries but not higher boundaries. In the British English example glottalization occurs before higher but not lower boundaries. Prosodically conditioned processes can be used to encode prosodic ranks.

But the choice of the line of the grid within which flapping occurs is not determined by grammar. There is a possible rendition of (a) in which flapping occurs in the entire utterance. It is this kind of variation in what line is picked as the relative line for certain rules to apply which is unexpected under theories that employ designated syntactic constituents to delimit prosodic domains; it is expected under the relational view.

The theory here is still able to capture some of the generalizations that were the motivation for the prosodic hierarchy. For example, if a rule applies within a domain P_i it necessarily occurs within all domains P_j, $j < i$. However, it assumes a greater flexibility with respect to how boundaries of a certain strength relate to syntactic nodes compared to theories that assume designated syntactic categories. This greater flexibility is indeed attested for rules such as flapping in English.

The two predictions of the relational rule can be summarized as follows: (i) when in a structure two constituents phrase together and form a single domain for some phonological rule that applies in ϕ, then by making one of them more complex and thus requiring a weaker boundary to be realized within this constituent they can be forced into separate phrases. Such effects are amply discussed in the literature, often under the heading 'branchingness', e.g. Zec and Inkelas (1990), Dresher (1994), Elordieta et al. (2003), and Prieto (forthcoming). The second prediction is that (ii), when two constituents do not phrase together in a single phrase ϕ, then by increasing the speech rate and/or embedding it deeper into a structure such that the boundary between the two constituents must be realized as being weaker than some other boundary, the two phrases should be able to phrase together.

The question is then how well these predictions are confirmed by empirical data. In order to test this in detail in the classic cases discussed in the literature on prosodic phonology, e.g. data from Chi Mwi:ni (Kisseberth and Abasheikh, 1974, Selkirk, 1986), Ewe (Clements, 1978, Selkirk, 1986), or Xiamen Chinese (Chen, 1987, Selkirk, 1986), more evidence with different numbers of levels of embedding and phrasings at different speech rates would be needed than are available in the original papers. I will not discuss the existing data here for reasons of time and space. The issue has to remain unresolved at this point, and future research has to bring to light on how closely syntactic and prosodic categories are tied together.

It is of interest in this context that some of the segmental rules that were originally taken to be categorical and used to define prosodic domains have turned out to be gradient once measurements were taken instead of using impressionistic evidence. This suggests that they cannot be used as evidence for designated categories since, on the contrary, they quantitatively distinguish degrees of boundary strengths. E.g., Esposito and Truckenbrodt (1998) reports that there at at least two degrees of lengthening involved in raddoppiamento sinttatico, a phonological rule in Florentine Italian that has played a big role in the shaping of theories on prosody—but crucially based on the assumption that it is a categorical test for a single type of prosodic domain, the phonological phrase. Similarly, gradient results were found in a yet unpublished study on raddoppiamento sinttatico conducted by Mirco Ghini (p.c.).

In theories that operate with designated syntactic categories, (Selkirk, 1986, Chen, 1987, Hale and Selkrik, 1987, Selkirk, 1995b, Truckenbrodt, 1995, 1999), each conjunct in a coordinate structure involving names should be mapped to the prosodic category corresponding to XPs (maybe by XP-alignment). Say the designated category is the phonological phrase, Φ. But this theory runs into problems in capturing the type of data pattern observed in chapter 2. The predicted phrasing for *every* coordinate structure would be flat:

(25) * (Lysander)$_{\Phi}$ (and Hermia)$_{\Phi}$ (or Demetrius.)$_{\Phi}$

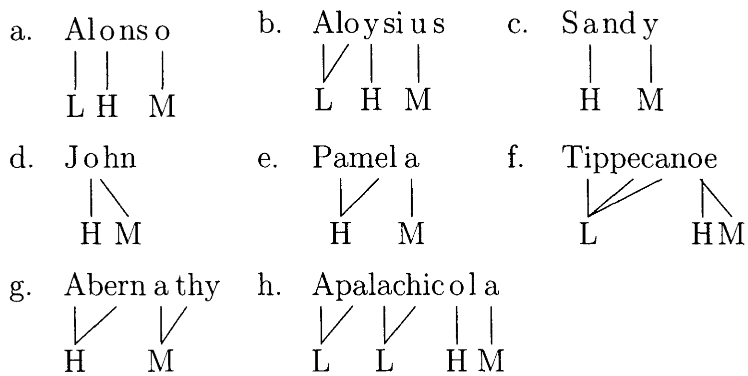
The hierarchical prosody that is actually observed constitutes 'XP-within-XP' effects. Similarly, the sentence coordinations investigated by Ladd (1988) (example (8)) and in Fery and Truckenbrodt (2004) constitute 'Sentence-within-sentence' effects. A hierarchical prosody is observed, although the elements are each of the *same* syntactic category, and hence should map to the same prosodic category. These Russian Doll effects point to a more recursive view of prosody, as it is pursued in Ladd (1986), Kubozono (1989, 1992), Ladd (1988, 1996) and Dresher (1994).

4.3.3 Tone Domains and Boundary Tones

How do the derived metrical grid structures relate to tonal elements such as boundary tones and pitch accents? I assume with Liberman (1975) and Gussenhoven (1990), Gussenhoven and Rietveld (1992) that tones are mapped to the metrical grid by principles that align independent intonational elements or 'tunes' to the metrical structures.

Liberman (1975) uses complex tunes and their alignment to motivate this view of how accents relate to prosody. One of his examples is the vocative chant Liberman (1975, 31). The melody that it consists of involves the following sequence of tones: [L H M]. Liberman's point is that the tune can be identified independently of the text, and should be represented separately. The alignment with the segmental material of the constituent that is chanted depends on its prosodic structure:

(26) Tune-to-Text Alignment



The generalizations about the placement of the tune are the following, according to Liberman:

(27)
a. High tone associates with main stress of the text, and with syllables intervening between main stress and the point where mid tone associates.
b. Low tone on syllables preceding the main stress—if there are any.
c. If there is a secondary stress following the main stress, mid tone associates with it, and with all following syllables.
d. If syllables following the main stress are all unstressed, the mid tone is associated with last one.

e. If nothing follows the main stress, then that syllable is "broken" into two distinct parts, the second of which receives the mid tone.

The grammar of tune-to-text alignment will not be addressed in this chapter. But an important point can be made by considering the nature of the conventions summarized above: they do not make reference to the fact that the examples that are used are words with a particular foot structure. Liberman (1975) could just as well have used bigger syntactic objects to make his point. The same alignment principles apply (I will ignore the finer alignment of the tune with the unstressed syllables below, just as the relative prominence within the pre-nuclear domain that aligns with the L-tones):

```
(28)  a. Lysander and Dem e tri u s!
            |              |   |
            L              H   M
      b. Lysander, Hermia and Dem e trius!
            |        |            |  |
            L        L            H  M
      c. Lysander and Dem e trius arrived!
            |              |        |
            L              H        M
```

The line that is relevant for the alignment with the intonational tune in the representation that I use here is the top line of the grid. How many lines the top line dominates depends on how complex a syntactic structure is. The theory of text-to-tune alignment was further developed in Pierrehumbert (1980).[6]

Gussenhoven (1990) and Gussenhoven and Rietveld (1992) similarly develop and test a theory of how tonal domains are constructed based on a prosodic representation. These studies assume prosodic constituents that are distinguished by labels, and yet they do not tie those labels to specific tonal correlates. Any size of prosodic constituent (foot, word, phrase) can end up providing the prominences and boundaries with which pitch accents and boundary tones are aligned. The parallel is that the top line in the derived prosodic representation is the relevant line to define tonal domains.

[6] A similar view of how tone placement relates to the grid is also at the heart of the account of African tone systems in Purnell (1997).

To summarize, there is some evidence that intonational tunes are not tied to particular lines in a prosodic hierarchy. Tonal elements are assigned to certain segments in the structure depending on where they link in the prosodic structure that they form part of.

Embedded intonational domains with different tonal properties, as they are reported in Beckman and Pierrehumbert (1986, et seq.) and assumed in the ToBI-annotation system (Silverman et al., 1992, Beckman et al., 2005) are in principle compatible with the theory presented here. But the same reasoning that was applied to domains of rule application applies here: the prediction is that tonal domains of a particular type cannot be tied to designated syntactic categories.

Some boundary tones actually carry meaning and have been treated as 'morphemes' (cf. Gussenhoven (1984), Pierrehumbert and Hirschberg (1990), and Selkirk (to appeara) and references therein.). They have been tied to speech act operators in Potts (2005). But then they do not give evidence for any particular prosodic unit—they simply attach to the prosodic domain they are affixed to, whichever line in the grid that may be. Similarly, sentence-final particles that encode similar meanings in many languages are not used as criteria for certain prosodic categories in those languages (although of course they might correlate with prosodic features).

That intonational patterns may really be part of the syntactic structure itself was first proposed in Stockwell (1960), who introduced phrase structure rules of the following kind: $S \rightarrow Nuc\ IP$, where IP stands for 'intonational pattern' and Nuc is then mapped NP VP, i.e. the actual sentence minus its intonation contour.

An updated version of this claim is to say that part of the intonation contour is the spell-out of syntactic heads of speech act operators that take sentences as their complements. Such operators were proposed and motivated in Jacobs (1991c)).

Summary

This section reviewed quantitative and qualitative evidence for prosodic structure. Quantitative durational effects at boundaries and pitch scaling effects can both be captured by the relational grid proposed in chapter 2. Both *boundary strength scaling* and *pitch accent scaling* give evidence for relative nested grouping of prosodic domains. The observed effects differentiate prosodic domains relative to each other, but do not distinguish different categories of prosodic domains.

Qualitiative distinctions of different types of domains, as are usually assumed in the literature on the 'prosodic hierarchy', are compatible with the presented theory. However, the relational theory is incompatible with the idea that particular syntactic categories map to particular prosodic categories. Instead, there should be flexibility in how the prosodic means are employed to realize the relative boundary ranks assigned by the syntax-to-phonology mapping. The assumption that certain syntactic constituents map to certain prosodic categories is common in the literature on prosodic phonology, but the state of the empirical evidence is not conclusive, and future inquiry will have to further investigate the issue.

One theory that seems particularly attractive in light of the evidence for recursion in coordinate structures is the one proposed in Dresher (1994, 39) based on evidence from Tiberian Hebrew. Dresher proposes that there are prosodic domains with different category labels, and motivates them based on the observation that they provide the domain for certain rhythmic and segmental phonological processes. At the same time, he presents evidence for recursion *within* those categories, and postulates that phonological rules do not have access to these within-category hierarchical distinctions. Similar within-category recursion was argued for in Ladd (1988).

So far, I have not discussed factors other than syntax that play a role in prosodic structuring. In particular, I have not discussed the role of eurhythmic principles on phrasing. For example, prosodic branchingness and principles that call for phrases of equal prosodic weight also seems to affect the choice in prosodic phrasing, as observed in Ghini (1993a,b). Eurhythmic principles are invoked in many theories of prosodic phrasing (i.a. Dresher, 1994, Selkirk, 1995b, Sandalo and Truckenbrodt, 2002, Elordieta et al., 2003, Prieto, forthcoming, Pak, 2005). Some effects of rhythm will be discussed in chapter 6, but a detailed discussion of eurhythmic effects goes beyond the scope of the present study.

Chapter 5

Boundary Strength Scaling

Abstract

Prosodic grouping reflects syntactic structure. This chapter presents evidence that utterances are structured hierarchically by imposing boundaries of various ranks, reflecting syntactic levels of embedding, as proposed in chapter 2. Prosodic boundaries are scaled relative to boundaries produced earlier in the utterance. This phenomenon that will be called 'Boundary Strength Scaling' in analogy to pitch-accent scaling, which similarly depends on the hierarchical prosodic organization of an utterance.

5.1 Scaling the Strength of Boundaries

This chapter reports on an experiment testing the boundary ranks in coordinate structures assigned by the algorithm presented in chapter 2. The generalization about boundary ranks proposed relates the strength of boundaries to syntactic levels of embedding. The experimental evidence shows that the model is better than a competing model from the processing literature that assigns boundaries more locally only looking at the two adjacent syntactic constituents (Watson and Gibson, 2004a,b) .

A closer look at the data reveals that speakers scale boundaries relative to boundaries produced earlier in an utterance. I propose that the implementation of prosodic structure allows for 'Boundary Strength Scaling'. Boundary Strength scaling makes global syntactic influences on prosody compatible with incremental structure building in on-line production.

5.2 Global Syntactic Effects

Syntactic, phonological, and processing factors influence prosodic phrasing. Early accounts of phrasing in the processing literature (Cooper and Paccia-Cooper, 1980, Grosjean and Collins, 1979) explain effects on prosody in reference to syntactic structure. Gee and Grosjean (1983), Ferreira (1988, 1993) map syntactic structure to a prosodic structure, allowing for certain mismatches, e.g. in order to account for the phrasing of function words.

Watson and Gibson (2004a,b) argue for a local processing-based model that uses only the length of the constituent to the left and right of a boundary (LRB-Theory) to predict its strength. The LRB thus greatly simplifies the mapping between syntax and prosody, and is compatible with incremental structure building since only local information is needed.

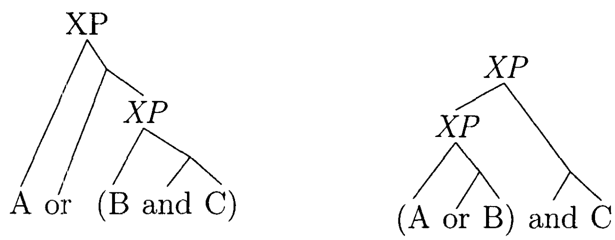
This section presents new evidence for a more syntactic approach to prosodic phrasing. In particular, it presents evidence for the boundary ranks assigned by the algorithm in chapter 2.

5.2.1 Coordinate Structures as a Test Case

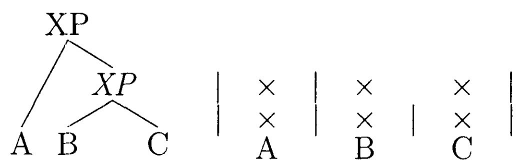
In coordinate structures, different bracketings are coded by parentheses in orthography, and by prosody in speech:

(1) Coordinate Structures

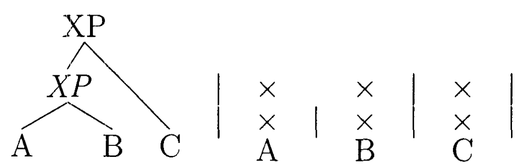
a. Right-Branching   b. Left-Branching



Prosodic phrasing reflects the bracketing of syntactic structure. In (1a), the boundary following A is stronger than the boundary following B:

(2) 

In (1b), the reverse is true:

(3) 

I am using the adapted version of the bracketed metrical grid introduced in chapter 2 to code prosodic structure. Each grid line is structured into feet (noted by the 'pipe'-symbol). A foot boundary at a given line in the grid implies a foot boundary on all lower lines. This leads to a strictly layered hierarchy of prosodic structure. The boundary rank, i.e., the strength of a given boundary, corresponds to the number of the grid line.

The data set used in this paper to test theories of prosodic phrasing consists of coordinate structures with 4 elements. There are exactly 11 different bracketings for such cases:

(4) a. Four elements at the top level:

Morgan and Joey and Norman and Ronny

b. Three elements at the top level:
Morgan or Joey or (Norman and Ronny)
Morgan or (Joey and Norman) or Ronny
(Morgan and Joey) or Norman or Ronny

c. Two elements at the top level:
Morgan or (Joey and Norman and Ronny)
(Morgan and Joey and Norman) or Ronny (Morgan and Joey) or (Norman and Ronny)

d. Two levels of embedding:
Morgan or (Joey and (Norman or Ronny))
Morgan or ((Joey or Norman) and Ronny)
(Morgan and (Joey or Norman)) or Ronny
((Morgan or Joey) and Norman) or Ronny

This data set serves to test theories of boundary assignment in prosody.

5.2.2 Two Theories

The LRB (Watson and Gibson (2004a,b)) relates the likelihood of an intonational phrase break at a given boundary to the complexity (measured in number of phonological phrases) of the two adjacent constituents:[1]

(5) The Left-Hand/Right-Hand Boundary (**LRB**) Hypothesis: The likelihood of a boundary between the LHS and RHS constituents increases with the size [counted in # of phonological phrases, MW] of each constituent. [due to processing cost of upcoming/preceding constituent; MW]

For the examples above, a stronger boundary occurs exactly where a complex constituent (underlined in the following example) follows (1a) or precedes (1b). The stronger boundaries are those represented by pipes on the top grid line.

[1]Effects of constituent length of the likelihood of a prosodic break based on impressionistic data have been observed as early as Bierwisch (1966a) and Martin (1971).

(6) Following and Preceding Complex Constituents

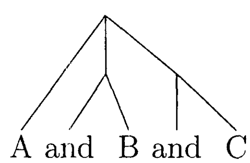
a. A or (B and C) b. (A or B) and C

```
| × | ×   × |      | ×   × | × |
| × | × | × |      | × | × | × |
  A   B   C          A   B   C
```

One advantage of the LRB over previous theories, according to Watson and Gibson (2004a), is that it is compatible with incremental processing: it is sufficient to look at the two adjacent syntactic constituents in order to determine whether or not a certain type of boundary occurs.

In order to account for coordinate structures in which more than 2 conjuncts are connected with a 'flat' prosody, the LRB would have to assume a syntactic representation with n-ary branching. The prosodic breaks between the conjuncts are then predicted to be equal to one another.[2]

(7) Flat Tree, Flat Prosody




```
| ×   ×   × |
| × | × | × |
  A   B   C
```

The connector 'and' would be ignored in evaluating the complexity of constituents, since they do not form separate phonological phrases. At each boundary, both the preceding and following constituent are simplex. The likelihood increases with every additional phonological phrase within an adjacent constituent. The LRB assumes then that there is a notion of phonological phrasing that itself is explained by some other factor, and makes predictions about the likelihood of intonational boundaries.

The complete predictions of the LRB for the data set discussed here (the data set in (4)) is summarized in table 5.1.

The alternative theory to be considered here is one in which the rank of a boundary reflects the syntactic embedding:

[2]The connectors themselves form a constituent with the following conjunct. See chapter 2 for discussion of the syntactic representation of coordinate structures, and for a discussion of unbounded vs. binary branching.

		A	B	C	D
1	A or B or C or D		1	1	1
2	A or (B and C) or D		2	1	2
3	A or B or (C and D)		1	2	1
4	(A and B) or C or D		1	2	1
5	(A and B) or (C or D)		1	3	1
6	(A and B and C) or D		1	1	3
7	A or (B and C and D)		3	1	1
8	((A or B) and C) or D)		1	2	3
9	(A and (B or C)) or D		2	1	3
10	A or ((B or C) and D)		3	1	2
11	A or (B and (C or D))		3	2	1

Table 5.1: *Predictions of the LRB*

(8) Scopally Determined Boundary Rank (**SBR**):

If Boundary Rank at a given level of embedding is n, the rank of the boundaries between constituents of the next higher level is n+1.

In coordination structures, conjuncts/disjuncts at the same level are separated by boundaries of the same rank; if further embedding is introduced, boundaries of lower ranks encode the boundaries between the further nested constituents.

Again, if we use unbounded branching trees, there can be more than 2 constituents at one level of embedding, so coordinate structures with flat prosodies receive a prosody with equal boundaries separating the conjuncts.

In chapter 2, the boundary ranks of the SBR are derived by a cyclic syntax, using only binary branching trees. A 'flat' coordinate structure in which each conjunct is set off by a boundary of equal rank is one that is created in one single cycle; an 'articulated' coordinate structure with several different prosodic boundary ranks is one that was created in separate cycles.

The choice between binary branching trees with a derivational syntax and unbounded trees is not relevant for this chapter, and I will not consider it in the following.

The predictions for relative boundary strength for the examples in (6) and (7) are the same for the LRB and the SBR.

The SBR, however, predicts long distance effects for more complicated cases. The complexity of a constituent non-adjacent to a boundary can influence its relative rank.

Consider the following two examples, in particular the rank of boundary at the point where I put the placeholder '$\diamondsuit$':

(9) a. A $\diamondsuit$ or B || or (C | and D)

b. (A | and B) || or C $\diamondsuit$ or D

The LRB predicts that the boundaries at '$\diamond$' are weak, since both adjacent constituents are simplex. The SBR predicts that these boundaries are strong, since they separate constituents at the higher level of embedding. This global long-distance effect would seem to be incompatible with incremental structure building.

The complete prediction for the data set in (4) are summarized in table 5.2. The cells in which the LRB and SBR differ are highlighted by boldface:

		A B C D		
1	A or B or C or D	1	1	1
2	A or (B and C) or D	2	1	2
3	A or B or (C and D)	**2**	2	1
4	(A and B) or C or D	1	2	**2**
5	(A and B) or (C or D)	1	**2**	1
6	(A and B and C) or D	1	1	**2**
7	A or (B and C and D)	**2**	1	1
8	((A or B) and C) or D)	1	2	3
9	(A and (B or C)) or D	2	1	3
10	A or ((B or C) and D)	3	1	2
11	A or (B and (C or D))	3	2	1

Table 5.2: *Predictions of the SBR*

This paper presents evidence that (i) long-distance effects are real, as predicted by the SBR; (ii) the SBR is compatible in principle with incremental structure building; and (iii), boundaries are scaled in strength relative to already produced boundaries.

In a production experiment, the following types of dialogues were presented using powerpoint:

(10) Experimenter: Who will solve the problem first?

Subject: Morgan and Joey are working together. Norman is working on his own. and so is Ronny. So one of the following will solve it first:

Morgan and Joey or Norman or Ronny.

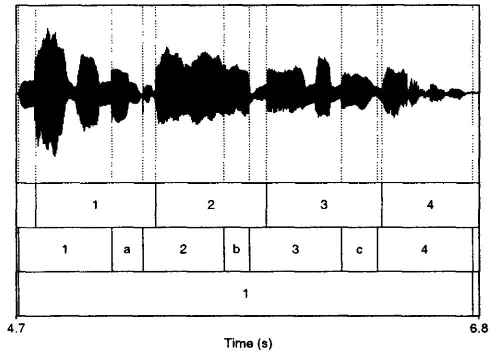


Figure 5-1: *Annotation Scheme*

The task was made easier by visual illustrations of the situations, further clarifying the grouping. The entire dialogue was recorded. The fragment answer (in boldface) was repeated three times; two sets of proper names were used for each of the 11 bracketings. 4 unpaid subjects were recorded. The data were annotated for word boundaries (second tier of annotation in (Fig. 5-1), and durations were measured using the Praat speech software (Boersma and Weenink, 1996).

The length of each constituent was normalized relative to the total duration of the answer. The mean of the three repetitions was computed, and constituent length was taken as a measure of the strength of the upcoming boundary, under the uncontroversial assumption that final pre-boundary lengthening increases with boundary strength (Price et al. (1991), Wightman et al. (1992) and references therein).

5.2.3 Syntactic Long-Distance Effects

Are there global effects of syntax on prosody that do not reduce to effects of adjacent constituents? In the following I will present evidence from a production experiment that there are. The evidence is in tune with the SBR but presents a problem for the LRB.

Effects of Adjacent Constituents

The LRB was designed to predict the likelihood of an intonational/intermediate phrase break[3]. The present study looked at boundary strength as reflected by duration of the preceding constituent. The advantage of this measure is that the annotation does not presuppose a theory of phrasing, and no labeling of prosodic categories (such as intonational phrase or intermediate phrase as in a ToBI-labeling) is necessary.

Watson and Gibson (2004a,b) used a different measure, the likelihood of intonational phrase break at a given boundary. This section presents evidence that the measure used here taps into the same phenomenon. Quantitative durational effects reflect the length of adjacent constituents. The results confirm an effect of adjacent constituents reported in Watson and Gibson (2004a,b). While Watson and Gibson (2004a,b) used the likelihood of intonational boundaries as the dependent measure, the quantitative measure of duration effect investigated.

First the effect of the length of the constituent following the boundary was tested by looking at the boundary strength after the first proper name in the structures in table 5.2.3.

						LRB	SBR	Cond
Baseline	1	A ◇ and	**B**	and C	and D	1	1	0
	2	A ◇ or	**B**	or C	or D	1	1	0
Simplex	3	(A ◇ and	**B**)	or C	or D	1	1	-1
	4	(A ◇ and	**B**	and C)	or D	1	1	-1
Complex	5	A ◇ or	(**B**	**and C**)	or D	2	2	1
	6	A ◇ or	(**B**	**and C**	**and D**)	3	2	1

Table 5.3: *Simplex vs. Complex Constituent Following the Boundary*

In the structures (1–4) in table 5.2.3, the constituent following the first connector is simplex and consists of only one proper name. In (5–6), the constituent following the first connector is complex and consists of 2 and 3 proper names respectively. The length of the first constituent (A), normalized relative to the overall duration of the utterance, was used as a measure of the strength of the following boundary.

[3]The two levels were conflated since inter-annotator reliability did not warrant making the distinction.

The predictions of the two theories for the three conditions 'Baseline', 'Simplex', and 'Complex' are the same: both SBR and LRB predict 'Baseline' and 'Simplex' to the same, and 'Complex' to be different from each of them.[4]

The omnibus ANOVA for the normalized length of constituent depending on the variable 'complexity' was highly significant $F(1,4) = 85.13$, $p < 0.000$. The means are plotted with the standard error in Fig. 5-2. The labels -1, 0, 1 reflect the value in the column 'COND' in table 5.2.3.

To test the effect of constituent complexity, pairwise post-hoc comparisons (Tukey) were conducted. Each pairwise comparison between 'baseline', 'simplex', and 'complex' was significant.

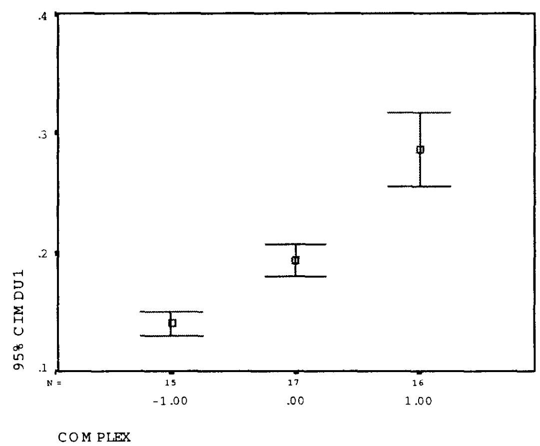


Figure 5-2: *Complexity of Following Constituent*

The contrasts between the conditions 'baseline' and 'simplex' on the one hand and 'complex' on the other are as expected by the LRB, and replicates the results of Watson and Gibson (2004a). They are also expected under the SBR.

However, the observed difference between 'baseline' and 'simplex' is unexpected both for the LRB and the SBR. In both conditions, the two constituents adjacent to the first boundary are simplex, and a boundary of equal strength is expected for the

[4]There is a difference for the predictions of the two theories with respect to the items within condition 'complex' (5 *vs.* 6): the LRB predicts an effect of the length of the following constituent, while the SBR does not. I will return to this difference between the SBR and the LRB in section 5.3.

LRB; they are on the lowest level of embedding, so a boundary rank of 1 is expected for the SBR.

The results for the two conditions differ, however. It matters whether or not *later in the utterance* there is a stronger boundary coming up (as in 3,4) or not (as in 1,2). This constitutes a long-distance effect, and shows that it is *not* sufficient to only consider two adjacent constituents at a given boundary. I will return to this effect in 5.3.

A second test was conducted to evaluate whether the complexity of the constituent *preceding* the boundary matters. Here, we are consdering the boundary that separates the last conjuncts. The boundary strength was again measured by relative pre-boundary lengthening, that is in this case lengthening of the proper name 'Norman', which directly precedes the boundary. The data is summarized in table 5.4.

								LRB	SBR	Cond
Baseline	1	A	and	**B**	and	C	◇ and D	1	1	0
	2	A	or	**B**	or	C	◇ or D	1	1	0
Simplex	3	A	or	B	or	(**C**	◇ and D)	1	1	-1
	4	A	or	(B	and	**C**	◇ and D)	1	1	-1
Complex	5	A	or	(**B**	**and**	**C**)	◇ or D	2	2	1
	6	(**A**	**and**	**B**	**and**	**C**)	◇ or D	3	2	1

Table 5.4: *Simplex vs. Complex Constituent Preceding the Boundary*

In the structures (1–4) in table 5.4, the constituent preceding the last connector is simplex and consists of only one proper name. In (5–6), this constituent is complex and consists of 2 and 3 proper names respectively. This time, the normalized length of the first constituent (C) was used as a measure of the strength of the following boundary.

The omnibus ANOVA for the normalized length of constituent depending on condition was highly significant $F(1,4) = 19.67$, $p < 0.002$. The means are plotted with the standard error in Fig. 5-3. The labels -1, 0, 1 reflect the value in the column 'COND' in table 5.4.

Pairwise post-hoc comparisons were conducted to test whether the three conditions were statistically different. As predicted by the LRB, both 'Baseline' and

'Simplex' were different from 'Complex'. The SBR also predicts this difference.

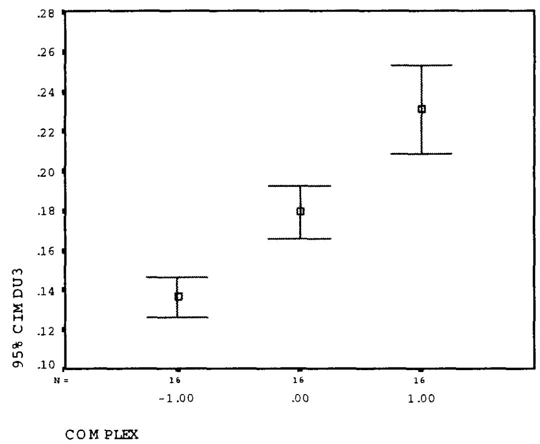


Figure 5-3: *Complexity of Preceding Constituent*

In addition, there was a significant difference between 'Baseline' and 'Simplex'. It matters whether or not a stronger boundary precedes in the structures. Once again, this effect is unexpected by the either the LRB or the SBR. And once again, it constitutes a long-distance effect.

The results presented so far confirm the claim in Watson and Gibson (2004a) that boundary strength is influenced by the complexity of adjacent constituents. While Watson and Gibson (2004a) were concerned with the likelihood of an intonational phrase break, the measure used in this study is quantitative lengthening.

The tests also showed initial evidence for long-distance effects. The effects observed here were not predicted by either the LRB or SBR. I will return to the question of how they can be explained, and suggest that the SBR can account for them once we understand boundary rank as a relative rather than an absolute notion.

5.2.4 Effects of Non-Adjacent Constituents

One major difference between the SBR and the LRB resides in the predictions for the examples in (9), repeated below:

(9) a. A ◇ or B || or (C | and D)

b. (A | and B) or || C ◇ or D

The LRB predicts a weak boundary at '◇' in (9):

(11) a. Morgan | or Joey || or Norman | and Ronny.

b. Morgan | and Joey || or Norman | or Ronny.

The SBR predicts a strong boundary:

(12) a. Morgan || or Joey || or Norman | and Ronny.

b. Morgan | and Joey || or Norman || or Ronny.

The data set used to decide between theories and the coding of conditions is summarized in Table (5.5).

								SBR	LRB	Cond
1	A ◇	or	B	or	C	or	D	1	1	11
2	A	or	B	or	C ◇	or	D	1	1	11
3	A ◇	or	B	or	(C	and	D)	2	1	21
4	(A	and	B)	or	C ◇	or	D	2	1	21
5	A ◇	or	(B	and	C)	or	D	2	2	22
6	A	or	(B	and	C) ◇	or	D	2	2	22

Table 5.5: *Data set for Testing Long-Distance effect*

The omnibus ANOVA for 'CND' was highly significant: $F(2,6) = 57.27$; $p < 0.000$. The means are plotted against the standard error in Fig. (5-4). The label '11' means that in this condition both theories predict a boundary of rank 1 and label '22' means that both theories predict a label of rank 2. The interesting condition is '21'. Here, the SBR predicts a boundary of rank 2, and the LRB a boundary of rank 1.

Post-hoc Tukey comparisons show differences between the 11 vs. 21 ($p < 0.045$) and 11 vs. 22 ($p < 0.008$), but not between 21 vs. 22 ($p < .66$) (cf. Fig. (5-4)). The results are as predicted by the SBR. The boundaries at '◇' in (9) are strong, and they are strong due to the complexity of a non-adjacent constituent.

There are syntactic long-distance effects on prosody. The LRB does not account for the effect of the complexity of a non-adjacent constituent, but the SBR does, although not in all cases.

The next section will show evidence that boundaries are scaled in strength relative to upcoming and preceding boundaries. These results will then be used in section 5.3

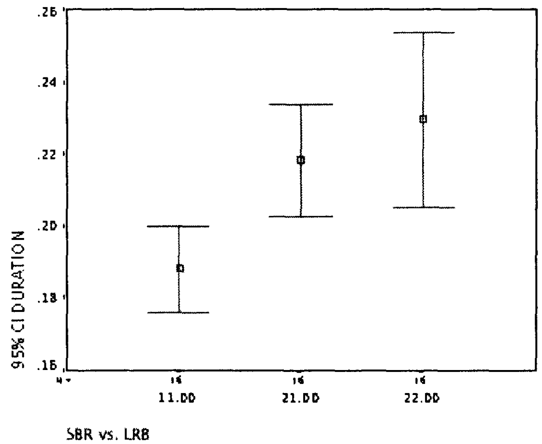


Figure 5-4: *Long Distance Effects*

to define a normalization based on the SBR, which is able to account for all of the effects reported in this section.

5.2.5 Looking Ahead

There is another kind of long-distance effect. In the realization of the first constituent, the complexity of the upcoming structure far beyond the adjacent constituent seems to play a role.

The metrical representation assigned in chapter 2 provides only relative boundary ranks. The phonetic and phonological implementation should reflect these boundaries, or at least should not contradict them. Durational cues are an important factor in the perceived boundary rank. If speakers anticipate the upcoming structure, then they might adjust the realization of the first constituent such that it allows them to realize the rest of the structure in such a way that the relative boundary ranks are correctly implemented within the phonetic range of their articulators.

The expectation is that boundaries are realized as stronger or weaker depending on how many lower or higher ranks will have to be realized later. The experiment was run in a way that certainly allowed the subjects to anticipate and plan the answer carefully. If there are look-ahead effects, they should be measurable in the length of the *first* constituent. Varying the length of the first constituent can be used to set up for upcoming stronger or weaker boundaries. The test-data used to probe for this effect are the following:

													BND
1		A	or	B		or		C		or	D	Baseline	0
2		A	and	B		and		C		and	D	Baseline	0
3		A	or	B		or	(	C		and	D)	Lower Rank	-1
4		A	or	(B		or		C		and	D)	Lower Rank	-1
5		A	or	(B		and	(	C		or	D)	Lower Rank	-2
6	(	A	and	B	)	or		C		or	D	Higher Rank	1
7	(	A	and	B		and		C	)	or	D	Higher Rank	1
8	(	A	or	B	)	and		C	)	or	D	Higher Rank	2

Table 5.6: *Look-Ahead*

The omnibus ANOVA for BND was highly significant: F(4,12) = 2.28; $p < 0.000$. The plot of the means reveals that the length of the first constituent varies by and large as expected with the number of upcoming weaker/stronger constituents:

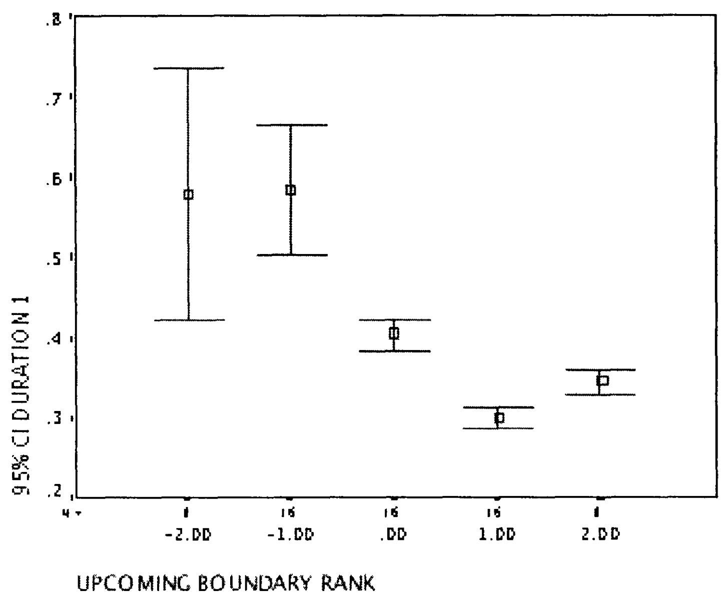


Figure 5-5: *Look-Ahead Effect*

This look-head effect might reflect speech planning, and thus be due to considerations that lie at the heart of the proposal in Watson and Gibson (2004a).

The LRB, however, predicts an effect that relates to the length of an upcoming constituent—independent of the degree of embedding. We can use the following data set to test whether this effect is real.

		LRB	SBR	Length
1	A or B or (C and D)	1	2	1
2	A or (B and C) and D	2	2	2
3	A and (B or C or D)	3	2	3

Table 5.7: *Length of the Upcoming Constituent*

The SBR predicts an equal boundary rank at the relevant boundary. The LRB predicts that the boundary should get stronger with an increasing length of the following constituent.

The omnibus ANOVA for the variable 'Length' approaches significance. F (1,4) = 4.80; $p < 0.056$. The means are plotted in figure 5-6. The means of the conditions summarized in the column NST in Table 5.7 are summarized in Fig. 5-6.

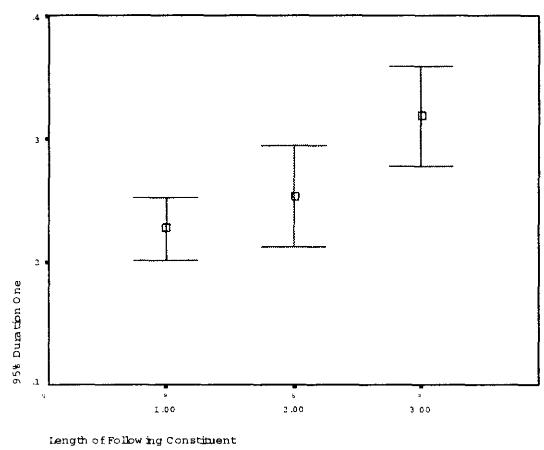


Figure 5-6: *Length of Following Constituent*

Just as predicted by the LRB, there does seem to be an effect of the length of the upcoming constituent, even if it is not quite significant. The effect is a look-ahead effect, in that it influences the realization of the first constituent, depending on the length of the second constituent.

5.2.6 Looking Back

The realization of prosodic structure may also be affected by the prosodic shape of structure produced earlier. Again, this effect takes more into account than just the preceding constituent.

Boundary strength scaling relative to an already produced boundary can be achieved by compressing or inflating constituents realized later in the sentence. The prediction is that the length of the *final* constituent should vary depending on whether or not it is in an early- or late-nesting structure. 'Early nesting' means that a constituent with a low level of syntactic embedding early in the utterance is followed by higher material; 'late-nesting' means that a constituent with a low level of embedding occurs late in the utterance. The data set to test for an effect of Boundary-Strength Scaling is summarized in table 5.8.

									NST
1	A	or	B	or	C	or	D	Baseline	0
2	A	and	B	and	C	and	D	Baseline	0
3	A	or	B	or	(C	and	D)	Late Nesting	-1
4	A	or	(B	and	C	and	D)	Late Nesting	-1
5	(A	and	B)	or	C	or	D	Early Nesting	+1
6	(A	and	B	and	C)	or	D	Early Nesting	+1

Table 5.8: *Late-Nesting vs. Early-Nesting and Look-Back*

The omnibus ANOVA for 'Nesting' approaches significance: $F(2,6) = 4.34$; $p < 0.068$, the means are as predicted (Fig. 5-7).

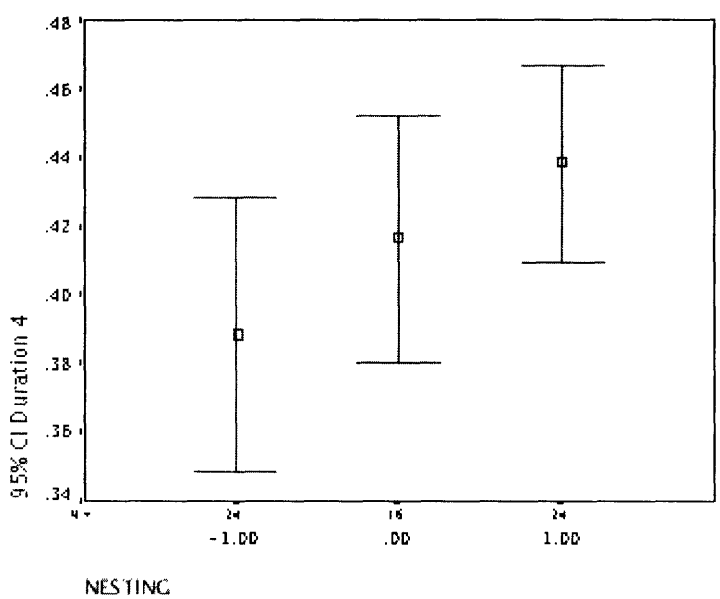


Figure 5-7: *Looking Back: Nesting*

The results are not quite significant. But if look-back effects are real, then boundaries can be scaled as stronger or weaker relative to already produced prosodic structure, in order to realize the relative boundary ranks that are imposed by the syntactic structure—this strategy can be used in the case of long-distance effects, if the look-ahead didn't make the preceding boundaries sufficiently strong in anticipation of the weaker boundary.

The variation in the duration of the *last* constituent is due to variation in the level of embedding of the last constituent—it is not due to pre-boundary lengthening. The last constituent is always at the end of the utterance, therefore the following boundary should always be a strong one.

There is a look-back effect that relates to the rank of the preceding boundary. The LRB predicts an effect that relates to the length of the preceding constituent, independent of the degree of embedding. We can use the data set in table 5.9 to test whether this effect is real.

											LRB	SBR	Length
1	(	A	or		B	) or	C		◇	and D	1	2	1
2		A	or	(	B	and	C	)	◇	and D	2	2	2
3	(	A	and		B	or	C	)	◇	or D	3	2	3

Table 5.9: *Length of the Preceding Constituent*

The SBR predicts equal boundary ranks at the relevant boundary ('◇') in the three structures. The LRB predicts that the boundary should get stronger with an increasing length of the preceding constituent.

The length effect is significant. The omnibus ANOVA for the variable 'Length' shows a clear effect. $F\ (1,4) = 13.12$; $p < 0.006$. The means are plotted in figure 5-8.

As predicted by the LRB, there is an effect of the length of the upcoming constituent. But note that the means do not relate the length of the relevant constituent with the length of the preceding constituents in a straightforward way. This suggests we are not not looking at the data in the right way yet. The remainder of the paper presents a different way of looking at the data, which captures all of the effects reported so far.

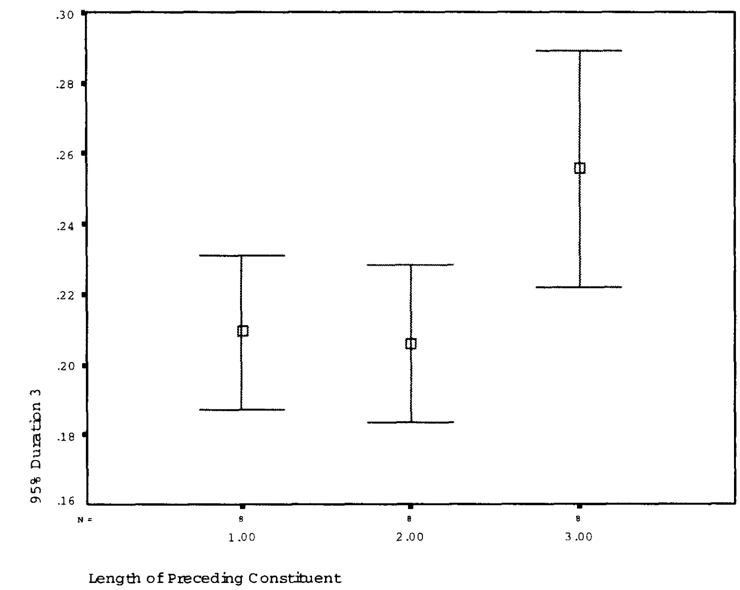


Figure 5-8: *Looking Back: Length*

5.3 Left-to-Right Scaling

What is a good model of the various kinds of lengthening effects observed in this chapter so far? The SBR predicts some but not all long-distance effects, the LRB predicts length-based effects, but fails to predict any long-distance effects. Both theories are not able to explain the full range of data.

Watson and Gibson (2004a) argue that since the LRB uses only local information, it is compatible with incremental structure building, and is thus superior to global syntactic approaches such as Cooper and Paccia-Cooper (1980), Gee and Grosjean (1983), Ferreira (1988). One premise of this argument is the assumption of a fixed prosodic hierarchy:

(13) Prosodic hierarchy (Selkirk, 1986, 384):

```
(------------------------------------------------------------)   Utt
(-------------------------------------)(---------------------)   IPh
(----------)(-------------------------)(---------------------)   PPh
(----------)(-------)(---------------)(------------)(-------)   PWd
(--)(------)(------)(--)(-----------)(--------)(--)(------)     Ft
(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)(--)       Syl
```

Under this assumption, a boundary is 'stronger' if it has the label of a prosodic constituent higher on the hierarchy, and 'weaker' with a lower label. It is crucial then to start out with the right level of prosodic constituent, so that one does not run out of categories in order to produce stronger or weaker boundaries later in the sentence. To illustrate the point, consider early- vs. late-nesting structures:

(14) a. Early Nesting:
(Morgan and Joey) or Norman or Ronny

b. Late Nesting:
Morgan or Joey or (Norman and Ronny)

In (14a), starting out with an intonational phrase leading to an intonational phrase break after 'Morgan' would make it impossible to signal a stronger boundary after 'Joey'—unless one placed an utterance end at this point. Conversely, in (14b), starting out with a phonological phrase break after 'Morgan' would make it impossible to signal a weaker boundary after 'Norman'—unless one mapped 'Norman and Ronny' to a single prosodic word.

The alternative to the prosodic hierarchy is a recursive Prosodic System, in which boundaries are ranked according to a metrical representation, without a difference in categorical label between the different ranks, as proposed in chapter 2. Tones are mapped to grids indirectly, obeying alignment principles, in which the top gridline plays a crucial role (Liberman (1975), Gussenhoven (1990)). The assumption here is that exactly all top-line grid-marks are realized as pitch accents, or at least other accents are severely reduced in pitch range.

(15)

	|	×		×		×		×	|
	|	×	|	×		×		×	|
	|	×	|	×	|	×		×	|
	|	×	|	×	|	×	|	×	|
		A		B		C		D	

This system allows for ***Boundary-Strength Scaling*** (BSS), in other words, boundaries can be scaled as stronger or weaker relative to other boundaries. Carlson et al. (2001) indeed report an effect that shows that two boundaries are adjusted relative to each other in the disambiguation of attachment ambiguities.

What is crucial in the theory proposed in chapter 2 is the relative rank of a boundary, not its categorical identity. BSS could be achieved either by look-ahead, which would assure that early boundaries are weaker/stronger than upcoming boundaries, or by look-back, which would scale a boundary as weaker/stronger relative to an *already produced* boundary.

This section shows that by normalizing the length of each constituent relative to the first constituent, and by coding boundary rank relative to the boundary following the first constituent (using the SBR), we can substantially improve the coverage of the data.

The idea is to normalize the data in such a way as would make sense if speech production is indeed incremental and left-to-right. The idea of a left-to-right organization of speech has been supported in Phillips (1996, 2003) who proposes that grammatical derivations as well should incrementally build structure in in a left-to-right fashion.

The left-to-right implementation of the SBR proposed here accounts for the long-distance effects, for 'looking-back'-effects, and also for the effects of constituent length. Furthermore, it allows us to define a metric of look-ahead effects.

5.3.1 A Normalization

The look-ahead effects, as reflected in the length of the first constituents, reflect a speaker's anticipation of the complexity of upcoming structure.

The observations in the previous section suggest another factor is at play. It looks as if boundaries are scaled relative to upcoming or preceding boundaries, in order to assure that they have the right relative rank. If the system really operates with relative boundary rank rather than absolute boundary strength, it would make sense for a speaker to start out strong when weaker boundaries are coming up and vice versa. The subsequent boundaries are then scaled relative to the first boundary accordingly.

The purpose of the look-ahead variation in the first constituent is to assure that the relative ranks can be realized given the physical range that a speaker has available to articulate relative boundary strength; it may also have the purpose of assuring that the produced boundary ranks are perceptually distinct.

If this is how the system works, we should first look at the data by controlling for the variation in the first constituent, and then verify whether the rest of the structure is indeed scaled relative to that first constituent.

A simple way of doing so is by normalizing the duration of all constituents relative to the first one. This can be done by dividing the length of every constituent by that of the first constituent. The length of the first constituent is then always 1—in other words, we have neutralized the look-ahead effect.

Now we should be able to observe a relative scaling of all other material relative to the first constituent. A constituent is going to be longer or shorter depending on whether the boundary rank that follows it is stronger or weaker than the boundary rank following the first constituent. This presupposes a theory of boundary rank assignment. I will use the SBR here, and show that it allows for a good model of the observed data.

The boundary ranks predicted by the SBR are recoded in such a way that it reflects boundary relative to the first boundary, which is the baseline or the 'zero boundary rank'.

The predictions of the left-to-right version of the SBR, or normalized SBR (henceforth NSBR), are summarized in table 5.10. The table gives the same ranks as the SBR, but instead of giving an absolute value, each rank is computed *relative to the rank following the first constituent.*

		A	B	C	D
1	A or B or C or D	0	0	0	
2	A or (B and C) or D	0	-1	0	
3	A or B or (C and D)	0	0	-1	
4	(A and B) or C or D	0	1	1	
5	(A and B) or (C or D)	0	1	0	
6	(A and B and C) or D	0	0	1	
7	A or (B and C and D)	0	-1	-1	
8	((A or B) and C) or D)	0	1	2	
9	(A and (B or C)) or D	0	-1	1	
10	A or ((B or C) and D)	0	-2	-1	
11	A or (B and (C or D))	0	-1	-2	

Table 5.10: *NSBR*

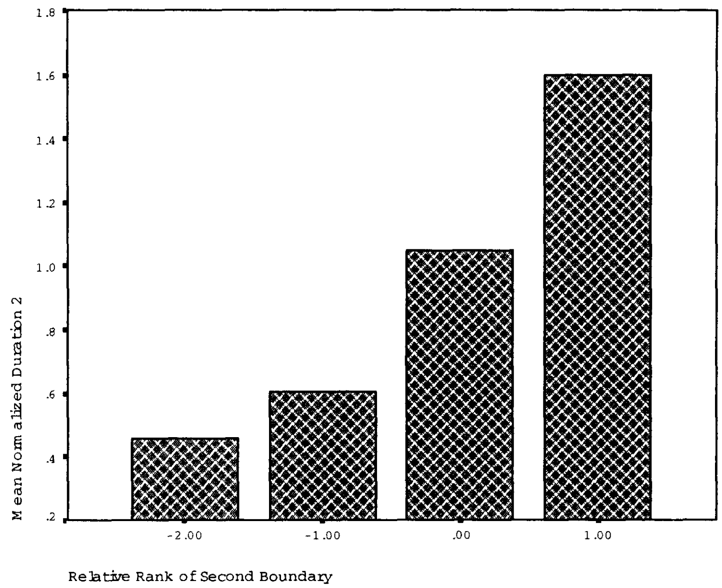


Figure 5-9: *Normalized Second Constituent*

In the following, I will illustrate some of the predictions of the new normalization, and illustrate that it provides a good model of how speakers scale boundaries.

Consider the graph in figure 5-9, which plots the means of the second constituent by the relative boundary rank.

Post-hoc comparisons show significant differences between three different levels, as shown in figure 5-10.

The left-to-right normalization proposed here is a substantial improvement over the original SBR, where normalization was achieved by dividing by the duration of the entire utterance. This becomes apparent when we compare the two theories by looking at the mean durations of the relevant constituent.

The original SBR predicts only two different ranks for the second boundary. The left-to-right SBR, on the other hand, predicts five different categories, and allows for a finer-grained understanding of the data.

A look at the third conjunct is similarly encouraging. Consider the means of the normalized length of the third conjunct, plotted against the strength of the following boundary in figure 5-11.

Post-hoc comparisons distinguish four different categories of boundary ranks (fig-

Normalized Duration of Second Constituent

	DBND1	N	Subset 1	Subset 2	Subset 3
Tukey HSD[a,b,c]	-2.00	4	.4609		
	-1.00	28	.6060		
	.00	32		1.0500	
	1.00	24			1.6012
	Sig.		.713	1.000	1.000
Tukey B[a,b,c]	-2.00	4	.4609		
	-1.00	28	.6060		
	.00	32		1.0500	
	1.00	24			1.6012

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = .104.

a. Uses Harmonic Mean Sample Size = 11.154.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

Figure 5-10: *3 Subsets defined by Post-hoc Comparisons*

ure 5-12).

The left-to-right normalization makes accurate predictions for the length of the conjuncts. In the following, I will illustrate how it can explain the long-distance effects and the various look-ahead and looking-back effects discussed above. The NSBR is a promising model of durational effects in prosodic structure. It is crucially based on the theory of metrical structure presented in chapter 2, which assigns a relative metrical structure that provides the relative boundary ranks.

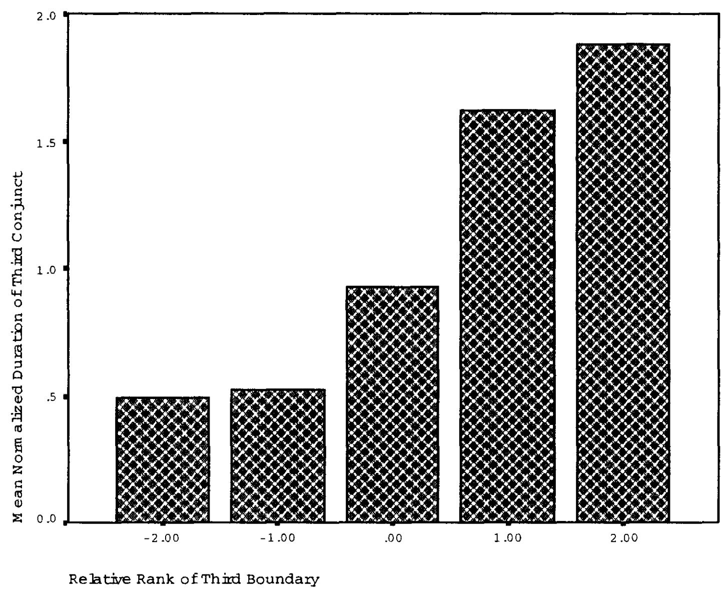


Figure 5-11: *Normalized Length of Third Constituent*

Normalized Duration of Third Constituent

	Rank Third	N	Subset 1	2	3	4
Tukey HSD[a,b,c]	-2.00	8	4955			
	-1.00	20	.5266			
	00	32		.9317		
	1 00	20			1 6244	
	2 00	8				1 8838
	Sig.		996	1 000	1 000	1 000
Tukey B[a,b,c]	-2.00	8	4955			
	-1.00	20	5266			
	.00	32		.9317		
	1 00	20			1.6244	
	2 00	8				1.8838

Means for groups in homogeneous subsets are displayed
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 4 47 3E-02

a Uses Harmonic Mean Sample Size = 13 115

b The group sizes are unequal The harmonic mean of the group sizes is used. Type I error levels are not guaranteed

c Alpha = 05

Figure 5-12: *4 Subsets defined by Post-hoc Comparisons*

5.3.2 Looking Ahead Again

The NSBR normalizes length according to the duration of the first constituent, thus keeping the length of the first constituent at the value 1 across conditions. Of course, as we have already seen, there is variation in the length of the first constituent.

The NSBR is only complete if we can find an insightful way of explaining this variation as well. I propose that a speaker aims to adjust the duration of the first constituent such that that the upcoming boundaries can be scaled relative to the first boundary within the limited range of possibilities that are available due to articulatory and perceptual limits. So if a boundary has to be stronger than upcoming structure, it makes sense to start out strong, but if stronger boundaries are coming than it makes start out weak.

A successful look-ahead strategy would then be one that counts the number of upcoming stronger and weaker boundaries, and adjusts the strength of the first boundaries accordingly. The strength of the first boundary is reflected in the length of the first constituent.

In particular, one possibility is that the look-ahead measure used by a speaker who does 'perfect look-ahead' is exactly the sum of all upcoming boundary ranks.

This is only one of many possible measures of look-ahead though. While the boundary ranks we tested in the previous section were assigned by the theory developed in chapter 2, the theory gives us no expectation whatsoever with respect to what measures look-ahead best. The purpose of this section is then to show that there are ways to account for look-ahead within the NSBR—but the discussion remains speculative.

Table 5.11 summarizes the sums of upcoming boundaries after the first one for each of the 11 conditions. The inverse of this value is the look-ahead value (LAV) for a particular structure.

If the LAV is positive, then the first constituent will be longer, if it is negative, then it will be shorter compared to the cases where the LAV is zero.

The theory predicts 6 different LAVs. Plotting the mean duration of the first

		A	B	C	D	Sum	LAV
1	A or B or C or D	0	0	0		0	0
2	A or (B and C) or D	0	-1	0		-1	1
3	A or B or (C and D)	0	0	-1		-1	1
4	(A and B) or C or D	0	1	1		2	-2
5	(A and B) or (C or D)	0	1	0		1	-1
6	(A and B and C) or D	0	0	1		1	-1
7	A or (B and C and D)	0	-1	-1		-2	2
8	((A or B) and C) or D)	0	1	2		3	-3
9	(A and (B or C)) or D	0	-1	1		0	0
10	A or ((B or C) and D)	0	-2	-1		-3	3
11	A or (B and (C or D))	0	-1	-2		-3	3

Table 5.11: *Look-Ahead Values*

constituent against this variable (figure 5-13) illustrates that the correlation is fairly good—except for the extreme values, where the correlation breaks down.

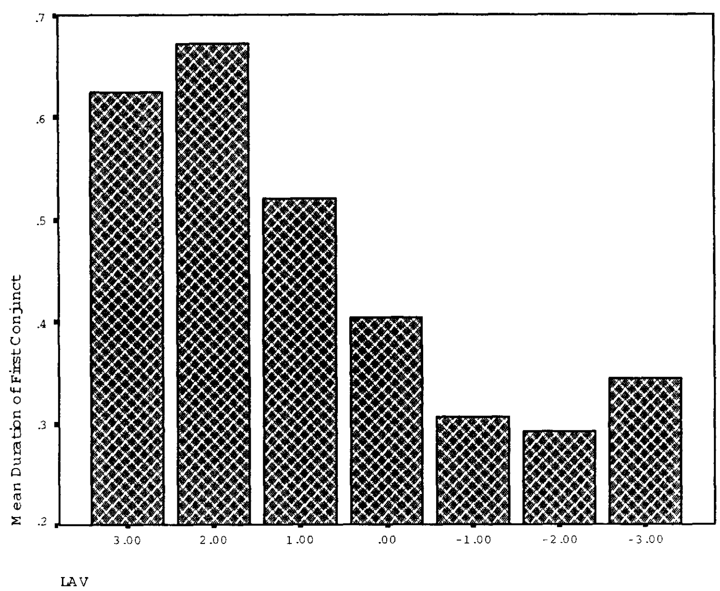


Figure 5-13: *LAV*

We can improve the correlation by the following stipulation: look-ahead only consider the structure of one cycle ahead. This is a reasonable stipulation if each cycle constitutes a processing unit, and look-ahead is constrained by how many processing units it can span. The predictions then change as outlined in table 5.12.

The means of the duration of the first constituent are again plotted against the expected look-ahead-value in figure 5-14.

		A	B	C	D	Sum	LAV
1	A or B or C or D	0	0	0		0	0
2	A or (B and C) or D	0	-1	0		-1	1
3	A or B or (C and D)	0	0	-1		-1	1
4	(A and B) or C or D	0	1	1		2	-2
5	(A and B) or (C or D)	0	1	0		1	-1
6	(A and B and C) or D	0	0	1		1	-1
7	A or (B and C and D)	0	-1	-1		-2	2
8	((A or B) and C) or D)	0	1	2		1	-1
9	(A and (B or C)) or D	0	-1	1		1	-1
10	A or ((B or C) and D)	0	-2	-1		-1	1
11	A or (B and (C or D))	0	-1	-2		-1	1

Table 5.12: *Sum of Boundaries Up and Including Next Cycle*

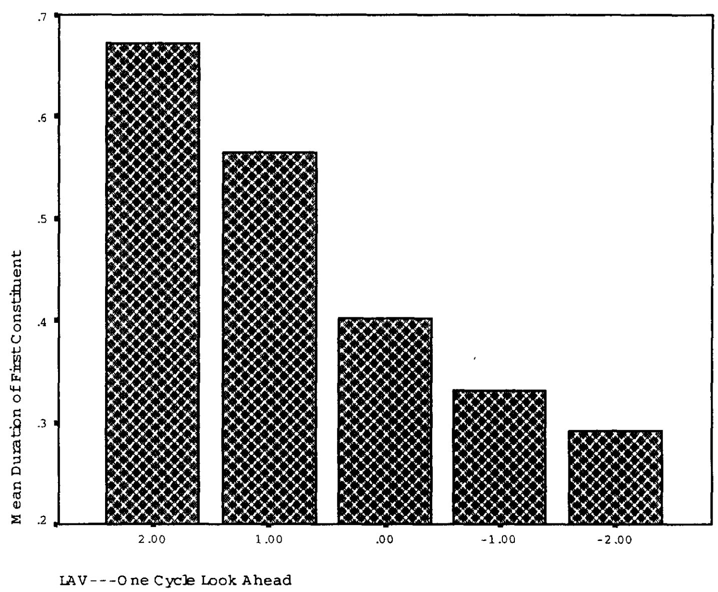


Figure 5-14: *LAV, 2 Cycle Look-Ahead*

The revised LAV theory then predicts 5 different boundary ranks. We can test this by running post-hoc comparisons on the mean duration of the first constituent in those 5 conditions. A Tukey-Post-B subsets test differentiates 4 categories in look-ahead effects, as shown in figure 5-15.

The NSBR also accounts for the apparent effects of the length of the following constituent. This effect was predicted by the LRB, but not by the original version of the SBR. The NSBR in conjunction with the look-ahead value makes the predictions summarized in table 5.3.2.

Mean Duration of First Conjunct

	LAVCYC	N	Subset 1	2	3	4
Tukey HSD[a,b,c]	-2.00	8	.2926			
	-1.00	28	.3323			
	.00	16	.4032			
	1.00	28		.5656		
	2.00	8		.6716		
	Sig.		.065	.084		
Tukey B[a,b,c]	-2.00	8	.2926			
	-1.00	28	.3323	.3323		
	.00	16		.4032		
	1.00	28			.5656	
	2.00	8				.6716

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.094E-02.
a. Uses Harmonic Mean Sample Size = 13.023.
b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.
c. Alpha = .05.

Figure 5-15: *Post-Hoc Comparisons on LAV Effect*

		LRB	SBR	Length	LAV
1	A or B or (C and D)	1	2	1	1
2	A or (B and C) and D	2	2	2	1
3	A and (B or C or D)	3	2	3	2

Table 5.13: *Length of the Upcoming Constituent*

The omnibus ANOVA for the variable LAV is highly significant ($F(4,1) = 34.94$; $p < 0.011$.

Long-Distance Effects Revisited

The NSBR combined with look-ahead values (LAV) accounts for a long-distance effect that remained unexplained in both in the LRB and the SBR. Table 5.14 summarizes the predictions of the three theories.

			LRB	SBR	LAV	Cond
Baseline	1	A ◊ and **B** and C and D	1	1	0	0
	2	A ◊ or **B** or C or D	1	1	0	0
Simplex	3	(A ◊ and **B**) or C or D	1	1	-2	-1
	4	(A ◊ and **B** and C) or D	1	1	-1	-1
Complex	5	A ◊ or **(B and C)** or D	2	2	1	1
	6	A ◊ or **(B and C and D)**	3	2	2	1

Table 5.14: *Simplex vs. Complex Constituent Following the Boundary*

The look-ahead value defined by the NSBR serves to make distinction between the three conditions. We can add the LAV-values within each condition and end up with a three-way differentiation. This accounts for the data observed in figure 5-2.

In fact, however, the NSBR makes a 5-way distinction between the 6 structures. This is exactly reflected in the distribution of the means (figure 5-16).

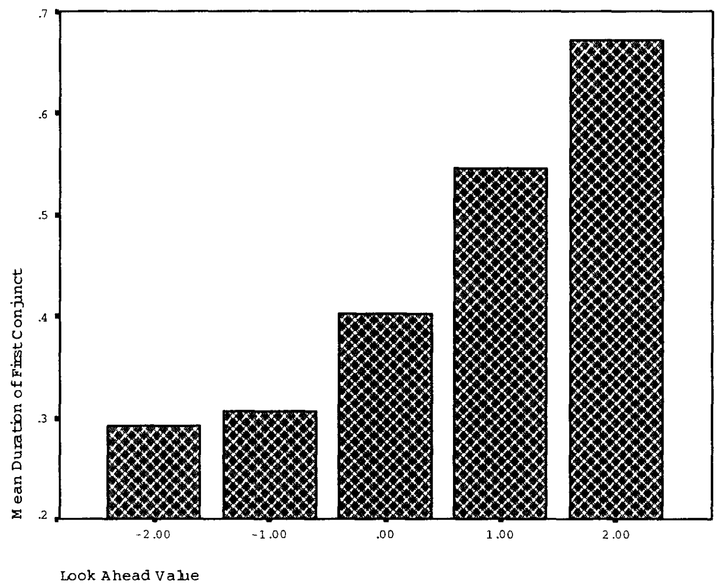


Figure 5-16: *5-Way Distinction Predicted by the LAV*

The predictions of the LAV are thus quite accurate even in a case where both LRB and SBR failed to account for the data.

The variation in length in the first conjunct is a look-ahead effect. The look ahead value (LAV) defined based on the sum of the upcoming boundary ranks, using the ranks defined by the NSBR, provides an adequate measure of the look-ahead actually observed. The theory accounts for a number of effects that go unaccounted both in the LRB- and the SBR-theory.

Consider now the predictions for the example that was problematic for the LRB, repeated below:

(9) a. A $\diamondsuit$ or B $||$ or (C $|$ and D)

b. (A $|$ and B) $||$ or C $\diamondsuit$ or D

The crucial boundary in example (a) is the one following the first conjunct. Since every constituent is normalized with respect to the first conjunct, no variation in length of that constituent is predicted based on the ranks assigned by the NSBR. The rank at ♢ is always zero. But based on the discussion in the last section, a look-ahead effect is predicted, depending on the sum of the upcoming boundaries, the LAV-value of the structures. The predictions for the variation in length before ♢ are summarized in table 5.15.

		SBR	LRB	LAV	Cnd
1	A ♢ or B or C or D	1	1	0	11
2	A ♢ or B or (C and D)	2	1	1	21
3	A ♢ or (B and C) or D	2	2	1	22

Table 5.15: *Data set for Testing Long-Distance effect*

The predictions of the NSBR with look-ahead are the *same* as those of the SBR: the duration of the first constituent should be higher for the conditions 21 and 22 compared to condition 11. This is borne out by the data, as discussed in section 5.2.4.

The case of (9b) is more complicated. Due to the new normalization scheme, a direct comparison between the three theories is not possible. Consider the predictions of the three theories (LRB, SBR, NSBR) in table 5.16.

		SBR	LRB	NSBR	Cnd
1	A or B or C ♢ or D	1	1	0	11
2	(A and B) or C ♢ or D	2	1	1	21
3	A or (B and C) ♢ or D	2	2	0	22

Table 5.16: *Data set for Testing Long-Distance effect*

In section 5.2.4, we saw that the SBR fares better than the LRB in predicting this pattern. The NSBR makes very different predictions. Condition 11 and 22 are predicted to be identical—thus differing in precisely the categories where both the LRB and the SBR made the same predictions. Since the data are normalized in a different way, we cannot directly compare the three theories.

We can test whether the predictions of the NSBR are borne out given the new normalization. Figure 5-17 illustrates that the means look as predicted.

Figure 5-18 illustrates that post-hoc comparisons confirm the predictions of the NSBR.

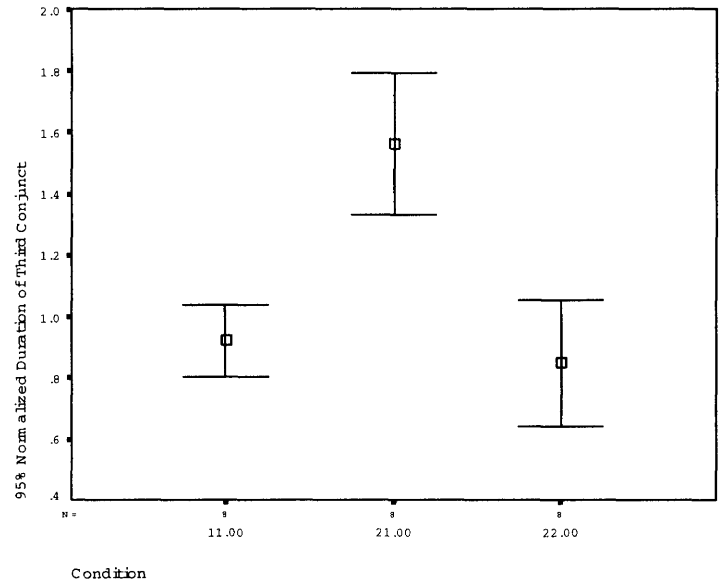


Figure 5-17: *Long Distance Effect in Normalized Data*

The results show that the NSBR makes correct predictions for the normalized length of the relevant third conjunct.

But what about the *actual* length of the third conjunct? The means of the third conjunct are summarized in figure 5-19—the third conjunct in conditions 2 and 3 have about equal length, while 1 is much shorter.

The reason the actual length does not reflect directly the relative rank of the NSBR is that the three structures in table 5.15 are associated with different look-ahead values. Only the look-ahead value of the third condition is positive, the others are zero. The different look-ahead value affect the length of the first conjunct, which in turn has an effect on the normalized length of the third conjunct. The actual length of the third condition is thus higher compared to the other two conditions than the graph in (5-17) suggests.

Normalized Duration of Third Constituent

	Condition	N	Subset 1	Subset 2
Tukey HSD[a,b]	22.00	8	.8491	
	11.00	8	.9223	
	21.00	8		1.5625
	Sig.		.655	1.000
Tukey B[a,b]	22.00	8	.8491	
	11.00	8	.9223	
	21.00	8		1.5625

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.694E-02.

a. Uses Harmonic Mean Sample Size = 8.000.

b. Alpha = .05.

Figure 5-18: *Post-Hoc Comparisons on Look-Ahead Effect*

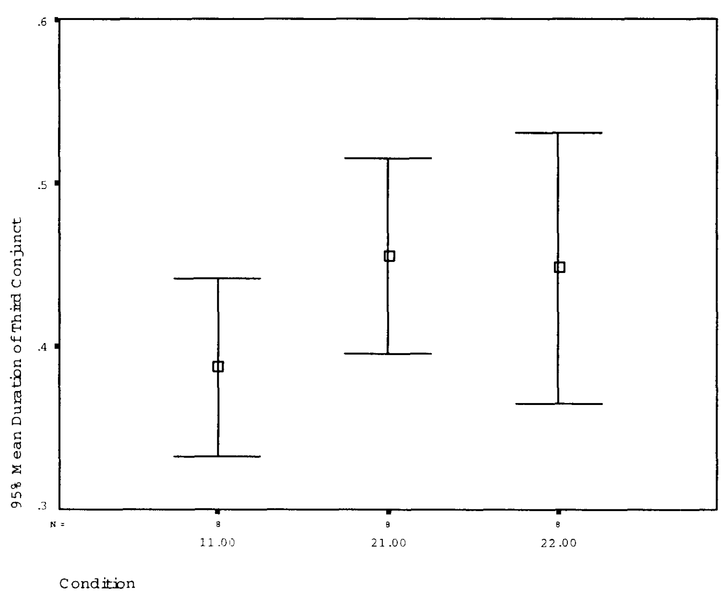


Figure 5-19: *Long Distance Effects*

5.4 The Prosody of Function Words

The comparisons so far only looked at length effects on the proper names in the structures. The reasoning was that the pre-boundary lengthening between conjuncts is a measure of the boundary strength. Another question of interest, however, is whether there are any length effects on the *coordinators.*

Coordinators are usually pro-cliticized[5], just like many other function words. Gee and Grosjean (1983, 433) simply assign a boundary of rank 0 to the boundary following them. This would predict however that there is no variation in the length of function words depending on where in the structure they occur. In the following example, e.g., 'or' and 'and' should have the same prosodic status:

(16) A or B or (C and D)

But within the relative theory of metrical structure in chapter 2, pro-cliticizing a function word may simply mean the following: the boundary separating it from the following conjunct is weaker than the boundary separating it from the preceding conjunct.

(17) Function Word Pro-Clititcization
If the boundary preceding a functor is of rank 1, then the boundary between the functor and its complement is n-1:

```
|        ×   |
|  ×  |  ×   |
   F    Host
```

In chapter 6, I will discuss how this 'prosodic subordination' of function words comes about. The prosodic structure assigned to (16) looks then as follows:

(18) Prosodic Structure with Functors:

```
| × |      ×  |      ×              ×  |
| × |  × | ×  |  × | ×              ×  |
| × |  × | ×  |  × | ×  |           ×  |
| × |  × | ×  |  × | ×  |   ×   |   ×  |
  A   or   B    or   C     and       D
```

[5]Although they may rebracket to the left, e.g. with 'n

The variance in the length of the connectors can then be conceived of as pre-boundary lengthening. The prediction is that the length of the connectors should co-vary with the length of the preceding constituent: If the preceding boundary is stronger, it means that the boundary separating the connector form the following conjunct is stronger. Figure 5-20 summarizes the means.

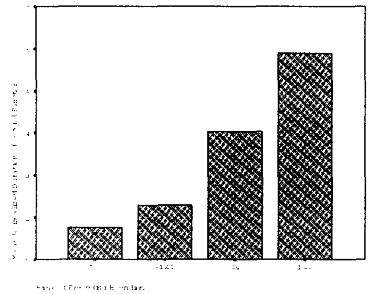


Figure 5-20: *Rank of Second Boundary, Second Functor*

Post-hoc Comparisons differentiate three different classes of functors, as illustrated in figure 5-21.

Mean Normalized Duration of Second Connector

	Bnd Rank	N	Subset 1	Subset 2	Subset 3
Tukey HSD[a,b,c]	-2.00	4	.1773		
	-1.00	28	.2302	.2302	
	.00	32		.4040	.4040
	1.00	24			.5890
	Sig.		.899	.112	.081
Tukey B[a,b,c]	-2.00	4	.1773		
	-1.00	28	.2302	.2302	
	.00	32		.4040	
	1.00	24			.5890

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 3.235E-02.

a. Uses Harmonic Mean Sample Size = 11.154.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

Figure 5-21: *Statistically Different Boundary Ranks*

Similarly accurate predictions are made for the last connector. Figure 5-22 summarizes the means plotted against the rank of the preceding boundary.

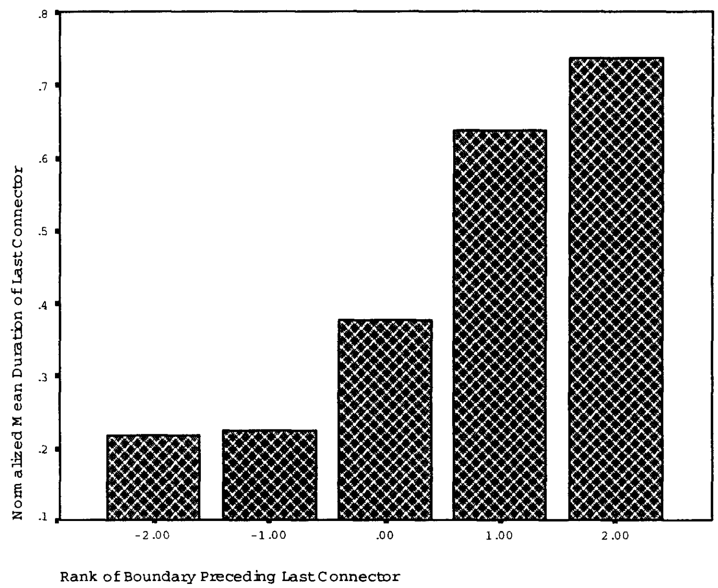


Figure 5-22: *Rank of Third Boundary, Last Functor*

Post-Hoc Comparisons differentiate 4 categories of lengthening for the last connector, as illustrated in figure 5-23.

A more extensive discussion of function words and how they receive their prosody within the theory outlined in chapter 2 will be discussed in chapter 6.

Normalized Mean Duration of Last Connector

	DBND2	N	Subset 1	2	3	4
Tukey HSD[a,b,c]	-2.00	8	.4955			
	-1.00	20	.5266			
	.00	32		.9317		
	1.00	20			1.6244	
	2.00	8				1.8838
	Sig.		.996	1.000	1.000	1.000
Tukey B[a,b,c]	-2.00	8	.4955			
	-1.00	20	.5266			
	.00	32		.9317		
	1.00	20			1.6244	
	2.00	8				1.8838

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 4.473E-02.

a. Uses Harmonic Mean Sample Size = 13.115.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

Figure 5-23: *Statistically Different Boundary Ranks*

Summary

This chapter showed evidence that the boundary ranks assigned according to the SBR are reflected in the prosody of coordinate structure. A recoding of the SBR in terms of a left-to-right implementation (the NSBR) led to a fairly accurate model of the relative strength of boundaries as it is reflected in duration.

The theory proposed here uses a relative notion of boundary rank instead of a categorical one as in the theory of prosodic phonology. The left-to-right implementation of the SBR, the NSBR, provides a model of prosody that is compatible with incremental production. Boundaries are scaled in strength relative to earlier boundaries.

A certain amount of look-ahead adjusts the strengths of the boundary after the first constituent in such a way that the subsequent boundaries can be realized in a reasonable range. Look-ahead requires anticipating the complexity of the upcoming structure. The precise measure of complexity used here was the sum of the upcoming boundary ranks—but whether or not this is the right measure is a question for further research.

The look-ahead in this experiment was fairly extensive, possibly due to the fact that the task was such that it facilitated anticipating the entire structure of the answer. A different kind of task might show less look-ahead and more look-back scaling relative to earlier material.

Part II

Prosodic Subordination

Chapter 6

Recursion and Prosodic Subordination

Abstract

There is an asymmetry in the assignment of prosody depending on linear order. The asymmetry is the following: when a functor precedes its complement, it can bear an accent, but when it follows its complement, it must be 'prosodically subordinated'. This asymmetry is at the heart of the distribution of nuclear stress, and the phonological difference between the pre- and post-nuclear domain. A principle of PROSODIC SUBORDINATION is proposed that accounts for the asymmetry and derives the correct prosodic grouping and the distribution of accents, including the location of nuclear stress. It can be observed across different domains of the grammar, from the prosodic relation between affixes and stems to the relation between matrix clauses and embedded clauses, and it applies to the relation between heads and complements and to the relation between modifiers and modifiees.

6.1 Nuclear Stress and Prosodic Asymmetry

The concept of 'nuclear stress' has played a crucial role in the development of theories of the syntax-phonology interface. This section discusses various problems with the original notion of nuclear stress and presents a systematic prosodic asymmetry. This asymmetry will be accounted for by the recursive application of PROSODIC SUBORDINATION. It will be shown to account for the prosodic grouping and nuclear stress in various syntactic domains of English, Dutch and German.

6.1.1 SPE

In SPE, for every phrase marker the nuclear stress rule (NSR) identifies exactly on vowel as most prominent. The vowel that native speakers perceive as most prominent in a given structure counts as its nuclear stress. The representation of stress in SPE is that of a phonological feature that can have any positive integer number as its value. The nuclear stress of an utterance is the unique vowel that has the value 1. Within every sub-constituent, there is exactly one vowel that has the lowest number. We can call this the vowel carrying nuclear stress within that sub-constituent.

A different notion of 'nuclear stress' or 'nuclear accent' is used in the intonation literature. Newman (1946, 174) singles out a heavy stresses as 'nuclear' if it 'acts as the nucleus of an intonational unit'. Within any intonational unit it is always the last heavy stress that counts as the nuclear stress. For the course of this section, I will stick to the SPE notion of nuclear stress, and I will return to Newman's idea about nuclear stress in the next section.

In SPE, just as in an earlier version of the theory presented in Chomsky et al. (1957), the stress value of a particular vowel is determined by recursively going through the structure and applying phonological rules. The three stress rules proposed for English in SPE are the Main Stress Rule (MSR), the Compound Rule (CR), or the Nuclear Stress Rule (NSR). The MSR assigns stress to mono-morphemic words and morphologically built complexes; the CR negotiates stress within compound structures; and the NSR finally negotiates stress in syntactic phrases.

These rules are recursively applied to surface structures, that is to the syntactic tree that is the output of a derivation. This basic idea of a recursive algorithm to assign stress was called the 'transformational cycle', and can be summarized as follows (SPE, 24–26):

> "Utilizing the principle of the transformational cycle, the speaker of English can determine the phonetic shape of an utterance on the basis of such rules as the Compound and Nuclear Stress Rules, even though the particular utterances may be quite new to him. [...] Once the speaker has selected a sentence with a particular syntactic structure and certain lexical items (largely or completely unmarked for stress, as we shall see), the choice of stress contour is not a matter subject to further independent decision. [...] With marginal exceptions, the choice [...] is as completely determined as, for example, the degree of aspiration".

Each of the stress rules promote one vowel over other vowels in prominence. This is achieved by 'subordinating' all other vowels in the structure by augmenting the stress number by 1. Which rule applies at a certain syntactic node depends on whether it is a word, a compound, or a phrase. In phrases, the NSR applies and demotes all stress numbers on each vowel except for the last vowel that carries a '1'.

The ultimate stress values of a phrase marker in English are determined by recursively applying the three stress rules to the tree structure of a phrase marker, starting at the terminal nodes:

(1) The Nuclear Stress Rule in SPE

$[_{NP}$ $[_{DP}$ Helen]	$[_{VP}$ $[_{V'}$ reads$_V$	$[_{NP}$	$[_{N'}$ strange$_A$	$[_N$ books$_N$]]]]	
1	1		1	1	*MSR*
			2	1	*NSR*
	2		3	1	*NSR*
2	3		4	1	*NSR*

This chapter is concerned with the type of data for which SPE would have considered the NSR to the relevant phonological rule, so I will not discuss the formulation of the MSR and the CSR. While the representation of prosody that is assumed will be used in this chapter is very different from the stress numbers in SPE, the notion of nuclear stress is still going to play a role, and the solution will again make use of the transformational cycle. The crucial difference will lie in the formulation of the principle that derives nuclear stress.

As for the NSR in SPE, it is sufficient for the present discussion to note that it has the effect that on any branching phrase, it is the last element that receives main prominence. In other words, nuclear stress is predicted to be always on the *rightmost element within phrases.* This prediction, however, turns out to be false.

6.1.2 Linear Order *vs.* Argument Structure

Newman (1946, 179) notes that in the following type of example, main prominence falls on the object, and not the final predicate[1]:

(2)
a. fávor to ask.
b. bréad to eat.
c. instrúctions to leave.

Bresnan (1971) discusses this and related data to show that the theory of nuclear stress assignment in SPE is in need of revision. Bierwisch (1968, 176) points out

[1] I only note nuclear stress for simplicity.

similar cases in German, and concludes that the rules of stress placement must make reference syntactic and semantic information:

(3) a. Peter betrachtet ein Búch.
Peter looks.at a book
'Peter is looking at a book.'

b. Peter hat ein Búch betrachtet.
Peter has a book looked.at
'Peter looked at a book.'

According to Newman, the crucial factor involved in what he calls 'middle stress', that is non-final nuclear stress, is argument structure: the 'logical object' precedes the predicate. Hence nuclear stress is final in the following examples, in which the 'logical object' follows the predicate:

(4) a. a desire to éat

b. the will to líve

c. instructions to léave (...= was instructed to leave)

That predicates, and heads more generally, are less prominent than their complements holds across a number of languages, as illustrated in Donegan and Stampe (1983), Jacobs (1992) and Cinque (1993). In Germanic languages this is particularly easy to observe since they are languages with mixed headedness. In English, functors predominantly take their arguments to the right:

(5) She painted some wálls.

But sometimes, the complement of a predicate can precede it. In English, in certain environments a predicate can follow its nominal complement. We can call these types of environments 'OV' environments.

(6) She had some wálls painted.

Whether or not a predicate in fact subordinates depends on the meaning of the predicate, the meaning of the object, and the interaction between the two. For example, while predicates in German OV environments tend to subordinate, sometimes they do not. This was observed in Höhle (1982) and Jacobs (1988, et. seq.). For example, in the following sentence it is equally possible to stress the predicate and the object:

(7) a. Der Árzt wollte einen Pátienten untersúchen.
the doctor wanted the patient examine

b. Der Árzt wollte einen Pátienten untersuchen.
the doctor wanted the patient examine

There does seem to be a subtle difference in meaning, but neither sentence requires narrow focus on the object or the predicate. The same is true for corresponding examples in English:

(8) a. The dóctor wanted to have a pátient exámined.

b. The dóctor wanted to have a pátient examined.

The difference has been related in the literature (e.g. in Krifka (1984) to the distinction between 'thetic' and 'categorical' statements—which is still very ill understood. I will return to this problem in section 6.3 and explore some of the differences between the two types of cases. For many cases, however, the predicate is subordinated and main stress falls on the argument.

The NSR in SPE predicts nuclear stress to be final in phrases, and thus cannot explain why nuclear stress falls on the object irrespective of linear order. Bresnan (1971) suggested to maintain the NSR proposed in SPE, and to derive the effect by changing the assumptions about how the phonological cycle interacts with the syntactic cycle. I will return to her proposal in section (6.3.3).

More recent approaches to sentence stress and prosody have incorporated Newman's idea and make reference to argument structure in deriving prosody. The generalization these approaches try to capture can be stated as follows (adapting the formulation in Schmerling (1976, 82)):

(9) Arguments are more prominent that functors.

There are different approaches to explain this observation. The first kind of approach involves the notion of 'focus projection'. 'Focus projection' is what Höhle (1982) termed the phenomenon that accents on certain constituents are compatible with focus on constituents that include them, sometimes up to and including the entire sentence. The NSR tries to capture the prosody of sentences with 'neutral' focus (a discussion of how this can be defined is given in chapter 7). Selkirk (1984) specifies explicit rules about which constituents can be in focus given a certain distribution of pitch accents. One of the rules of focus projection is that heads can inherit F-marking from arguments. This correctly captures that an accent on the object is sufficient to license 'focus' on the constituent containing both the head and the object, and allows the predicate thus to remain unaccented and hence lower in prominence. The approach in terms of 'focus projection' receives a detailed discussion in chapter 8.

The second type of approach attempts to derive the generalization from conditions on prosodic domain formation. Gussenhoven (1983) proposes a rule of accent formation, which integrates predicates into the accentual domain of adjacent arguments. Gussenhoven's proposal is reminiscent of the approach taken in Truckenbrodt (1995), who introduces a constraint Wrap-XP, which forces XPs (but not heads) to be contained in a prosodic constituent of their own.[2]

The third type of approach explicitly postulates a principle that assures that predicates are less prominent than their arguments, as e.g. in Schwarzschild (1999). In Jacobs (1991a, 1992, 1999), Arregi (2002), Wagner (2002a) a version of this principle is incorporated into a theory that assumes the transformational cycle and assigns a complete metrical representation to phrase structures.[3]

A commonality of all of these approaches is that they treat the head-final case

[2]Approaches based on left- or right- XP alignment (Selkirk, 1986) and the approach based on ϕ-formation and restructuring in Nespor and Vogel (1986)—while not incompatible—-offer no explanation for the generalization about complement prominence when a language has both head-initial and head-final constructions, and I will not discuss these approaches at this point. They do, however, correctly predict an asymmetry depending on the linear position of a head relative to its complement, to which I will turn shortly.

[3]Cinque (1993) proposes a different explanation for this generalization based on 'depth of embedding', and does not make direct reference to argument structure or the head-complement relation. This approach is discussed and criticized in Truckenbrodt (1993) and Arregi (2002).

and the head-initial case as mirror images of each other. The principles that assure that predicates are less prominent apply independent of linear order (e.g. Schmerling (1976, 82)). These symmetric approaches do not account for a systematic asymmetry that relates prosody and linear order.

6.1.3 Prosodic Asymmetry

The previous section illustrated that predicates are less prominent than their complements, whether they precede it or follow it. There is a crucial difference though between head-initial and head-final constructions: when the functor precedes its complement, it can optionally bear an accent, but if it follows its complement, it must be unaccented:[4]

(10) Did she change anything in her new apartment?

a. She páinted the wálls.

b. She had the wálls painted.

Newman (1946) observed that it is the last accent that is perceived as the more prominent accent. If this is correct, then placing an accent on the predicate shifts the prominence in the case when it *follows* the argument but not when it precedes it. The asymmetry observed here can thus be related to the more general observation that predicates are less prominent than their arguments in Schmerling (1976).

There are of course circumstances when the predicate can or even must receive an accent in even in the 'OV' order (11a).

(11) What did she do with those ugly pink walls?
She had the walls páinted.

[4]The context is such that the presupposition of uniqueness in the answer is fulfilled although the walls have not been referred to yet. 'The walls' is not contextually given. They are the walls pertaining to the apartment. The interaction between the presupposition of givenness marking and the presuppositions of definite noun phrases is discussed in chapter 7.

But this prosody requires the complement to be 'given' in the context. To be precise, the prosody in (11a) introduces the presupposition that that another proposition about 'the walls' is salient. This is the same presupposition that (12) introduces:

(12) What did she do with those ugly pink walls?
She páinted the walls.

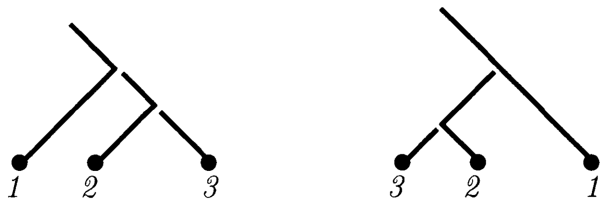
Whether or not an accent is placed on the predicate in the VO-order does not affect the information structure of the sentence. Gussenhoven (1983) reports experimental evidence that the presence/absence of placing an accent on the predicate has no influence on whether or not listeners are more likely to consider a sentence part of a dialogue in which the predicate was given or not.[5]

The presuppositions of givenness marking will be discussed in chapter 7. For the remainder of this chapter, I will only discuss sentences that have 'neutral' prosody, i.e. a prosody that does not impose a presupposition of 'givenness' on the context. The notion of 'neutral' prosody is difficult to define. It is made explicit in chapter 7.

The observation important here is that there is an asymmetry between VO and OV. This prosodic asymmetry can be further illustrated by looking at structures that consists of three elements that form a selectional sequence, such that the highest constituent Head_1 selects a complement that consists of a a head, Head_2, and its complement. In the following, I will draw syntactic trees such that the projecting side of a branching node is not interrupted. There are several different linearizations for these structures. Consider the following two cases:

(13) a. Right-Branching a. Left-Branching

[5]The experiment used a 'context retrieval task'. Listeners had to choose between two dialogues, one in which the answer had wide focus, the other in which it had narrow focus (a narrow focus question makes the predicate given. Some researchers have claimed that accent on the predicate is infelicitous with narrow focus on the object (Kadmon, 2003)). But an experiment reported in Gussenhoven (1983) found that while there is at least a tendency to use unaccented predicates in contexts with narrow focus on the object, the correlation was not clear cut. Furthermore, listeners were not able to use the accentedness of the predicate to retrieve the context in which the answer was uttered.



Instances of both linear orders are attested in German predicate sequences. Consider the following two examples:

(14) '...weil er ihr...
'...because he...

a. [versprách] [zu versúchen] [zu schwéigen]. $\acute{1}\prec\acute{2}\prec\acute{3}$

b. [zu schwéigen versprach zu versuchen.] $\acute{3}\prec 2\prec 1$

be silent try promise

'...promised her to try to be silent.'

The observation is that predicates that precede their complements receive a prosodic domain of their own, but predicates that follow their complement are prosodically subordinated. The metrical representation assigned to those structures by Jacobs (1991b) and Cinque (1993) is exactly symmetrical:

(15) Symmetric Representations

a. $1\prec 2\prec 3$

$$\begin{array}{l} \qquad\quad \times \\ (\,.\quad .\quad \times\ ) \\ (\,.\,(\,.\quad \times\ )) \\ (\times(\times(\times))) \end{array}$$

b. $3\prec 2\prec 1$

$$\begin{array}{l} \quad \times \\ (\ \ \times\ \ .\ \ .\,) \\ ((\ \times)\,.\ \ .\,) \\ (((\times)\times)\times) \end{array}$$

Cinque (1993, 245)

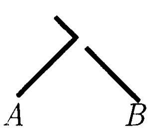
```
×××  ×  ×            ×  ×××
×    ×               ×   ×
     ×               ×
    (×)             (×)
```

Jacobs (1992)

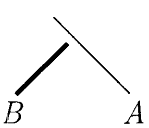
These symmetric approaches miss out on an important phonological generalization, namely that the two cases show a systematic prosodic asymmetry. The generalization about prosody and linear order can be summarized as follows.

Iin the following, I will draw trees with bold lines on the branches whose corresponding phonological representation do not undergo subordination; also, I am connecting the projecting side to the root and leave the non-projecting side unconnected.

(16) Prosodic Asymmetry

a. 

When a functor A *precedes* its complement B, the functor may be on a par with its argument or may be prosodically subordinated: [Á B́]; [A B́].

b. 

When a functor A *follows* the complement B , A is prosodically subordinated: B́ A (unless A is focused or B is given).

In the next section, I will discuss how the effects of prosodic subordination can be accommodated with the cyclic approach to syntax-to-prosody mapping presented in chapter 7. In section (6.4) I will present more evidence for the generalization from predicate sequences in English, Dutch, and German. The generalization will be shown to be robust under word order variation across dialects of West-Germanic. Section (6.5) extends the generalization to a much wider variety of cases.

6.2 Prosodic Subordination

In theory of syntax-to-projection presented in chapter 2 prosodic subordination remains unaccounted. This section presents an addition to the theory that allows prosodic subordination to recursively apply.

6.2.1 Concatenate and Project

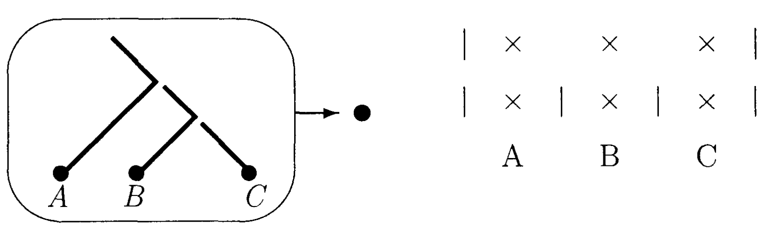
Consider again the simple case of sequences of predicates in English. They appear to form a right branching structure, and they have a 'flat' prosody, in that each predicate receives an accent and is set of from adjacent predicates by boundaries of equal strength. They are prosodically on a par. Compare the prosody in (17a) with (17b), which is very similar in the prosodic phrasing:

(17) a. They want | to try | to begin | to plan | to move.

b. to want, | to try, | to begin, | to plan, | and to move.

In chapter 3 it was argued that both types of structures are associative domains, and are therefore assembled in a single cycle. Flat prosodies are generally associated with right-branching structures that are assembled in a single cycle, and are mapped to a single foot in the top metrical grid line. This is illustrated in the following diagram:

(18) A or B or C: Derived in a Single Cycle



A single mapping principle was proposed in chapter 2 to derive prosodic structure: each cycle is mapped to a new top line grid line in the metrical representation upon completion. The mapping principle is the following:

(19) PROSODIC MATCHING

a. Concatenate

Concatenate the prosodic representation of the elements in the domain aligning their top lines and filling the columns where necessary.

b. Project

Create a new top-line grid line n by projecting all grid marks on line n-1, and mapping them into a single foot on line n.

This projection principle has the effect that each element in a cycle projects to the top line. That was the right result for coordination structures, since independent of the bracketing and the number of cycles, each conjunct always receives at least one accent.

Nuclear stress is not encoded here by projecting the element bearing nuclear stress to the top grid line, as is usually done (e.g. Halle and Vergnaud, 1987, Cinque, 1993, Arregi, 2002). Still, nuclear stress can be read off the structure. The generalization proposed in chapter 2 is the following:

(20) Nuclear Stress Generalization

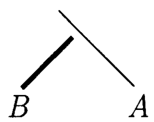
Within each foot, nuclear stress is perceived on the last of those grid marks that project highest.

The idea that the most prominent is the last among equal goes back to Newman (1946), and has been adopted by various authors since (e.g. Schmerling (1976, 86) and Truckenbrodt (1995, 113)).

This means that nuclear stress is not a primitive of the theory but is a derivative notion. In order to determine the nuclear stress of a structure, once needs to investigate the conditions of accent omission and prosodic subordination.

While nuclear stress is always on the last element when all that happens is PROSODIC MATCHING, some times nuclear is not final, as was discussed above. For example, in the case of the OV-word-order, something else happens: only some but not all material projects to the top line. I assume that the correct prosodic representation for prosodic subordination to be the following:

(21) Prosodic Subordination



| × |
| × | × |
B A

Constituent A is 'affixed' to B, i.e. to the prosodic domain it is subordinated to. Since the A does not project to the top grid line, it does not receive an accent. All and only top line grid marks are associated with accents, as outlined in chapter 2. This has the effect that nuclear stress is not perceived on the final element A in the structure, but on B. *Nuclear stress is perceived on the last element that has not been subordinated.* But under what conditions does subordination apply?

6.2.2 Cyclic Subordination

The assumption in chapter 2 is that syntax is only mapped to prosody at the end of a cycle. The usual way of relating the prosody of two cycles is to concatenate a number of elements at the top line and projecting each elements. I will assume that subordinating one constituent to another constituent is a relation between two sister constituents, where one sister is subordinated and 'affixed' to the other. I will treat subordination then as a case in which two sister nodes undergo a cycle and the mapping principle that applies is PROSODIC SUBORDINATION. The following principle has the effect of subordinated one sister to the other.

(22) PROSODIC SUBORDINATION
A is subordinated to B by concatenating A and B such that if A projects to line i, B projects to line i+1.

The question is now under what circumstances PROSODIC SUBORDINATION applies. PROSODIC SUBORDINATION is the only principle apart from PROSODIC PROJECTION in the the system. All factors, be they related to argument structure or related to information structure, factor into the derivation prosody only in deciding when either of the principles applies.

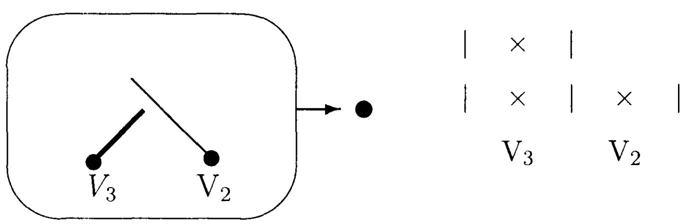
One factor that can trigger PROSODIC SUBORDINATION seems to be argument structure. The discussion in the previous section suggested that a functor follows its complement is obligatorily prosodically subordinated. We related this generalization to prominence: if the functor would not be subordinated, then it would project to the same grid line and become more prominent than its complement by virtue of the nuclear stress principle. The following principle is the driving force behind obligatory subordination:

(23) COMPLEMENT PROMINENCE

If A and B are sisters and A is the functor and B its argument, then B is more prominent than A.

This principle will force prosodic subordination whenever a functor follows its complement, due to generalization about prominence in sequences of unsubordinated constituents. If PROSODIC SUBORDINATION did not apply, then the functor would *not* be less prominent than its complement. The application of PROSODIC SUBORDINATION can be illustrated as follows:

(24) PROSODIC SUBORDINATION



The conditions under which PROSODIC SUBORDINATION applies in cases where the functor *precedes* its complement will not be discussed here. I will assume that it is optional for the time being, and indeed in many cases it is. Two factors seem to play a role, one is information structure, the other is rhythm. Some evidence for rhythmic effects will be discussed in passing in this chapter.

The following sections will apply the principle to a number of different types of cases and illustrate how the system works.

6.3 Nominal Arguments

This section discusses the prosody of complements and specifiers, and discusses issues relating to the interaction of movement and subordination.

6.3.1 Complements

Predicates preceding their complement can receive an accent in English (25a). This true both for infinitival and for DP-complements (25b).[6]

(25) a. She wánted to hélp to succéed.

b. She wánted to hélp to páint the hóuse.

These sequences of predicates and their prosody can be derived by assembling them in a single cycle and applying PROSODIC PROJECTION upon spell-out, as outlined above. Some times, even in English an object precedes its predicate and the expected asymmetry is observed, as discussed before:

(26) What did she want to change before moving in?
She wánted to have some wálls painted.

[6]The accent on 'paint' is certainly not obligatory, but it is possible. In VO-orders, the verb *can* bear an accent. The evidence for the phrasing of verb and direct object often given is the application of the rhythm rule:

(1) Evidence for Phrasing Kenesei and Vogel (1995)

a. ...in English: Rhythm Rule:
[They mánaged] [to óutclass] [Délaware's cantéen].

b. ...in German: No Accent on Verb
[Sie haben Délaware's Kantíne übertroffen].

Since here, the functor 'painted' follows its complement, 'walls', PROSODIC SUBORDINATION applies. Assuming that 'some' has already been subordinated relative to 'walls' at an earlier cycle, the following representation is derived:

(27) some wálls painted

```
          |  ×  |
|   ×   |  ×  |    ×    |

   some    walls   painted
```

There are more cases where functors are preceded by arguments. Consider intransitive predicates and subjects: subordination is often the unmarked option, both with unaccusative (a) and unergative (b) verbs.

It has been reported that unaccusatives tend to phrase with the subjects whereas unergatives don't (Selkirk, 1995b, Hoskins, 1996). This at least does not hold for *all* unergative and unaccusative predicates, as illustrated here. While it may well be that these subjects are not in the complement position of the lexical predicates involved, they are still the arguments of the the expression that follows them and they are not the projecting element. Therefore, they are the complement of the functors and the functors consequently subordinate. Subordination is even observed when there is more than one predicate (c):[7]

(28) Why is everyone in the palace so nervous?

a. [The quéen called]

b. [The quéen is expected to come.]

There are a number of cases that show that functors are not always subordinated when they follow an argument. As is well known, the subject in an intransitive sentence sometimes phrases separately from the verb, e.g. when it encodes information that

[7]The DPs here are new information definites. The presupposition of uniqueness is fullfilled by word knowledge: there is usually only one queen per palace, if there is a queen at all. This is why using an indefinite here would sound funny, since it suggests that the default assumption that there is only queen if any is not fulfilled. Definiteness is only very indirectly related to information status. See discussion in chapter 7.

counts as 'given' in the context. The following context is set up to facilitate wide focus in the embedded clause—but with a backgrounded or 'given' subject.

(29) What did you say the dean did?
I just said that [The déan] [arríved].

This prosody is only felicitous if the argument in question is contextually given. An argument counts as contextually given if there is a proposition involving the same argument (or a co-referent one) that is salient in the discourse. This kind of effect and how it can be accounted for will be discussed in chapter 7. Other examples do not constitute discourse-old information. Consider the following:

(30) Why did they close the factory?

a. [The fáctory] [went bánkrupt]

b. Gasolíne evaporated.

c. [A wórker] [evá́porated].

'The factory' is given in (30a) and thus the verb receives main stress. 'A worker' in (30c) however does not have to be 'given' but still nuclear stress falls on the verb. The indefinite is arguably treated as a member of a set inferred from the background (the workers of the factory). Another factor may be that the predicate 'evaporate' is simply very unexpected in combination with the argument 'a wórker', as opposed to 'gasoline in (b), for example. Just as in (a), the verb receives an independent accent.

Similar contrasts as in the case of unaccusative verbs can be constructed with unergative verbs. In (31b), the subject is again an indefinite that is interpreted as a member of set made salient by the discourse (i.e. a partitive relating to a set in the background).

(31) Why did they interrupt the play?

a. [A chíld was crying].

b. [An aćtor] [was crýing].

The stage-level, individual-level distinction is directly reflected in prosody (Schmerling (1976), Diesing (1992), Kratzer (1995)), and here discourse new-/oldness is not at stake. Individual-level predicates resist subordination:

(32) a. [Your éyes are red].

b. [Your éyes] [are blúe].

The issue that in argument-predicate sequence prosody is to some degree flexible and under ill-understood circumstances the predicate can also receive an accent does not only arise in the case of subjects. As observed in Höhle (1982), Jacobs (1988), Jacobs (1991b), Uhmann (1991), and Jacobs (1999), there are some object-verb sequences in which the predicates receives an accent, and it is not obvious that in those cases the object must is presented as 'given'.[8]

(33) a. Der Árzt wollte einen Pátienten untersúchen.
the doctor wanted the patient examine

b. Der Árzt wollte einen Pátienten untersuchen.
the doctor wanted the patient examine

There is clearly a subtle difference in meaning, but it is hard to definite exactly the criteria that distinguish them. The hope is that once properly understood, it will turn out that depending on the prosody the functor/argument relations are affected and that this is why the prosody is different. I will not explore the issue further in this chapter.

6.3.2 Specifiers

A notorious problem for the assignment of prosody in general (cf. discussion in Cinque, 1993) and for the present proposal in particular is the prosody of specifiers. The following case presents another case of a two-place functor.

[8]The notion often invoked in order to describe the difference between those two sentence types is the distinction between 'thetic' and'categorical' judgements (see e.g. Krifka (1984) and Lambrecht (1994, 137) for English and German). The problem with this distinction is that it has resisted a precise formalization and the difference remains hard to pin down (see Ladusaw (1994) for an overview).

(34) The wóman wore the hát.

Under the assumption that it is 'wore the hát' that takes the subject 'the wóman' as an argument (Marantz, 1984), the asymmetry in (16) would lead one to expect subordination of 'wore the hát' under 'the wóman'—a patently false prediction of the system described so far.

One possible way out is to question the functor-argument relation. The subject might be introduced by a separate head, as discussed in chapter 3, following Schein (1993), Kratzer (1996), Pylkkänen (2002). But it may still be that this head is 'affixed' to the transitive verb. Schein (1993) e.g. suggests that agents are introduced by INFL, which at least morphologically is affixed to the verb. In this case, the problem persists, since still the subject is the argument of the following functor, the VP headed by the predicate.

The problem could be resolved only (i) by positing that it is the specifier which in fact projects, following the proposal in Starke (2001), who takes specifiers to be complement-taking projectors; (ii) by stipulating that the head that introduces the subject is really a suffix on the subject; or (iii) by following (Anagnastopoulou and Alexiadou, 2001) in positing that only one argument can remain the projection of a functor and exploiting the movement to of the subject to derive a relation between the subject and its sister that does not require subordination.

The generalization is that for transitive functors that take two arguments subordination is only possible between head and complement. Jacobs (1991a, 21) proposes to stipulate this as a restriction on subordination: a predicate can subordinate (in his terminology, the object 'integrates') if no other argument has integrated into the predicate. He notes that compound stress rule seems to capture a same generalization.

Arregi (2002) uses the definition of complement and specifier in Bare Phrase Structure (Chomsky, 1994) to capture the effect. A complement is the sister of a non-branching head; a specifier is the sister of a branching head. The generalization about prominence is then that if B is the head and A and B are sisters then A is prominent unless B is branching. This is precisely the compound rule proposed in

Libermann and Prince (1977).

For Arregi (2002), specifiers do not receive nuclear stress is due to the fact that the constituent they attach to is internally branching. The solution in terms of 'branchingness' requires some elaboration.

There are many cases where branching nodes *do* subordinate. One involves cases where an argument precedes several predicates:

(35) Maria hat <u>zu schwéigen</u> *versucht zu versprechen.*
Maria has to be silent tried to promise

Also, predicates that are compounds and are thus branching can subordinate:

(36) Er wollte das Zímmer staubsaugen.
he wanted the room vacuum-clean
'He wanted to vaccuum-clean the room.'

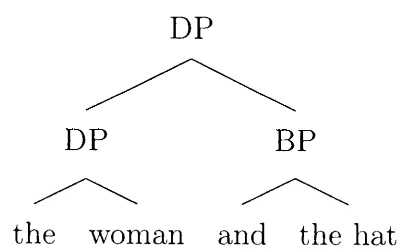
Only in cases where the non-projecting element is the second argument of a transitive functor does the principle of subordination fail to apply, but not in cases where the complement of an embedded predicate is the sister of a branching cluster of predicates. It seems that 'branchingness' has to be defined in such a way that it only counts other arguments of the heads in question in assessing branchingness. This is the generalization offered in Jacobs (1991a).

A similar problem arises in the case of coordination. Again, 'and' can be seen as a two-place function. The prosody is similar to the previous example:

(37) Coordination
The wóman and the hát.

If we do not have a special condition to prevent specifiers from subordinating, the analysis that I would have to adopt in order to assign the right prosody is the following:

(38)



This type of analysis was in fact proposed for coordination in Munn (2000). The first conjunct projects, and the second conjunct including coordinated is an adjunct to the first. But it is not clear whether this kind of analysis could be generalized to *all* specifiers.

In order to capture the facts regarding specifiers I incorporate the insights from Jacobs (1991a) and Arregi (2002) and modify the principle of COMPLEMENT PROMINENCE as follows:

(39) COMPLEMENT PROMINENCE

If A and B are sisters and A is the functor and B its argument, then B is more prominent than A unless A already contains an argument (i.e. is branching in the sense of BPR).

6.3.3 Subordination and Movement

Bresnan (1971) adopts the NSR proposed in SPE, i.e. NSR is always assigned to the right-most element within a phrasal domain, and advances a quite different type of solution to non-final nuclear stress. She relates cases like (40a), in which a predicate follows its argument, to cases like (40b), in which more material follows the direct object.

(40) a. George has pláns to leave.

b. Helen left diréctions for George to follow.

The ingenious solution that Bresnan (1971) proposes to reconciliate the data with the NSR proposed in SPE is that the NSR applies *before and after* the object move to the sentence initial position. This has the effect the direct object receives nuclear stress in its in-situ position, i.e. all other words get demoted by 1. This copy is then deleted and the moved copy is the only one in the structure with number 1. In other words, the stress numbers that are assigned are exactly those that are assigned to a structure in which the direct object is in-situ and receives nuclear stress. Once again, argument structure is crucial—it is the base location of the object that is used to assign the correct stress, and it is in this position because it is the internal argument of the predicate.

(41) The Nuclear Stress Rule and Movement

$[_{NP}$ directions	$[_S$ for	George to	follow	directions.]]	
1		1	1	1	*MSR*
		2	2	1	*NSR*
				$\emptyset$	
1	2		2		*NSR*

An approach that merely states that predicates are less prominent than arguments (e.g. Schmerling (e.g. 1976) cannot account for (40b), since the subject of the embedded clauses is also subordinated.

In Bresnan's system, non-final nuclear stress in phrases must necessarily involve movement, in other words all head-final structures that show nuclear stress on the non-head must be derived from underlying head-initial structures. This approach has the advantage that there is no reference to argument structure in stating the NSR itself.

The observation that predicates are less prominent than their complements even when they follow them has an interesting consequence in the Bresnan-system: all cases in which nuclear stress falls on the complement in an OV structure, the order must be derived from underlying VO,; more generally, each head-final structure with nuclear stress on the complement must be derived from an underlying head-initial order.

The present approach is compatible though with the effects of movement that are discussed in Bresnan (1971), although the derivation of the observed facts works differently. Consider the following representation for (40b):

(42) [instructions [λx. for George to follow x]]

The top node of the tree has two daughters, the DP 'instructions' and a one place predicate, 'for George to follow x'. The DP is the complement of the λ-predicate, and consequently it is prosodically subordinated. The entire relative clause is the functor and it follows its complement.[9]

The principle of subordination can apply to syntactic nodes of any size. We will see other examples in which movement causes the subordination of remnant material, including the subordination of predicate sequences, of complex modifiers, of affixes, and of entire matrix sentences.

6.4 Predicates and Prosodic Asymmetry

This section presents evidence for the generalization in (16). Functors that follow their complement or part of it are prosodically subordinated.

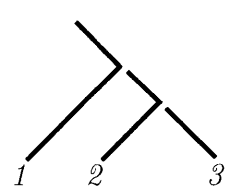
So far the main concern was data from English. I will now consider comparative data from different dialects of West Germanic. These languages differ in their prosody. However, once linear order is taken into account, most of the apparent prosodic differences (initial *vs.* final stress) reduce to syntactic differences (head-initial *vs.* head-final), and are as expected if the generalization in (16) holds.

6.4.1 One, Two, Three

Consider the case of predicate clusters in Dutch:

(43) Dutch Predicate Cluster: Final Stress

...[dát hij] [wílde hélpen vérven].
that he wanted.to help.to paint



[9] It is crucial that the relative clause is counted as a one-place predicate, and it is not treated as branching for the purposes of COMPLEMENT PROMINENCE.

The predicates are ordered according to their embedding, starting with the highest predicate. The expectation is then that the prosody should be just like in the English case: a sequence of predicates that are prosodically on a par. The predicted prosodic structure looks as follows:

(44)

×	×	×
×	×	×
wilde	helpen	verven

The actual output for (43) usually contains fewer accents than given here: The accent on 'helpen' is usually dropped, as indicated in (45a). Another possibility is to only place an accent on the last predicate (45b):

(45) a. ...[dát hij] [wílde helpen vérven]

b. ...[dát hij] [wilde helpen vérven]

Some predicates might be lacking accents due to rhythmic considerations. The distribution of accents on predicates seems to be sensitive to how closely adjacent accents are. This was observed also in Schmerling (1976, 98ff). This could be modeled by a phonological constraints forcing clash reduction that have the following effect:

(46) Clash Reduction

×	×	×		×		×
×	×	×	→	×	×	×
wilde	helpen	verven		wilde	helpen	verven

One indication that rhythmic considerations play a role is example (47b). If a preposition separates the last two predicates, they are separated enough to both maintain their accents.

(47) ...[dát hij] [wilde hélpen met vérven]

It seems that in a clash configuration some grid marks are deleted. The examples in (47a) and (47b) suggest that a relevant factor is whether two grid marks are separated by a foot on some lower grid line. Clash removal gets rid of the accent on 'helpen' in (45a) and on 'met' in (47):

(48) Rhythmic Patterns
Clash Removal in 'wilde helpen met verven'

∅	×		×
×	×	∅	×
×	×	×	×
wilde	helpen	met	verven

Clash Removal in 'wilde helpen verven'

×	∅	×
×	×	×
wilde	helpen	verven

The rhythmic nature of accent-placement in the pre-nuclear domain is further evidenced by (49a vs. b) and (49c vs. d) respectively.

(49) Hij zèi dat hij...

a. ...vérfde. $\acute{1}$

b. ...wil de vérven. $1 \prec \acute{2}$

...that he wanted to paint

c. ...wìlde helpen vérven. $\grave{1} \prec 2 \prec \acute{3}$

...that he wanted to help to paint

d. ...wìlde kunnen helpen vérven. $\grave{1} \prec 2 \prec 3 \prec \acute{4}$

..that he wanted to be able to help to paint

e. ... wìlde kunnen mògen helpen vérven. $\grave{1} \prec 2 \prec \grave{3} \prec 4 \prec \acute{5}$

... wìlde kunnen mogen helpen vérven. $\grave{1} \prec 2 \prec 3 \prec 4 \prec \acute{5}$

that he want be.allowed can help paint

'He says that he wants to be allowed to be able to help to paint.'

It seems that in deciding which accent to maintain and which accent to remove, there are at least three factors to be considered: the last accent is never dropped; the first accent is not dropped, unless there are only two accents or *all* pre-nuclear accents are dropped[10]; medial accents are often all dropped, but if not then an alternating

[10]The literature reports a tendency to maintain accents at the beginning of domains (Bolinger, 1958).

pattern is chosen.

The following example illustrates that rhythm must be evaluated rather globally. In a matrix sentence in Dutch, the first predicate is considered for the computation of rhythm on the predicate sequence, although it has moved to first position and is not part of the cluster:

(50) Hij wílde mogen kunnen helpen vérven.
he wants to be allowed to be able to help to paint.
'He wants to be allowed to be able to help to paint.'

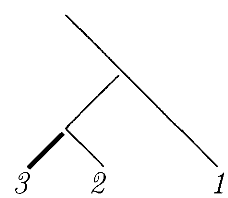
The reason I hesitate to formalize the rhythmic principles (which seem to require a rather global evaluation of eurythmy) is that it is not clear based on just those examples what the real pattern is. A closer investigation involving predicates of different shapes and sizes is necessary. This is a question for further research.

6.4.2 Three, Two, One

Consider now the German counterparts of the Dutch predicate clusters:

(51) German Predicate Cluster: Initial Stress

...dass er málen helfen wòllte.
...that he paint help want.



Main stress in German falls on the first predicate. No unreduced accents can occur in the post-nuclear domain unless one of the predicates is contrastively focused—in which case it would receive main prominence. The material following the accent on 'malen' has be severely pitch reduced if not deaccented. As indicated, there are intuitions about secondary stress in the post-nuclear domain.

German also differs from Dutch in the linear order of predicates, apart from the linear location of main word stress: while the predicates in Dutch are ordered according to embedding, the order in German is the exact inverse. The distribution of accents is then as expected.

There are secondary stresses after the nuclear one, but these cannot be realized as pitch accents. This is true independent of the number of predicates that follow:

(52) Er sag|te dass er...

a. ...mál|te. $\acute{1}$

b. ...mál|en wollte.. $\acute{2} \prec 1$

c. ...mál|en helfen wòllte. $\acute{3} \prec 2 \prec \grave{1}$

d. ...mál|en helfen können wòllte. $\acute{4} \prec 3 \prec 2 \prec \grave{1}$

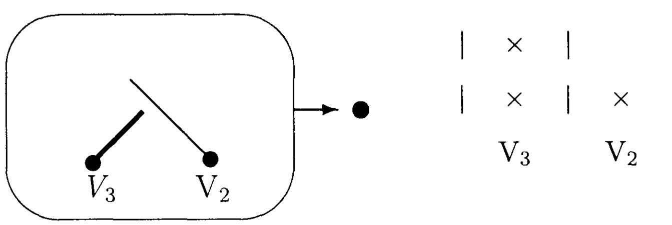
e. ...mál|en helfen kòennen dürfen wòllte. $\acute{5} \prec 4 \prec \grave{3} \prec 2 \prec \grave{1}$

that he want be.allowed can help paint

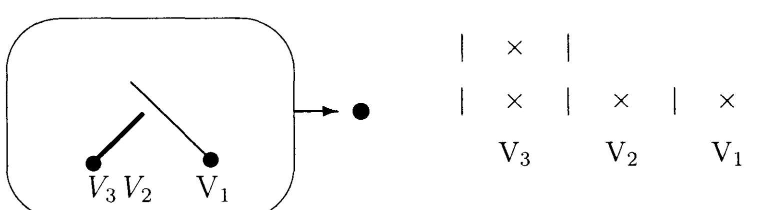
The fact that secondary stress in the tail of these clusters is rhythmic suggests that the prosodic structure in the tail is flat, and subject to rhythmic restructuring.

In order to derive the prosody in clusters with these linear orders, each predicate that is preceded by its complement must be subordinated. The derivation of (51) works then as follows, applying PROSODIC SUBORDINATION two times:

(53) First Cycle



(54) Second Cycle



The different word order compared to Dutch might be base-generated or one order might be derived by movement from the other—as long as movement involves λ-abstraction and no other operators it will leave the generalization in tact. Suppose the German order is derived from the Dutch order, then V_3 is still an argument of its sister:

(55) $[\ V_3\ ]\ [\ \lambda x.\ [\ V_2\ x\ ]\ ]$

If conversely the Dutch order is derived from the German order, then again, the functor-argument relations are not changed, only the word order is:

(56) $[\ [\ x\ V_2\ ]\ \lambda x\ ]\ [\ V_3\ ]$

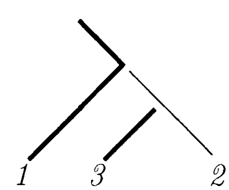
According to the theory presented here prosody does not give any evidence about which if any of the two orders is more basic.

6.4.3 Mixed Orders

That Dutch and German indeed do not differ in their prosodic systems becomes apparent in predicate cluster orderings that are attested in both languages:[11]

(57)

a. ...dat Ján Maríe$_{given}$ kan$_1$ gezíen$_3$ hebben$_2$.
that Jan Mary could seen have
'that Jan could have seen Mary.'

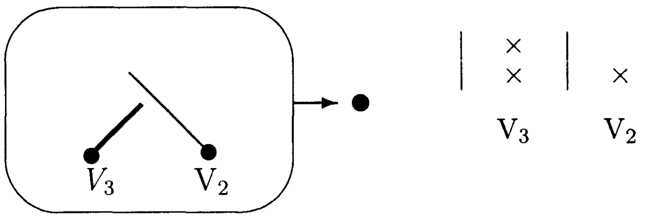


b. ...wéil sie ihn hat$_1$ málen$_3$ wollen$_2$.
because she him has paint wanted
'that she has wanted to paint him'

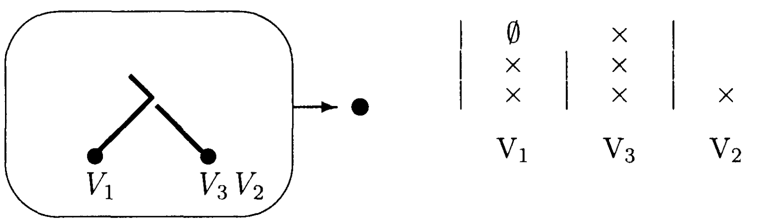
[11]Here, the DP argument preceding the cluster is made 'given' (old information), in order to prevent subordination of the cluster.

In this example, the second predicate 'hebben' is preceded by its complement 'gezien', which effects its subordination. The complete derivation involves a step of PROSODIC SUBORDINATION and a step of PROSODIC PROJECTION, followed by clash removal:

(58) First Cycle



(59) Second Cycle and Clash Reduction

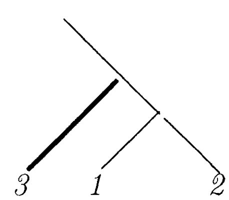


The next example illustrates a case where it seems that only a sub-constituent from the complement precedes one of the predicates (this word order is not accepted by all native speakers of Dutch, but is apparently generally accepted in Flemish). Predicate 1 is followed by predicate 2, but preceded by predicate 3, which semantically is part of its argument.

(60)

a. ...dat Ján Maríe gezíen$_3$ kan$_1$ hebben$_2$.
that Jan Marie seen can have
'that Jan can have seen Marie.'

b. ...wéil er es káufen$_3$ wird$_1$ können$_2$.
because he it buy will can
'because he will be able to buy it.'



This type of order can be derived by positing movement:

(61) $[\ V_3\ ]\ [\ \lambda\ x.\ [\ V_1\ V_2\ x\ ]\ ]$

The derivation proceeds then with a step of PROSODIC PROJECTION assigning a prosody to $[\lambda x.V_1 V_2 x]$ creating a one-place functor, followed by PROSODIC SUBORDINATION. The derived representation looks as follows:

(62) Prosody for 3 $\prec$ 1 $\prec$ 2

```
| × |
| × | ×     × |
| × | ×  |  × |

  V3   V1    V2
```

The following cases of particle climbing further illustrate that entire predicate sequences can treated as functors and be subordinated.

(63) Climbing up the Cluster (cf. Evers, 2001, and ref. therein)
Het labyrinth waar we hem niet over...

a. zullen hoeven laten **ná** denken. $\grave{1}\prec 2\prec 3\prec \acute{5}\prec 4$

b. zullen hoeven **ná** laten denken. $\grave{1}\prec 2\prec \acute{5}\prec 3\prec 4$

c. **ná** zullen hoeven laten denken. subordination $\acute{5}\prec 1\prec 2\prec 3\prec 4$

about will need let think
'The labyrinth about which we won't let him reflect.'

I assume again that movement of the particle creates a one-place predicate:

(64) $[\ V_5\ ]\ [\ \lambda\ x.\ [\ V_1\ V_2\ V_3\ V_4\ x\ ]\ ]$

As expected the entire functor following the particle is prosodically subordinated.

6.4.4 Predicate Sequences and DP Arguments

Subordination of predicates following arguments can also be observed in Dutch and German, even when multiple predicates follow an argument.

In Dutch, the entire cluster subordinates when the DP complement of the 'deepest' predicate precedes it, as it usually does. The prosody is exactly parallel to the case of particle preposing discussed above: the entire sequence of predicates is turned into a single functor that then takes the direct object as its complement:[12]

(65) a. ...dat hij [een múur$_6$ wilde$_1$ mogen$_2$ kùnnen$_3$ helpen$_4$ vèrven$_5$.]
that he a wall want allow can help paint
'he says that he wants to be allowed to be able to help to paint a wall.'

b. ...wéil er [ein Bíld$_6$ malen helfen kòennen dürfen wòllte.]
because he a picture paint help can be.allowed wanted

If the direct object encodes 'given' information, the predicates are not deaccented, however:

(66) a. ...dat hij [de múur$_6$ wílde$_1$ mogen$_2$ kúnnen$_3$ helpen$_4$ vérven$_5$.]
that he a wall want allow can help paint

b. ...wéil er [das Bíld$_6$ málen helfen können dürfen wollte.]
that he the picture paint help able allowed wanted

'...that he wants to be allowed to be able to help to paint a wall.'

This effect will be investigated in chapter 7.

[12]The post-nuclear stresses, both in Dutch and German, cannot receive an accentual realization unless the relevant predicate is focused or the preceding argument constitutes given information. A focused predicate, however, would receive nuclear stress. All predicates following the direct object are subordinated, that is, they are at least severely reduced in pitch range, if not altogether deaccented. The rhythmic pattern in the tail is the *same* in Dutch and German, despite the different word orders. This is as expected, since cyclic subordination derives a flat tail following the direct object for both languages.

Discussion

Across the different languages, the prosodic generalization about predicate sequences is the same: the last accent falls on the complement that is selected last. This is captured here by the interaction of the principle of COMPLEMENT PROMINENCE with PROSODIC SUBORDINATION.

The generalization that nuclear stress falls on the most deeply embedded predicate can be captured applying various different approaches (Cinque, 1993, 269ff), Arregi (2002), or Jacobs (1991a, 1992)). The prosodic asymmetry in (16) discussed here however is not accounted for in those earlier approaches.

The generalization seems to be the same in English, Dutch, or German. While these languages differ in the precise tonology that is mapped onto the grid structures, the metrical grid is derived by the identical mechanism. The variation lies in the word order.

There is more evidence that the prosodic systems of those languages are identical. We can replicate all the possible orders of predicates across dialects in a *single* language. The following paradigm shows three of the possible orders of a particular predicate sequence in German. When predicates are ordered according to embedding as in (a), this order is often taken to involve 'extraposition'. Different orders are possible, however, in so-called 'restructuring' environments (e.g. in b,c).[13]

(67) '...weil er ihr...
'...because he...

a. [versprách] [zu versúchen] [zu schwéigen]. $\acute{1} \prec \acute{2} \prec \acute{3}$

b. [versprách] [zu schwéigen zu versuchen]. $\acute{1} \prec \acute{3} \prec 2$

c. [zu schwéigen zu versuchen versprach.] $\acute{3} \prec 2 \prec 1$

d. [zu schwéigen versprach zu versuchen.] $\acute{3} \prec 2 \prec 1$

[13]There are many syntactic differences between 'extraposition' and 'restructuring' that I will not address in this paper—restructuring derives what appear to be mono-clausal constructions that, e.g. , facilitate scrambling between clauses, and allow pronouns that are arguments in the lower clause to be affixed on the matrix verb in second position, etc. (cf. Wurmbrand, 2003). Restructuring does not always result in a different word order between the predicates (Haider, 1994).

be silent try promise

'...promised her to try to be silent.'

The example (67a) is similar to predicate clusters in Dutch (49), in that main stress is rightmost and secondary accents precede the main one. The fact that the median predicate does not necessarily lose its accent rhythmically as in the Dutch example (43) (although it may in fast speech), maybe due to the fact that there is unstressed phonological material—the preposition—intervening, preventing a clash. This is similar to the pattern observed in the Dutch example in (47), where also a preposition separated two predicates.[14]

Bech (1955/57) discusses stress and syntax of predicate clusters in a dialect that allows the following:[15]

(68) Unacceptable in Standard German, yet speakers have intuitions about main stress...

a. ??man wìrd ihn hier können líegen bleiben lassen.

b. ??dass man ihn kann lassen líegen bleiben.

c. ??dass man ihn hier wird können lassen líegen bleiben
that one him here will can let lie stay
'that one will be able to let him stay lying here.'

[14]Since English does not allow cluster formation and has a rather fixed word order between predicates, only the evidence from DP's preceding predicates discussed in the next subsection serves to show subordination in this language. A look at complex nominals, however, shows further evidence for the asymmetry at work:

(1) a. [a tráiner] [of sláyers] [of vámpires]

b. [a tráiner] [of vámpire slayers]

c. [a vámpire slayer trainer.]

[15]Compare: Voicing assimilation in English plurals is not segment-based: Speakers know that the plural of Bach is Bachs. (Morris Halle) Thanks for David Pesetsky for pointing out the parallel in the argument.

Speakers of Standard German know a syntactic generalization about stress—not a dialect-specific stress rule. The evidence discussed so far in this chapter suggests that the grammar of English and Dutch involves the exact same generalization about prosody.

The system has nothing to offer in order to explain *how the different linear orders come about.* This is a research question that is the topic of much current debate, and I will not endeaver to explore the issues here (for discussion see the collection of articles in Kiss and Riemsdijk (2004)).

6.5 Further Evidence for Asymmetry

This section presents further cases of prosodic subordination. The uniting factor is that they involve functors taking arguments, and the functors are spelled out before the argument.

6.5.1 Modifiers

Modifiers in Germanic differ from arguments with respect to their prosody relative to a following predicate (Gussenhoven, 1984, Krifka, 1984, Selkirk, 1984, Jacobs, 1992, Truckenbrodt, 1993). Consider first the case of arguments (the relevant predicates are in italics, the relevant arguments are underlined).

(69) Arguments

a. [Sie hát] [einen Tángo *getanzt.*]
   she has a.acc Tango danced
   'She danced a tango.'

b. [Sie íst] [in den Gárten *getanzt.*]
   she is in the.acc garden danced
   'She danced into the garden.'

Predicates following DP-arguments (69a) and PP-arguments (69b) are subordinated. Subordination means that they are realized with a reduced pitch range, if not altogether deaccented. They are prosodically 'affixed' to the preceding domain.[16]. Consider now the case of modifiers (again, I put the functor in italics, and I underline the argument of the modifier, i.e. the modified constituent):

(70) Modifiers

a. [Sie hát] [*den ganzen Ábend*] [<u>getánzt.</u>]
she has the.acc entire evening danced

'She danced the entire evening.'

b. [Sie hát] [*im Gárten*] [<u>getánzt</u>].
she has in.dat garden danced

'She danced in the garden'.

Modifiers in (70) are mapped to their own accentual domain, separate from the accentual domain of the verb.[17] The difference between modifiers and arguments outlined so far is only one half of the pattern, however.

A closer look at the prosody of modifiers reveals an asymmetric pattern depending on linear order—in fact, a pattern that is exactly parallel to the asymmetry observed between predicates and arguments (Wagner, 2004a).

[16]One could say that the arguments 'incorporate' into the prosodic domain of the verb that selects them; or conversely, that the verb becomes affixed to the prosodic domain of the argument. Indeed, structures that involve actual incorporation or affixation in German show similar prosodic patterns. I will not explore this point further in this paper.

[17]The following examples were presented in Cinque (1993, 250, 254) with nuclear stress on the adjunct—this, however, is not the appropriate prosody in neutral contexts. Unless the modifier is focused, nuclear stress in these examples falls on the verb, as was pointed out already in Truckenbrodt (1993), Jacobs (1999), and Ishihara (2001).

(1) a. ...[dass Kárl] [ein Búch] [mit Múehe] [lésen kann].
that Karl a book with problems read

b. ...[dass Frítz] [gút] [kóchen kann].
that Fritz well cook can.

The word order in (1) actually seems to require marked focus on the predicate. The more neutral order is: 'dass Karl mit Múehe ein Búch lesen kann'. The example may then actually not be relevant for establishing the neutral stress pattern, since both word order and focus structure are not neutral.

When a modifier precedes the modified constituent, two separate prosodic constituents are derived, as was illustrated in (70); however, when it follows the modified constituent, the modifier is subordinated. Consider as a first example the case of post-posed sentential modifiers:

(71) Postposed Adjuncts

a. [Sie hát] [getánzt, *den ganzen Abend*].
she has danced the.acc entire evening
'Sie hat danced the entire evening.'

b. [Sie hát] [getánzt, *im Garten*].
she has danced in.dat garden
'She danced in the garden'.

The prosodic relation between modifiers and modified constituent is similar to that between predicates and arguments. Earlier treatments (e.g. Krifka, 1984, Gussenhoven, 1983, Selkirk, 1984, Jacobs, 1992) focused on the contrast between modifiers and arguments, and thus missed out on the parallel between modifiers and predicates.[18] More evidence for prosodic asymmetry in modification is presented in Wagner (2004a).

The prosodic parallel between modifiers and predicates can be linked to a semantic parallel: it is the modifier that is the functor and that takes a complement, the modifiee (Bierwisch, 2003). But if modifiers are the functor, then 'Functor First' requires them to be spelled out before the modifiee.

Modifiers that precede their complement can also be subordinated and remain unaccented, just like predicates in VO configurations. Again, whether or not they are depends on rhythmic factors:

[18]It is not obvious whether it is legitimate to call a phrase like 'the entire evening' a modifier, and group it with complement taking predicates, since maximal projections are not usually taken to be able to take complements. Starke (2001) however argues for treating entire specifiers as projecting the feature of a functional projection and thus being the projector. Another way to make sense of this is to adopt the analysis of adverbials proposed in Alexiadou (1997) and Cinque (1999), who view adverbials as specifiers of functional projections. It is not 'the entire evening' itself then, which is the modifier, but a functional head that takes the DP as its specifier.

In (72a), it is possible to omit the accent on the manner adverb 'gut' in (72a) without changing the information structure of the sentence. This seems harder in the case of a longer adverb, that places more unstressed material between its main stress and the following verb (72b).

(72) Sie gláubt, dass er
she believes that he

a. [gut kóchen kann].
well cook can.

b. [áusgezeichnet] [kóchen kann.]
excellent cook can.

Selkirk (1984, 231) states that modifiers can be stressed without affecting the information structure. The example used also involved a modifier that *precedes* its modifiee:

(73) Jóhn bought a red tíe.

The expectation based on the proposal here is that modifiers that precede the modifiee can but do not have to be accented, while modifiers that follow their complement must be deaccented.

Some modifiers do not seem to obey the generalization. They are accented although they follow what they modify. But there is reason to believe that the modifiee *is not their complement*. They differ in their syntactic relation to the modifiee, and have a structure similar to the relation between a secondary predicate and its argument.

An instructive case to look at are locatives. Locatives have different readings, again depending on where in the structure the locative is placed. Consider the following wide-scope locative, which Maienborn (2001) characterizes as setting the frame for the proposition expressed by the rest of the sentence. The presence of an additional sentential locative modifier forces the frame-setting reading (74), since two locatives of the same type would be contradictory:

(74) In Italy, Lothar bought his suits in France. (Maienborn, 2001, 197)

The intended reading is one where the frame setting locatives narrows the time under consideration to 'the time when Lothar was living in Italy', while the lower locatives specifies where the events of buying suits took place. There are two possible linearizations of frame-setting locatives, and again the prosodic asymmetry is observed:[19]

(75) a. [In Ítaly,] [Lóthar bought his súits in Fránce].

b. [[Lóthar bought his súits in Fránce] in Italy].

Frame-setting locatives contrast with lower locatives such as 'in France' in the examples above, which receive an accent sentence finally . Another difference between the two types of locatives is that lower Locatives cannot occur in sentence-initial position:

(76) a. *[[In Fránce,] [Lóthar bought his súits] in Italy].

b. *[[In Ítaly] [In Fránce,] [Lóthar bought his súits].

Frame setting locatives pattern just as is expected of functors: when they follow their complement, they are subordinated, but not (necessarily) when they precede their complement.

Low Locatives do not show the pattern expected of functors. But VP-final modifiers also differ in their syntax, and there is good reason to assume that they are not functors that take the VP as their complement. Larson (2005) argues that they are event predicates, that contain an event variable, and assigns a right-decending structure, just as in coordination structures. If Larson's coordination analysis is correct, then the accentedness of low locatives is expected. This point is discussed in detail in chapter 3.

A final issue that again arises both for predicates and for modifiers is that of specifiers. Just as in the case of transitive predicates, an issue arises with transitive modifiers. Consider the nominal adjuncts involving 'with'. 'With' would have to be analyzed as a modifier taking two arguments:

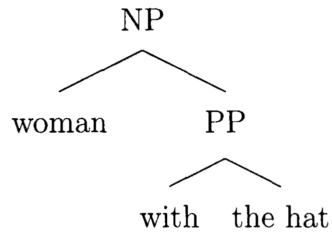
[19] Jacobs (2001) and Maienborn (2001) note that frame-setting modifiers are similar in syntactic position and in their semantics to topics. The same parallel may apply to the intonational realization of these modifiers and topics. I will not explore this in any detail here.

(77) The wóman with the hát.

Given the analysis so far, the modifier should be the main projection of the construction. 'with' combines with 'the hat' forming the modifier 'with the hat', modifying the head noun 'woman'. We would expect then subordination of 'with the hat', since it follows the modifiee—a prediction that is quite apparently false. Why does the principle of subordination fail to apply?

Note that even if 'with' is the main predicate, the entire constituent is not a PP, but an NP. The determiner 'the' selects a noun, not a prepositional phrase.

(78)



Of course, if this is the correct structure and the projector is really the head noun 'woman', the problem with respect to the prosody disappear: the projector precedes the non-projector, so two accentual domains are expected.

While 'woman' is the argument of the functor 'with the hat', the functor remains unsubordinated and receives its own accentual domain since it is its sister that projects.

6.5.2 Affixes

The principle 'Functor First' is also at play when syntax combines morphemes with each other. There is a well-known asymmetry between prefixes and suffixes which is observed across the Germanic languages but also attested cross-linguistically (e.g. Cohn, 1989, for Indonesian). Prefixes can form a prosodic domain of their own, while suffixes prosodically subordinate, i.e. they are realized with an extremely reduced pitch range if they carry accents at all:

(79) Stress: Prefixes vs. Suffixes

a. [prè] [émptive] *vs.* [[trúst] wòrthy]

b. [nòn] [hỳper] [artìculate] *vs.* [[[fríend] ship] wòrthy]

In Germanic, there is a correlation that suffixes tend to encode heads that determine the syntactic category while prefixes are modifying heads that leave the category of the stem unchanged. Since both modifiers and derivational heads can be conceived of as functors, this difference is not relevant for the prosodic facts we are interested in.

Just as with other functors, e.g. those entering predicate clusters, suffixes can be recursively added, and each of them is prosodically subordinated:

(80) Recursive Affixation

kost	stem
kost bar	'expensive'
kost barkeit	'valuable object'
kost barbkeitslos	'without a valuable object'
kost barkeitslosigkeit	'state of being without a valuable object'
kost barkeitslosigkeitshalber	'because of a state of being without—'

The rhythmic distribution of intuitions about secondary prominence in the tail of derivatives suggest that the suffixes are on a par with respect to each other, just as is it is the case in predicate clusters.

6.5.3 VP-Preposing

The asymmetry presented so far relates prosody and linear order: if a functor *precedes* its complement it may but does not have to subordinate; if it *follows* its complement it is obligatorily subordinated. The asymmetry was first illustrated based on predicates and infinitival and nominal complements; it has been illustrated in a variety of other cases, including function words, instances of modification, and affixation.

A question not touched upon in this chapter so far is how the different linear orders between functors and their complements come about. Under an asymmetric view of syntax (e.g. Haider, 1993, Kayne, 1994) one order is basic and the other order is derived. Consider the following prosodic pattern of VP remnant fronting in German:

(81) a. [Péter [hat der María] [einen Kúss gegeben]
Peter has the Mary a kiss given

b. [[Einen Kúss gegeben] hat Peter der Maria.]
a kiss given has Peter the Mary

c. [[Einen Kúss gegeben] hat Peter der [Mariía]$_{Focus}$]
a kiss given has Peter the Mary

The fronted VP receives an accentual domain, the remaining part of the sentence is subordinated—unless there is a focused constituent in the remnant material. This, of course, is exactly the asymmetry observed earlier for predication and modification.

It seems tempting then to analyze the reorderings at stake in all examples as cases of movement. A comprehensive syntactic account of the derivation of linear order in tandem with prosody remains to be developed.

6.5.4 Matrix Clauses

The asymmetry can be observed for functors and complements of varying size. Just as it is attested in the combination of morphemes, it is also attested in the combination of sentences. In order to illustrate this point, I will use examples from German.

A matrix clause can either precede or follow an embedded clause. When it precedes it, both clauses are mapped to separate prosodic feet that are on a par:

(82) [Anna sagte zu Ulli] [wahrscheinlich werde es regnen]
anna said to ulli probably will it rain
'Anna said to Ulli that it will probably rain.'

However, when the matrix clause follows its complement, is is prosodically subordinated. (I mark the subordinated part by underlining):

(83) [wahrscheinlich werde es regnen] sagte Anna zu Ulli]
probably will it rain said anna to ulli
'Anna said to Ulli that it will probably rain.'

Matrix clauses can be recursively nested:

(84) [wahrscheinlich werde es regnen] meine Hans behauptete Anna]
probably will it rain thinks Hans claimed Anna
'Anna claimed that Hans thinks it will rain tomorrow.'

How can we account for this pattern? A matrix clause contains a predicate that takes an embedded clause as its complement. An important fact is that complement clauses obligatorily extrapose in German. This can be shown looking at cases where there is a clause final participle:

(85) a. Hans hat gesagt, dass es morgen regne.
Hans has said that it tomorrow rains

b. * Hans hat dass es morgen regne gesagt.
Hans has that is tomorrow rain said

Matrix clauses can then be viewed as functors and embedded clauses as their arguments once extraposition has applied:

(86) λx. Peter said x (
[It will rain tomorrow.])

Since the matrix clause is the functor, and the embedded clause the argument, the linear asymmetry is now just an instance of the more general asymmetry described by (16). For a discussion of the syntactic and prosodic properties of these reorderings see Wagner (2004b).

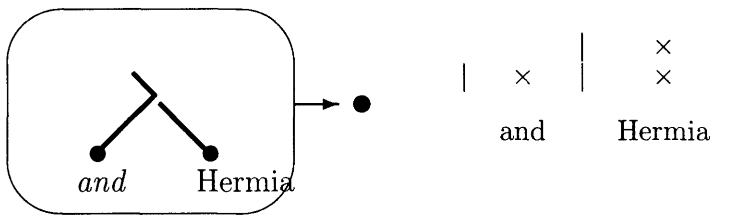
6.5.5 Function Words

In the discussion of coordination in chapter 2 it was left open how to account for the prosody of the connectives 'and' and 'or'. Each conjunct receives an accent, but connectors usually do not:

(87) Lysánder and Hérmia.

The subordination of function words is just another instance of the subordination of functors more generally. What is particular with function words is that usually, subordination takes place even when the functor precedes the argument. This is optionally possible with *any* functor.

(88) a. First Cycle



Prepositions can quite generally subordinate and remain unaccented:

(89) a. of Demétrius

b. from Áthens

c. has líved

But note that the subordination of prepositions depends on syllable count, suggesting that this again a rhythmic effect, while post-positions obligatorily subordinate:

(90) a. entláng des Mississíppis
along.location the Mississippi

b. # den Mississíppi entláng
the Mississippi along.direction

Adpositions thus give more evidence for prosodic asymmetry. Some function words have completely stressless allomorphs, e.g. 'v, ans 's for 'have' and 'is'. The hypothesis that I want to advance here is that function words can be only be subordinated or stressless when they occur in a position that *is independently prosodically subordinated*. In other words, there is nothing special about function words—they simply occur in weak positions, just like other functors. Stressless allomorphs are special in their underlying form, but they are not a different grammatical category.

If this is the correct account of the subordination of function words, then only functors that take complements should reduce in this way. Intransitive prepositions (Emonds, 1976) indeed do not subordinate (thanks to Draga Zec for suggesting intransitive prepositions as an example). This prediction is borne out:[20]

[20] A similar contrast exists with focus particles. While 'only' may subordinate, 'too' does not and is accented:

(91) a. He wént awáy.

b. She lóoked aróund.

c. She wálked ín.

However, there are function words, or at least allomorphs of function words, that can never be accented or receive any prominence. I will assume that these are lexical allomorphs that are marked as prosodically deficient.[21] One possibility is to assume that these are lexical allomorphs specified to occur only in weak positions. These allomorphs develop for some lexical functors but not others because they occur more frequently.

Another fact that we can explain now is why the prosodic 'size' of functors should depend on their level of embedding, a result from chapter 5. Consider the following structure. The prosodies of the connectors i differ in at what grid line they form a foot. The 'or's form feet at a higher line compared to the 'and'.

(92) A or B or (C and D)

```
| × |     × |     ×       × |
| × | × | × | × | ×       × |
| × | × | × | × | × |     × |
| × | × | × | × | × | × | × |
  A   or  B   or  C  and  D
```

In chapter 5 I presented evidence that this is reflected in their phonetic realization: the degree of lengthening correlates with the strength of the upcoming boundary. Function words are prosodically subordinated *to the next lower grid line.*

(1) a. Máry only láughed.

b. Máry laughed, tóo.

But here, it is less obvious what the syntactic struture at least of 'too' is.

[21]There are also function words that lean to the 'wrong' side.

(1) a. a friend'v Jóhn

b. John's in Berlín.

Many approaches to function words, e.g. Selkirk (1996), make function word specific constraints. The implementation works by positing alignment constraints that make reference to the distinction lexical word only (*vs.* function word). Since function words are invisible to alignment constraints, they end up dangling on to the prosodic constituents they are adjacent to:

(93) (fnc (lex $)_{\omega})_{\phi}$

The class of function words is assumed to be defined based on independent syntactic grounds. The account does not explain though why it is that only function words subordinate that have a complement. One factor could be that they are phrase final. But many function words do subordinate phrase finally:

(94) a. What are you tálking about?

b. Who are you thínking of?

Phrase finality does have an effect here though: phonetically reduced versions of function words are often not allowed when the complement is extracted, even when the preposition subordinates (see Selkirk (1996) and references for evidence):

(95) * Who are you thinking'v

But still, the fact remains though that intransitive but not transitive functors subordinate as illustrated by (94) *vs.* (91) , and this fact receives a straightforward explanation within the allomorphy approach proposed here, under which the use of a reduced function words is parasitic on the it being independently subordinated.

Another issue is the notion 'function word' itself. It is not clear that there is a definition of function words that separates them from other functors in such a way that correlates with prosodic behavior. One criterion offered in Abney (1987) for function-wordhood is the open class/closed class distinction. But this distinction is not very helpful in predicting prosody. Prepositions are a closed class, and lexical verbs are an open class. But lexical verbs with the same or nearly the same segmental content have exactly the same prosody as prepositions:

(96) a. hìnder den Tórmann!
block the goalie
'Block the goalie!'

b. hìnter den Tórmann!
behind.direction the goalie
'Behind the goalie!'

Some prepositions (of, for...?), namely those that have prosodically sharply reduced allomorphs seem to be function words and are completely stressless, but others (along, behind...) don't seem to be any different from lexical functors. Another criterion for assigning a morpheme to the class of function words that Abney (1987) discusses is prosodic reduction—which of course we cannot use here. Another less clearly defined notion is lack of lexical content. Unless there is any *syntactic* definition of exactly function words that are *prosodically* special, the theory is circular.

It seems more adequate then to treat Function words just as any other functor, but stipulate that those functors that occur in subordinated positions (i.e. those that take complements) can develop stressless allomorphs that are obligatorily stressless.

Which function words develop those allomorphs might relate to frequency, or *probability* as in (Jurafsky et al., 2000). If the idea that frequency or probability is relevant is correct, this might explain why it seems that whether or not words have *alternatives* in a given slot matters (over *vs.* under *vs* on...). In many contexts, there are few or no alternatives, and thus the probability of what can occur in those slots is very high (auxiliary positions, instances of 'of' and 'to'). There is thus a potential explanation for the special status of some functors that does not require making a syntactic distinction between function word and lexical word.

A look at more sets of facts involving function words e.g. the one discussed in Selkirk (1996) and Zec (2002) would be useful at this point. The intricacies are mostly related to the fact that function words can be stressless and be incorporated into the prosodic word domain of their host. I will not discus word morphology in any detail, and leave this for some other occasion.

The main points of this section were that (i) function words are subordinated just were other functors are and show the same asymmetry depending on linear order;

(ii) only functors that take a complement have weak allomorphs, but not intransitive ones—these are the ones that do not undergo subordination since they are the complement of a predicate; (iii) weak allomorphs of functors cannot occur in positions that are not independently subordinated.

Summary

This chapter presented a generalization about prosodic asymmetry. Functors that follow their complement are obligatorily prosodically subordinated, but functors that precede their complement do not have to be.

The asymmetry can be captured by the interaction of syntactic recursion with two principles of syntax-to-phonology mapping, PROSODIC MATCHING and PROSODIC SUBORDINATION.

A look at predicate sequences illustrated that apparent prosodic differences between English, Dutch, and German reduce to independently motivated syntactic differences. The analysis can be applied to a variety of constructions involving functors of different kind and size.

The Cyclic application of PROSODIC MATCHING and PROSODIC SUBORDINATION derive nuclear stress: the last unsubordinated constituent counts as the nuclear stress. This is an architectural difference to alternative approaches to nuclear stress who usually first assign nuclear stress and then specify special rules to capture the difference between the post-nuclear and the pre-nuclear domain. The advantage of the approach in terms subordinating and then deriving nuclear stress indirectly is that it lends itself easily to incorporate information structural effects that can also determine the choice between PROSODIC MATCHING and PROSODIC SUBORDINATION.

Chapter 7

Givenness and Locality

Abstract

Contextually given material is deaccented under certain circumstances. This chapter proposes a new generalization about when givenness marking is possible. The generalization requires a notion of RELATIVE GIVENNESS. The prosodic marking of givenness shows the same asymmetry that is observed in cases of predication. This can be accounted for by using the tool of Prosodic Subordination from chapter 6.

7.1 Relatively Given

Constituents that encode information that was made salient by prior context often remain unaccented. Consider a case in which the direct object is deaccented:

(1) Mary sat at her desk. John walked in. What happened next?

a. # She kissed JOHN.

b. She KISSED John.

Context can lead to deviations from the otherwise expected or 'neutral' stress pattern. The last accent in (1b) falls on the verb, not on the direct object, as would be expected had there been no mention of Mary in the context. What counts as 'given' and which (if any) stress pattern counts as 'neutral' has been a major issue in the literature on accent placement.

In recent years the issue was often framed in terms of the notion 'givenness', and the aim has been to specify conditions on deaccentuation (Selkirk (1995b), Reinhart (1995, et seq.), Williams (1997), Schwarzschild (1999), Büring (to appear), Sauerland (2005)).

The proposal here is based on a notion of RELATIVE GIVENNESS. I present evidence that constituents that are marked as given are marked as given *relative to something else*. The idea is implemented using a presuppositional approach to givenness marking.

7.1.1 The Problem

Contextually given material is often prosodically subordinated or 'deaccented' in Germanic. A key example from the recent literature is (2) (Schwarzschild, 1999, 146):[1]

(2) John drove Mary's red convertible. What did he drive before that?

[1]In this chapter, I will often only mark the *last* accent in every example by capitalization, i.e. the accent that is perceived as the nuclear stress of an utterance. The presence or absence of pre-nuclear accents is not discussed unless it is specially mentioned. E.g. in the example below, (b) is claimed to be infelicitous regardless of whether or not 'blue' is accented.

a. He drove her BLUE convertible.

b. # He drove her blue CONVERTIBLE.

The constituent 'convertible' is marked as given. Sauerland (2005) proposes a presuppositional account of this effect. Other presuppositional approaches to this kind of contextual effects on prosody are Chomsky (1971) and Rooth (1992b), who use the notion of 'focus'. Sauerland (2005) differs from these accounts among other things in that it is not the *focused* but the *given* constituent that introduces a presupposition. The notion 'focus' plays no role this approach.

Presuppositions are requirements that certain elements impose on the context. If these requirements are not met, then a sentence containing the element remains without truth value. I will use the notation in Sauerland (2005), who puts the presupposition of an expression in the numerator of a fraction and the value of an expression in the denominator:

(3) $\lambda x. \frac{presupposition(x)}{value(x)}$

According to Sauerland (2005), the given predicate 'convertible' in (2a) carries a G-marker.

(4) He drove a BLUE convertible$_G$.

This operator G introduces the presupposition that the predicate 'convertible' is given, and requires there to be a true proposition of the form 'x is a convertible' in the context:[2]

(5) $[\![\text{G}]\!] = \lambda f^{et} \frac{\exists \in D^e . f(x)=1}{f}$

The presuppositional notion of 'givenness' explains why givenness marking is obligatory when possible. This is ensured by the principle 'maximize presupposition'

[2]The presupposition is really that the predicate f(x) must be given. Presupposing the existence of a true predicate seems to be too strong.

(Heim, 1991). The same point is made in Sauerland (2005), and applies equally to other presuppositional approaches, e.g. Rooth (1992b).[3]

This presupposition, however, is too weak. It would also be fulfilled in the following context, in which deaccenting is infelicitous:

(6) Mary's uncle, who produces high-end convertibles, is coming to her wedding. I wonder what he brought as a present.

a. He brought a [CHEAP convertible].

b. # He brought [a RED convertible].

c. He brought [a red CONVERTIBLE]

It seems that what is wrong with (6b) is that 'high-end' is not a relevant alternative to 'red', and therefore a contrast does not make a lot of sense. Marking 'convertible' as given is fine in (6a), since 'cheap' and 'high-end' are alternatives to each other.

However, in order to even state this, we have to use a vocabulary that falls outside of the scope of the givenness presupposition in (5), namely we have talked about 'contrast' and 'alternative'.[4]

That indeed 'high-end' and 'blue' are not alternatives can be motivated independently. One piece of evidence that 'high-end' does not qualify as an alternative for

[3]In Schwarzschild (1999)'s approach, this is achieved by the constraint 'Avoid F' which minimizes F-markers in syntax. The explanation in terms of 'maximize presupposition' seems more insightful, since it links the requirement to a more general phenomenon, e.g. the choice between definite and indefinite article discussed in Heim (1991). Schwarzschild (1997) proposes an 'Attentiveness Maxim', which is similar in spirit to 'maximize presupposition. I do not have the space to outline this approach in detail here.

[4]The same problem arises for the account of givenness in Schwarzschild (1999), at least according to the discussion on page 152/153. Sentence (1a) is argued to require the antecedent in (1b):

(1) a. eat a GREEN apple

b. eat an apple

Since apples have colors (b) should be a sufficient antecedent for (a). By analogy a 'high-end convertible' should be more than sufficient as an antecedent in (6) since entails that the givenness of 'convertible'.

'red' comes from association with focus in exclusives:

(7) Mary only likes RED convertibles.

Sentence (7) does not rule out that Mary might like high-end convertibles, unless the context is such that it made salient a partition of convertibles into red ones and high-end ones. This shows that 'high-end' is not usually considered an alternative to 'red', although of course a more elaborate context might make a partition of convertibles into red ones and high-end ones salient.

It seems that the sentence in (6a) does not just require that the predicate 'convertible' is given; it also requires that there must be an alternative x to its sister, such that [x convertible] is given. 'High-end' does not qualify as an alternative to 'red', so there is a presupposition failure. But why would it be relevant that there is an alternative to the sister of the constituent marked as given? There is no such expectation based on the presupposition in (5).

7.1.2 The Solution

The requirement that there be a contextually given alternative for the sister of the constituent that is marked as given can be called the 'sister-restriction'.

The sister-restriction points to a stronger presupposition than the one in (5): In order to be marked as given, a constituent has to be *given relative to its sister*. Marking a constituent x as given introduces the presupposition that there is alternative y' to its sister y such that the constituent $[y'x]$ is given. I define then the operator G_R as follows:

(8) Relative Givenness

$$[\![\mathrm{G}_R]\!] = \lambda x.\lambda y.\frac{\exists y' \in Alt(y): [\![y'x]\!] is\ given}{[\![yx]\!]}$$

We can now define a preliminary notion of 'neutral stress', which is a useful notion to specify a special set of contexts to investigate prosody.

(9) Neutral Stress (to be revised in (45))

A stress pattern is neutral if it does not trigger a presupposition of relative givenness for any sub-constituent.

Not just lexical items but also complex nodes can be marked as given. Consider the case of prosodic subordination of entire VPs:

(10) Mary praised John. What happened next?

[ANNA [praised John.]]

Within the VP, 'John' and its sister are not given relative to each other respectively. So the relative prosody between the two is as in the neutral case: they are on a par prosodically. But at the next higher node, the VP is given relative to 'Anna'. The alternative to 'Anna' that is salient is 'Mary', and 'Mary praised John' is given. Therefore the entire VP is prosodically subordinated.

7.1.3 Relative Givenness and Rooth (1992)

The notion of relative givenness employed here is very similar to the theory of focus in Rooth (1992b). In this theory, a focus operator '$\sim$' is used to account for focus effects. It takes an unpronounced anaphoric element as its complement and a syntactic constituent of any size as its second argument, and introduces the presupposition that there is an antecedent matching the second argument.[5]

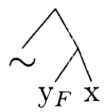
In order to determine whether there is a matching antecedent in the context, the F-marked constituents are replaced by variables, and any constituent that meets the shape of this 'template'—which is usually referred to as the 'presuppositional skeleton'—qualifies as an antecedent. The presuppositional skeletonwas orginally proposed in Jackendoff (1972), and employed in the theories of focus interpretation in Rooth (1985), Kratzer (1991) and Rooth (1992b).

The operator '$\sim$' from Rooth (1992a) essentially introduces the presupposition that there is an antecedent of the shape $[y' \ x]$, where y' is an alternative to y. For

[5]This is implemented using the focus semantic value of the second argument. I will not give the details of this analysis.

the special case in which '$\sim$' takes scope over exactly two sisters one of which is F-marked, the result is essentially the notion of relative givenness in (8) used here, where the F-marked constituent is the second argument of G_R

(11) G_R(x)(y) expressed in Rooth (1992b) terms:



The key to the focus theory proposed in Rooth (1992a) is to consider not just the F-marked constituent itself, but the entire constituent in the domain of '$\sim$', i.e. the scope of the focus. This notion of the scope of the focus has been put to use in phonological theorizing by Truckenbrodt (1995).

The modification proposed here is that I stipulate that the presuppositional skeleton needed always consists exactly of two sisters. I implement this idea however by taking the two sisters to be the argument of the operator G_F. The new proposal attempts to do what Rooth (1999, 239) suggested to do: it packages the entire theory of alternative semantics into the meaning of an operator.

The question in Rooth (1999, 239) is the following: "What function is denoted by the intonational prominence morpheme?". The answer proposed here is that the morpheme is G_F and its denotation is (8). The intonational import of G_F may either zero or purely tonal—at least it does not have any segmental content. How it effects the accentuation/deaccentuation of its first argument will be addressed below.

Sauerland (2005) proposes an even greater simplification, and gets rid of the presuppositional skeleton altogether in proposing the account based on the presupposition in (5). The evidence from sister-effects suggests that this presupposition is not strong enough, and that a minimal presuppositional skeleton is necessary. This is captured by the presupposition of RELATIVE GIVENNESS.

The approach in Schwarzschild (1999) also considers the node dominating the given constituent in the evaluation of an expression, and also operates with the presuppositional skeleton, just as Rooth (1992b). One problem is that it does not impose a requirement on the sister of the given constituent, namely that the antecedent must involve a true alternative to the sister.[6]

The definition of *relative givenness* provides a way to account for the contrast between (2) and (6). The presupposition introduced by givenness marking can be evaluated by just looking at the sister constituent.

7.1.4 What are Alternatives?

The focus theory in Rooth (1992b) and the definition of RELATIVE GIVENNESS use the notion of 'alternatives'. But what are alternatives?

It seems that individuals are usually potential alternatives to each other. Predicates are more selective in choosing their company in the alternative set. One criterion to test for alternative status is association with 'only', as was illustrated before.

For adjectives, one-replacement is also a test for alternative status. The following answer cannot be used to state the that Mary's uncle brought a any old blue convertible—it has to mean that he brought a blue high-end convertible:

(12) Mary's uncle, who produces high-end convertibles, is coming to her wedding.
I wonder what he brought as a present.
He brought a BLUE one.

I This seems to be due to the fact that 'blue' is not an alternative to 'high-end'. Compare the following example:

(13) She likes red conVERtibles, but he brought a BLUE one.

Here, the direct object in the second clause does not mean 'red blue convertible'. One-replacement affects the entire NP structure of the antecedent, but it does not necessarily include alternatives.

There are thus tests for alternatives, and there is also a little more to be said about the structure of the alternative set. One first observation (see also Rooth (1992b)) is that alternatives have to be different. Consider:

[6]Sauerland (1999) suggests a restatement of Schwarzschild's account in Roothian terms, which would actually also impose a presupposition involving alternatives, but does not explore the consequences.

(14) Mary has a blue convertible. What kind of car does John have?

a. Guess what: It's a blue CONVERTIBLE.

b. # Guess what: It's a BLUE convertible.

'Blue' is not an alternative to 'blue', and an antecedent of the shape [blue convertible] consequently does not satisfy the presupposition induced by saying [BLUE convertible]. Although every constituent in 'a blue convertible' is given, no constituent is marked as given, since no constituent is relative to some other constituent.

Another important point is that what counts as an alternative depends not just on the item in question itself but also *on the sister that is marked as given relative to it.* Consider the following two examples:

(15) a. She has a new bicycle.
Natural Partitions of bicycles that 'new' is part of:
{new, old}, {new, former}, {new, used}

b. She has a new boyfriend.
Natural Partitions defined over boyfriends that 'new' is part of:
{new, old}, {new, former}, #{new, used}

While it is certainly possible to talk about used boyfriends, it is quite marked. The reason is that {new, used} is not a natural partition of boyfriends. This is why this set does not become salient when talking about 'new boyfriends'. The joke in talking about 'used boyfriends' lies in applying a partition that is natural for objects one can possess and sell. What is a natural partition depends on what we know about the world and how we partition things in general.

'New' and 'old', and 'new' and 'used' are natural alternatives when the noun that the adjectives modify denotes a predicate that is used to refer to objects that can be possessed and sold. 'High-end' and 'red' are not part of a natural partition of cars. The correct definition of 'alternatives' goes far beyond the scope of this paper.

Jackendoff (1972, 242ff) already observed that determining alternatives (or the 'presuppositional set' in this terminology) requires conceptual structure and knowledge of the world. He states among other things that they have to be a coherent set under discussion.

The presupposition of relative givenness captures the fact that the sister constituent plays a role in determining what counts as an alternative.

7.1.5 Accommodating Givenness

Sometimes givenness marking is possible although apparently the presupposition of relative givenness is not fullfilled:

(16) What kind of convertibles does she like?
She likes BLUE convertibles.

In this context, there is no salient alternative to 'blue' that is given. But the question under discussion is about kinds of convertibles, and the answer 'blue' evokes colors as a way of partitioning convertibles. In other words, the partition into colors (and therefore alternatives to 'blue') is made salient in virtue of the use 'blue' and the fact that 'kinds' are at issue in this context. 'Blue' is used to pick out a cell in a partition of convertibles. It is not just modifying the convertible. The presupposition can be accommodated. Consider also the following example:

(17) John's aunt, who is incredibly rich and owns a bicycle factory came to his wedding. I wonder what she brought as a present.

a. Guess what: She brought a brand new BICYCLE.

b. ?# Guess what: She brought a BRAND NEW bicycle.

The answer in (b) requires there to be an alternative for 'brand new' to be salient. An obvious alternative would be 'used'. But 'used' does not seem to be pertinent here, given that the aunt is so rich, why would a 'used bicycle' be salient option here? The facts are different in the following dialogue:

(18) John's aunt, who is incredibly rich and owns a bicycle factory came to his wedding. I wonder what she brought as a present.

a. Guess what: She brought a used BICYCLE.

b. Guess what: She brought a USED bicycle.

Both answers are possible, but (b) expresses a comment on the present that (a) does not express: The answer (b) evokes an alternative to 'used'—an obvious one being 'new'. By using version (b) of the answer, the speaker insinuates that 'new bicycle' is somehow salient in this context: maybe he or he believes that being super-rich and being the owner of a bike-factory, it is cheap of the aunt to bring a *used* bicycle as a gift. The presupposition can easily be accommodated.

Marking a constituent as given does quite generally not require that a relevant alternative to the sister was actually *mentioned* in the discourse. It could be that the speaker assumes that the listener shares certain background assumptions and can thus accommodate the presupposition:

(19) Why do you think he hasn't played soccer before?
He THREW the ball into the goal.

In (19), the speaker presupposes that the listener understands that in soccer, one generally scores by *kicking* the ball into the goal. Again, it is easy to accommodate the presupposition that [kick the ball into the goal] is salient. The relevant alternative to the predicate 'throw'—'kick'—is obvious and relevant in the context.

Rooth (1992a) observes cases of 'suggested contrast'. These are again cases where unexpressed alternatives are evoked:

(20) a. in MY opinion

b. in the OLD days

c. in THIS country

It is interesting to observe that at least for (a,b), it is not so easy to think of a context in which stress would fall on the head noun. The 'marked' stress seems to be the default stress in those examples. This type of effect will be used to motivate a revision of the definition of 'neutral' stress in section 7.2.3.

7.2 Additional Arguments

Relative givenness makes two predictions: (i) Givenness marking is always relative: a constituent can be marked as given only if it is given *relative to something else*, in the way explicated by the givenness presupposition above; and (ii) the presupposition that is introduced by a constituent that is marked as given can be determined *just by looking at its sister.* Givenness is evaluated very locally. This section summarizes more evidence that bears on these predictions.

7.2.1 All-Given Contexts

There are two types of all-given contexts. The context in (21) does not license givenness marking of the VP since it does not provide an alternative subject such that [subject′ VP] is given.

(21) Last week the newspaper reported that after the game all that happened was that the coach praised John. I wonder what happened after this week's game.

a. The coach [praised JOHN]].

b. # [The COACH [praised John]].

An answer involving 'again' would be preferred in this context, but even using again does not change the observation:

(22) Again, the only thing that happened was that the coach [praised JOHN]].

Of course if 'again' follows the sentence, then the entire sentence can be subordinated relative to it.

(23) [[The coach praised John] AGAIN].

But this can be explained by the fact that [the coach praised John] can be marked as given here relative to 'again'.[7]

[7] A simple way to prevent the problem of whether or not 'again' should be used is to modify the example and say something like 'it was wrongly reported that the coach praised John'

Marking the VP 'praised John' as given as in (21b) would trigger the presupposition that there is an alternative y' to the subject 'the coach' that is salient in the discourse, and [y' praised John] is given. This is not fulfilled, hence (21b) is infelicitous. The stress pattern that emerges is the neutral one.

Consider now the following all-given context. Again every constituent including the entire 'the coach praised John' is made salient by the context. (24) provides an alternative y' to the subject such that [y' VP] is given. Consequently givenness marking is possible and hence obligatory and the VP is deaccented.[8]

(24) Last week the newspaper reported that after the game all that happened was that the coach or the manager—I forget which—-praised John. I wonder what happened after this week's game.

a. # [The coach [praised JOHN]].

b. [The COACH [praised John]].

Schwarzschild (1999, 166) discusses a related example.

(25) Did Karen get the money or did Marc get the money?

a. KAREN got the money.

b. * Karen got the MONEY.

Here, marking [got the money] as given as in (25a) is possible, and hence obligatory. There is an alternative y for 'KAREN' such that 'y got the money' is given, the presupposition of relative givenness is fulfilled.

Schwarzschild (1999) suggests that either the statement 'Karen got the money' or the entire question can in principle serve as the antecedent for deaccentuation, yielding (25a,b), but only (a) is grammatical. In order to explain that only (a) is grammatical, Schwarzschild (1999, 166) introduces a notion of question-answer congruence, which requires to use the question as the antecedent.

[8] In this context, the adverb 'again' cannot be used since based on the context, its presupposition is not satisfied.

But even so, there is a remaining problem. It is not clear why the question would not qualify as an antecedent for deaccentuation in [Karen got the x] in (b). Usually material within the question is quite readily available to license deaccentuation in the answer, even when they are yes/no questions.

(26) Did Karen get the money?
No, JOHN got the money.

Sauerland (2005), following the discussion in Schwarzschild (1999), evokes a notion of 'salience' to capture the choice in antecedence. Antecedents for givenness marking have to be salient. The added convention is that the existential presupposition introduced by the question under discussion is particularly salient.

(27) a. John [got the money]-G 'got the money' is more salient!
b. John-G got the money.

Sentence (a) presupposes for Sauerland (2005) that someone got money, while (b) presupposes that John is salient in the discourse. The former presupposition is more salient, according to the convention that the presupposition introduced by the question under discussion is very salient.

But using the presupposition of the question under discussion to arbitrate between different all-given contexts is not always possible. In the following context, once again all sub-constituents are given but givenness-marking emerges:

(28) John's aunt owns a factory that produces extremely high-end and extremely low-end bicycles. I wonder what she brought as a present to his wedding.
a. # Guess what: She brought a low end BICYCLE.
b. Guess what: She brought a LOW-END bicycle.

In the approach in Sauerland (2005), there is no way to decide between the two possibilities, since they are not distinguished by whether or not they follow from the presupposition of the question under discussion:

(29) a. # brought a [low end]-G BICYCLE.

b. brought a LOW END bicycle-G

The presupposition introduced by (b) would be that there is a proposition about a bicycle in the context and (a) presupposes that there something that is low end is the context. There seems to be no way of choosing between those options based on the presupposition in Sauerland (2005).

The proposal here predicts for (a) to require that there is an alternative x such that [low-end x] is given, which is not the case, so it is ruled out. (b) presupposes that an alternative x exists such that [x bicycle] is given, which is satisfied, since [high-end bicycle] is given.

What is at stake in the all-given examples is whether or not some constituent is given *relative to some other constituent*, just as expected based on RELATIVE GIVENNESS.

7.2.2 Givenness and Movement

A problem arises for the following case:

(30) John was in the kitchen. Mary walked in. What happened next?

a. # John kissed MARY.

b. John KISSED Mary.

'Mary' is marked as given. It would seem that a presupposition ought to be triggered that there must be a salient alternative x for 'kissed', such that 'x-ed Mary' is given. This, however, is not the case, and thus (30a) should invoke a presupposition failure.

One possibility that comes to mind is that proper names are always always deaccented, whenever the individual they refer to is given in the context. This is the prediction based in Sauerland (2005). The presupposition that is assumed there for individuals that are marked as given is that they have to have been introduced into the conversation:

(31) $[[G]] = \lambda\ x^{e} \frac{\exists i.g(i)=x}{x}$

This presupposition seems appropriate for (30). But it is not strong enough for the following scenario:

(32) What evidence do you have that Mary was involved in the burglary?

a. They [arrested John and MARY.]

b. They [arrested JOHN AND Mary.]

c. They [arrested JOHN and Mary.]

Marking 'Mary' as given in a coordinate structure introduces a presupposition that is not satisfied just by the fact that 'Mary' is a salient discourse referent in this context. Sentence (b) would be felicitous if there is an alternative x for 'and', such that 'x Mary' is given, as in the following context:

(33) Did they arrest only Mary?
No, they arrested JOHN AND Mary. (Alternative: 'only'; Given Constituent: 'only Mary')

Sentence (32b) may also be ok in a context where it is given information John was arrested:

(34) John was arrested. But what evidence do you have that Mary was involved in the burglary?
[They [arrested JOHN AND Mary.]]

Sentence (32c) would require that there is an alternative x to 'John', such that 'x and Mary' is given, as in (35):

(35) Did they arrest Bill and Mary?
No, they arrested JOHN and Mary. (Alternative: Bill; Given Constituent: 'Bill and Mary')

We are faced with a situation where the weak presupposition proposed in Sauerland (2005) is appropriate for (30), and the strong presupposition of *relative givenness* is appropriate for (32). How can we resolve this conundrum?

The solution that I want to propose for this problem is the following: The DP in (30) can move, and in particular it moves to adjoin to a higher constituent, in order to facilitate givenness marking. The idea is that 'Mary' is really adjoined to the entire proposition:

(36) [[λx. [He kissed x.]] Mary]

Then 'Mary' is marked as given, but its sister is the entire proposition. It therefore does not have to be marked given relative to the predicate 'arrest', but it is marked as given relative to the entire proposition. This type of movement thus weakens the presupposition associated with givenness marking.

A set of the shape [x Mary] is salient in this context, where x stands in for propositions with an open argument slot. It is made salient by having introduced the referent Mary into the conversation. I assume that the set is a subset of the question under discussion. The set includes every proposition that contains Mary as an argument.

The explanation that the proper name in (32) does not get marked as given is that movement out of coordinate structures is not possible. Givenness must henceforth be evaluated locally. The givenness presupposition introduced (that there be an alternative to 'and') is too strong . There is thus an interaction between the strength of the presupposition introduced by a constituent marked as given and island effects.

The kind of movement relevant here might be the same that is involved in right-node raising. RNR also differentiates between the two examples:

(37) a. The Sheriff arrested—, but the Judge acquitted— Mary.

b. * They arrested John and— but convicted Bill and— Mary.

That information structure can re-bracket a syntactic subject is also assumed in Steedman (2000). The present approach motivates the rebracketing: it facilitates givenness marking by weakening the presupposition.[9]

[9]The principle 'Maximize Presupposition' would then create the following expectation: when the stronger presupposition is satisfied, givenness movement should not occur. This seems to be

The Movement account explains the puzzle of why DPs referring to contextually given individuals sometimes deaccent and sometimes do not. They can only deaccent if they are 'relatively given' to their sister constituent—and what the sister constituent can be adjusted by movement. An example discussed by Ladd (1080, 81) further illustrates the point:

(38) Has John read Slaughterhouse-Five?
He doesn't réad books.

Here, there was talk about books in the context. Once 'books' moves and attaches to the proposition, givenness marking is possible and hence obligatory. Note that every constituent in this example is given. Still, it the direct object that can be marked as given relative to the entire proposition, since the contexts makes 'books' and propositions about them salient. But now consider the following example:

(39) Has John read Slaughterhouse-Five?
He doesn't read Poetry or NOVELS.

While 'novels' is made salient in this context, poetry is not, so 'poetry or novels' as a whole are not given.

This approach to givenness and movement can rationalize the pervasive cross-linguistic tendency that given constituents undergo movement[10] (e.g. object shift, pronoun-movement to 'C'). These operations create new possibilities for givenness-marking.

7.2.3 Inherently Unstressed Elements

A resilient problem in accounting for sentence stress is that some items seem to resist nuclear stress. An example is the word 'something' in English (Bresnan, 1972).

(40) a. Helen had WRITTEN something.

incorrect. Given material (pronouns, given DPs) moves whenever possible. I will not discuss this problem here.

[10]Or the equivalent rebracketing in the grammatical theory of choice.

b. Helen had written a BOOK.

Bolinger (1972) argues that accents fall on 'newsworthy' constituents, and claims 'something' is generally not accented because it is not newsworthy (40). But what is more newsworthy about 'some food' compared to 'something' in (41)?

(41) Ad Neeleman (p.c.):

a. She ate some FOOD.

b. She ATE something

Bolinger's appeal to 'newsworthyness' is circular, unless criteria are specified to assess whether a constituent is newsworthy or not independent of accentedness.

The present proposal explains the difference in accent placement in (41). The direct object can move, and the presuppostions that are then derived are in (42).

(42) a. There is an alternative proposition in the context about some food.

b. There is an alternative proposition in the context about something.

Presupposition (42b) is trivially satisfied, so 'something' can always be marked as given; presupposition (42a) is only satisfied when there was talk about food, which is only the case in the appropriate context.

The proposal here accounts for why 'someone' and 'somewhere' etc. are 'inherently destressed'. They can be marked as given in contexts although there is no antecedent for givenness marking. This is due to the fact that the presupposition that is introduced by marking them as given is trivial. In explaining the contrast in (41), the present theory goes beyond Bolinger's approach, while capturing the intuition that marking 'something' is related to the poverty of lexical content that makes it easy to accommodate the presupposition.

Other items that seem to be inherently unstressed are indexicals such as 'here', 'I', 'you', 'now'. This is not surprising, since in any utterance a speaker, and addressee, a location, and a reference time can be taken as given, since the presupposition that marking them as given are easy to accommodate.

A similar rationale can be made for other examples that Bolinger (1972) discusses. They also invoke presuppositions that can be accommodated:[11]

(43) Bolinger (1972, 636–637)

a. I'm going to the dóctor's place. vs. I'm going to the doctor's bárn.

b. I'm doing it for Jóhn's sake. vs. I'm doing it for Jòhn's wélfare.

Bresnan (1971) suggests that these items are semi-pronouns. As Bolinger (1972) observes, this approach runs the risk of being circular, at least if this label is used as a mere diacritc for not receiving stress. Again, the presupposition that marking those items as given introduces may simply be easy to accomodate.

In outlining the data on prosody I often made use of a notion of 'neutral stress' in order to refer to the prosodic of a sentence that does not impose restrictions on the context. While a definition of 'neutral stress' does not play any role in the official theory, it is a useful notion when speaking about the relation between syntax and prosody controlling for information structure. The discussion in this section suggests a revision of the definition of 'neutral stress', repeated here:

(9) Neutral Stress

A stress pattern is neutral if it does not trigger a presupposition of relative givenness for any sub-constituent.

The sentence in (40a) is 'neutral', in that stressing the direct object would require a much more restricted context than not accenting it—but it is not neutral according to (9), since it triggers a givenness presupposition. A context in which 'something' is accented is the following:

[11]Some of the items that can be deaccented can even be omitted:

(1) I'm going to the dóctor's. vs. I'm going to the doctor's bárn.

Clearly, these are recoverable in this context. There are in fact approaches that try to relate (de-)accentuation directly to predictability or expectedness, e.g. Jurafsky et al. (2000). The idea here is compatible a role of predicability, but what has to be predictable is not just 'place', but also the presupposition of relative givenness introduced by marking it as given.

(44) What did you say she wrote? A novel? Helen had written SOMETHING. But I don't know what.

Clearly, it requires a rather specific context to yield an accent on 'something'. According to Höhle (1982), one can define a 'neutral' stress pattern as the pattern that is compatible with *most* contexts. We can capture this intuition by the following refinement of the original definition of 'neutral' in (9):

(45) NEUTRAL STRESS

A stress pattern is neutral if it only triggers relative givenness presuppositions that are easily accommodated in most contexts.

The fact that 'something' is an 'unstressed' item is explained by the same mechanism that explains deaccenting more generally.

A further prediction is then that in languages that lack givenness deaccentuation (e.g. Spanish, Italian), the equivalent of 'something' should not be deaccented. This prediction is borne out by the facts (cf. Ladd (1996) and references therein). Italian lacks deaccentuation of given material, as is illustrated by the following contrast between an Italian sentence and a close English counterpart:

(46) a. le inchieste servono a mettere a posto cose andate fuori pósto

b. the investigations help to put back in place things that have gone óut of place.

And it also lacks the property of having unaccented 'something' etc.:

(47) a. Ho sentito qualcúno.

b. I héard someone.

'Relative Givenness' can account for why items like 'something' are generally unstressed in English, by the very mechanism that explains givenness deaccentuation more generally. Moreover, it correctly predicts that languages without givenness deaccentuation put stress on items such as 'something'.[12]

[12]The revised definition of neutral stress also does justice to the fact that sometimes marked stress seems to be the default. Consider again some of the examples that Rooth (1992a) reports as

7.2.4 Interaction with Other Presuppositions

Some lexical items tend to be prosodically subordinated, and yet the presupposition that they introduce is not easy to accommodate in any context. Consider pronouns such as 'him' or 'there'.

(48) The princess KISSED him there.

When given these sentences out of context, native speakers are likely to deaccent the pronouns. But out of context, it is not clear what 'him' and 'there' are intended to refer to, and hence there is not other salient proposition about 'him' or 'there' that is salient.

These sentences cannot be used unless the context satisfies certain presuppositions that are introduced by the lexical items 'him' and 'there'. They require there to be a unique salient discourse participant that is male or that there is a unique salient location. A speaker confronted with these sentences without a context that satisfies these presuppositions will either try to accommodate (or make up) a context that satisfies the presuppositions of the pronouns or reject the sentence.

Say one accommodates a context in which the presuppositions of the pronouns are satisfied and there is a unique referent for 'him' an 'there'. Then the givenness presupposition needed to mark them as given in (48) is satisfied, since there must be some other statement about 'him' and 'there' that is salient. The pronouns thus seem to be inherently unstressed because the presuppositions that allow their use usually entail that the presupposition of marking them as given is fullfilled.

The prediction is of course that similar to the deaccentuation of proper names, the givenness presupposition introduced by subordinating pronouns should depend on the

---

involving 'suggested contrasts':

(1) a. in MY opinion ...as opposed to somebody else's opinion that is at issue.

b. in the OLD days ...as opposed to nowadays.

These expressions are usually used to express the contrast that is made explicit after the examples above, and in fact they sound odd with main stress on the head noun.

syntactic context. For example, if a pronoun is embedded in a coordinate structure, then the presupposition of relative givenness is not necessarily satisfied just by virtue for the fact that the presuppositions of the pronoun are satisfied. This is borne out by the facts. Consider the following example, which shows the same pattern as the one involving the proper name:

(49) What evidence do you have that Mary was involved in the burglary?

a. They [arrested John and HER.]

b. They [arrested JOHN AND her.]

c. They [arrested JOHN and her.]

The treatment of pronouns is then in no way different from the treatment of full DPs. In fact, their presuppositions can interact with the presupposition of relative givenness in a similar way as the in the case of pronouns. Consider the following dialogue:

(50) What happened next? The princess kissed the frog.

This answer presupposes that there is a unique frog. The answer suggests then that there was talk about a frog in previous discourse. When confronted with this dialogue without further context, it is not unlikely that a speaker who is asked to read it might deaccent 'the frog', unless the sentence is understood as a way of expressing that the princess kissed the frog as opposed to somebody else.

But not every definite DP is going to have this property of placing such a restriction on the context. Consider the following discourse:

(51) Why was there an uproar in the palace?
The king slapped the queen.

Given that the scenario is the palace, it is likely that if we introduce a king at all that it would be a unique king and if we introduce a queen at all that there is a unique queen, because palaces usually house only one king and one queen, if they house any at all. The use of the indefinite article seems odd in this case since it would suggest that there are several kings and queens around. Here, using the definite article thus does not necessarily require a context in which either king or queen have been mentioned.

The relation between definiteness and givenness is an indirect one then. The presupposition of uniqueness is not necessarily satisfied by previous mention in the immediate context but can be shared by world knowledge that can be taken as part of the common ground. Some DPs do not even require a particular scenario are those where context-independent world knowledge tells us that uniqueness is satisfied, such as 'the sun' or 'the president of the United States'. This issue is further discussed in the context of scrambling in chapter 8 and Wagner (2002b).

7.2.5 Relative Clauses and Relative Givenness

Relative clauses are often prosodically subordinated. This was observed i.a. in Bresnan (1971, 258):

(52) Mary liked the propósal that George left.

This can be accounted for if the relative clause is a one-place function derived by movement:

(53) λx. that George left x

It subordinates just like other functors that follow their complement. Nuclear stress, that is the last accent, falls then on the head of the relative clause. But this, as was critically observed in Berman and Szamosi (1972) and Lakoff (1972), is not always the case.

Bresnan (1972, 337) illustrates the problem with the following two examples, which according to Stockwell, when presented out of context, differ in how they tend to be pronounced:

(54) a. Introduce me to the man you were tálking about.

b. I'll lend you that bóok I was talking about.

According to Bresnan, there is also a semantic difference between the relative clauses: if the relative clause receives nuclear stress, then it is used to pick out one out of a set of alternatives defined by the head noun. In this case the head of the relative clause is what she calls a 'concealed partitive'. In the sentence with nuclear stress on the head noun, however, there is no such partitive reading.

We can make sense of this intuition with the notion of relative givenness. Not subordinating the relative clause amounts to marking the head of the relative clause as being given. The presupposition is then that there must be an alternative its sister, the relative clause, such that head and relative clause are given in the context. In other words, there must be an alternative restrictive relative clause that picks out a different individual in the extension of the predicate that heads the relative clause.

In the example above, this means that there must be some other man salient which could have been picked out by a salient different function. The requirement is then that the salient set of men contains at least one other men, and this arguably what Bresnan's intuition of being a 'concealed partitive' is about.

7.3 Givenness and Prosodic Subordination

This section makes explicit how givenness marking is implemented. The new observation is that the prosodic subordination due to givenness shows a close resemblance to the prosodic subordination of functors. The proposal is then that the two phenomena are actually part of the same generalization, and the same principle accounts for both effects, and predicts that languages which have one also should have the other.[13]

7.3.1 Givenness and Prosodic Asymmetry

As discussed in chapter 6, functors are often not accented, although they constitute new information (55a). Non-accentuation of the complement is usually infelicitous,

[13]The notion prosodic subordination is used in this chapter to account for the facts that are often discussed under the label 'post-focus reduction' 'citepIshih03. The idea is that it is the same theoretical object as the prosodic subordination discussed in chapter 6.

unless it is given (55b).

(55) After Mary arrived, what happened next?

a. She praised JOHN.

b. # She PRAISED John.

Linear order matters. When a functor *precedes* its complement, it can bear an accent:

(56) After Mary came, what happened next?

a. [She [praised JOHN.]]

b. [She [PRAISED JOHN]]

The German OV-word-order sharply contrasts with the English VO-word-order with respect to this optional accent-placement. In German, it is impossible to place an accent on the verb in a wide-focus context:

(57) After Mary arrived, what happened next?

a. [Sie [hat [HANS gelobt]]].

b. # [She [hat [HANS geLOBT]]].

The contrast is due to the different word order, and also emerges in English in OV-contexts:[14]

(58) Did she request any changes in the new office?

a. She wanted to have some WALLS painted
She wanted to have the some walls PAINTED
(unless there was talk about walls)

Conversely, German VO patterns like English VO and allows for an an accent on the predicate:

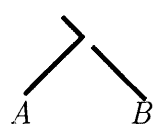
[14]A more thorough discussion of OV contexts in English would be necessary to illustrate that they indeed pattern just like German OV contexts. This is beyond the scope of the this chapter. Compare the discussion in chapter 6.

(59) a. Sie lobte HANS.
she praised Hans

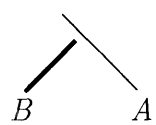
b. Sie LOBTE HANS.
she praised Hans

This asymmetry is discussed in more detail in chapter 6. It can be summarized as follows:

(60) PROSODIC ASYMMETRY

a. 

When a functor A *precedes* its complement B, the functor may be on a par with its argument or may be prosodically subordinated: [Á B́]; [A B́].

b. 

When a functor A *follows* the complement B , A is prosodically subordinated: B́ A (unless A is focused or B is given).

Consider now the case where the direct object encodes given information in English. Placing an accent on 'JOHN' is infelicitous in this context:

(61) Mary sat in her office. John walked in. what happened next?

a. # She praised JOHN.

b. She PRAISED John/him.

Again, if we compare this to German, an interesting asymmetry emerges:

(62) Mary was in her office. John walked in. What happened then?

a.	[She [práised John.]]	a.	[Sie [hat [Hans gelóbt]]].
b.	# [She [praised Jóhn]].	b.	[Sie [hat [Háns gelóbt]]].

In English, it is impossible to place an accent on the direct object in this context. It would be the last accent, and hence be more prominent were it accented. In German, however, it is possible to put an accent on the direct object when it is given.

Similar asymmetries are observed with adjectives. When the head noun can be marked as given relative to the adjective, then it cannot carry an accent, but when conversely the adjective can be marked as given relative to the head noun, it can carry an accent. Experimental evidence illustrating this asymmetry is reported in Ito et al. (2004).

What is striking about the asymmetry in givenness marking is that it looks like the exact mirror image of the asymmetry observed for neutral stress. In VO-structures an additional accent is impossible, and in OV-structures it is. What unites the two cases is that an optional accent is possible *exactly when it would not be the most prominent accent.*

Another parallel between subordination in functor-argument sequences and in givenness marking regards so called 'function words'. Both functors and material that tends to encode information that is 'given', e.g. pronouns, by virtue of the presuppositions they impose on the contexts they can be used in. These items tend to develop weak allomorphs, a property they share with functors. The unifying factor is that both types of elements undergo prosodic subordination.[15]

The development of weak allomorphs is of course another parallel between functors and given constituents. Function words subsume those two classes of words.

7.3.2 Prosodic Subordination

The prosodic asymmetry motivates the distinction of two ways of how the prosodies of two sisters can combine.

(63) Summary of the Syntax–Phonology Mapping:

[15]An interesting questions arises in how to code weak pronouns and clitics ((cf. Cardinalettie and Starke, 1994)) in the present system. They are obligatorily reduced and cannot occur in contexts where they would not be subordinated, such as in coordination structures. Similarly, there are completely stressless allomorphs of some functors that cannot be used in cases where the functor has to carry some stress. I will not go into the details of how this could be made to work.

a. PROSODIC MATCHING

Match A and B in their prosodic status if A precedes B: [Á B́]

b. PROSODIC SUBORDINATION

Subordinate A to B: [B́ A]

One principle that negotiates which of the two options applies is COMPLEMENT PROMINENCE from chapter 6:

(64) COMPLEMENT PROMINENCE

If A and B are sisters and A is the functor and B its argument, then B is more prominent than A unless A already contains an argument (i.e. is branching in the sense of BPR).

These principles apply recursively to the nodes of a tree. Thus they derive a grid for a tree of any given complexity, with a potentially infinite number of lines on the grid. Subordination effectively affixes the subordinated constituent to the preceding prosodic constituent. The following assumption is what allows us to test the system with respect to accent placement:

(65) Top Line grid Marks are realized as Accents.

The proposal is a version of the approaches in the tradition of the 'transformational cycle', (e.g. Chomsky and Halle, 1968). The idea is that relative prosody is decided locally between sisters. The item bearing nuclear stress is simply the last unsubordinated element in an utterance.

The difference is that nuclear stress is considered to be a derivative notion: the nuclear stress is simply the last accented constituent. The generalization at the heart of the observed asymmetry is then that arguments have to be more prominent than functors. Since it is the last accent that is perceived as more prominent, this forces subordination in the case of the OV-word order but not in the case of the VO-word order.

Apart from the functor-argument relation, givenness also affects the choice of whether two cycles are related by matching or by subordination. The generalization again can be captured by using the derived prominence and adding a second principle:

(66) a. A Given constituent is less prominent than the constituent relative to which it is marked given.

b. Functors are less Prominent than Arguments.

We can consider these constraints on the output of cycles, and the correct pattern is derived if the givenness-related constraint outranks the functor-related constraint. This is not the only way to think abou the interaction of givenness-subordination and functor-subordination. I will return to this problem in the next section.

These two principles have the effect that subordination will be obligatory when (i) a given constituent follows the constituent it is marked given relative to or (ii) a functor follows its argument, and the functor is not given relative to the argument.

These constraints capture the pattern of givenness marking observed above: given constituents prosodically subordinate when they follow the constituent relative to which they are marked as given, but they may or may not subordinate when they precede it.

7.4 A Unification?

There is a way to unite the two asymmetries. If the given constituent was the functor, and it's sister the argument, then givenness-deaccenting could be viewed as another instance of functor-subordination.

7.4.1 The Givenness-Operator

Consider again the givenness operator G_R. The operator takes the given constituent as its complement, and takes it's sister as a second argument. It's denotation is repeated here:

(8) Relative Givenness

$$[[G_R]] = \lambda x.\lambda y.\frac{\exists y' \in Alt(y):[[y'x]] is\ given}{[[yx]]}$$

Similar to 'only', the givenness operator is a two-place function. The first argument is the given constituent, the second argument is the constituent relative to which it is marked as given. In cases where A is the functor, and B is the predicate, it leaves the functor-argument relation constant:

(67) Who did she praise?
She [[G_R(praised)] (John)]

But when the given constituent is the right-hand sister, then it changes the functor-argument relation compared to the neutral case:

(68) What did she do to Hans?
She [(praised) [G_R(Hans)]].

We can think of G_R as being an affix on the given constituent. In some sense, G_R(Hans) thus becomes a functor, taking an argument. Since it follows its complement, prosodic subordination applies and the given constituent is obligatorily realized with a reduced pitch range.

This approach on givenness is similar to that in Steedman (2001) in that it allows us to change the direction of composition for information structure reasons.[16]

7.4.2 Languages without Subordination

Functors subordinate when they follow their complement. A given constituent forms a functor with the operator G_R that takes its sister as a complement. If givenness-subordination and predicate subordination are really part of the same generalization, then we expect to find more evidence for a correlation of the two processes.

[16]The operator 'G_R' may well have a phonological output, but one that only affects tonology, and without segmental content. But see Calhoun (2004) and Calhoun (2005) for a critical discussion of claims about the tonal difference between theme and rheme pitch accents.

In section 7.2.3 I reported an observation from Ladd (1996) that Italian lacks at least certain cases of deaccentuation of given material:

(69) a. le inchieste servono a mettere a posto cose andate fuori pósto

b. the investigations help to put back in order things that have gone óut of place.

Italian also lacks the property of having unaccented 'something', a fact that was derived from the same principle that is responsible for givenness deaccentuation:

(70) a. Ho sentito qualcúno.

b. I héard someone.

But Italian also lacks predicate subordination:

(71) a. Ho un libro da leggérre.

b. I have a bóok to read.

These cases might be due to the same property of Italian—the lack of prosodic subordination of functors.

It is not clear though what exactly is different about Italian and English such that it has the effect that prosodic subordination does not apply in configurations where it applies in English. One possibility is that there are syntactic differences that yet have to be discovered. Another is that the mapping principle of prosodic subordination is simply missing.

The typology is more complicated however. Ladd (1996) reports that under some circumstances, there is givenness-marking even in Itallian. Also, in Italian, there is prosodic subordination in cases that are often called 'contrastive focus' or 'exhaustive focus'. Since the approach to givenness proposed for Germanic here subsumes cases typically considered 'contrastive' and cases considered to be due to 'givenness', it is not clear that unification proposed here is rich enough express the subtle cross-linguistic differences.

Ronat (1982, 34) reports clear cases of givenness-accenting for French:

(72) Paul regarde les informations tous les soirs; Marie est jalouse de la télévison.

The phrase 'de la télévision' can be deaccented, just as in the corresponding English sentence:

(73) Paul watches the news every night; Mary is jéalous of the televison.

Zubizarreta (1998, 75) shows that French does not have predicate subordination, as opposed to English:

(74) a. J'ai un travail a FAIRE.

b. I have a JOB to do.

It seems then that subordination with due to givenness and subordination of functors are separate phenomena. I am not aware though of a language that has predicate subordination but no givenness-subordination. The precise relation between the two phenomena remains to be discovered.

Givenness Marking vs. Givenness

This chapter outlined the generalization about when a constituent can be marked as given. It can be captured by the notion of *relative givenness*. Givenness marking itself can be seen as another instance of prosodic subordination.

In some early accounts of sentence stress information structural effects were modeled by a shift or overwriting of the syntactically determined 'default' stress (Chomsky, 1971, Stockwell, 1960, 1972). Other approaches proposed the *prior* assignment of major prominence to the verb blocking the application of the syntax sensitive stress rules that would otherwise apply (Newman, 1946, Bresnan, 1972).

In the present system, nothing special has to be said about the mapping to prosody. The effect of information structure is indirect, and can be explained in reference to semantics of givenness marking and the shift in functor-argument relations.

Sometimes, constituents that are given cannot be marked as given in this technical sense because the strong presupposition of relative givenness is not satisfied. It may well be that they are realized slightly reduced or faster, since they are contextually salient. Gradient effects have indeed been reported in the literature.

Bard and Aylett (1999) and Bard et al. (2000)) for example report that in addition to deaccentuation, there is at least one other process that can affect given material even when it is not deaccented, which is reducing token length and decreasing articulatory detail, and plausibly relates to the accessibility of antecedents in the discourse. The assumption here is that while these additional gradient effects of accessibility differ from prosodic subordination in that they are not categorical and that they do not introduce a semantic presupposition.

The grammatical marking of a constituent as given encodes a presupposition that imposes a restriction on what the speaker considers to be the common ground at this point in the conversation (Stalnaker, 1978). A successful givenness marking thus takes into consideration the listener's knowledge. The gradient reductions reported in Bard et al. (2000), however, occurred even when the antecedent of the constituent in question was only accessible to the speaker but not to the listener.

The empirical prediction is then the following: Absence of presence of gradient reduction due to salience, likelihood, or expectedness should not affect the felicity of an utterance given a context, but at most influence intelligibility.[17] Failure to mark presuppositions that are satisfied or introducing presuppositions that are not satisfied affect the felicity of an utterance. RELATIVE GIVENNESS is about the presence/absence of presuppositions and felicity conditions on sentences in context. It is not about gradient effects.[18]

[17]In an experiment reported in Welby (2003) gradient acceptability judgements are reported. One source of gradience in cases in which the presuppositions that the intonation encodes differ is how hard it is to accommodate the presupposition that is encoded.

[18]This chapter did not discuss cases of multiple focus, and of different kinds of foci. A discussion of different kinds of focus and how they are realized when they cooccur is in Selkirk and Kratzer (2005). These are beyond the scope of this chapter, but potentially pose a problem for the approach I am taking, which essentially conflates contrastive focus and information focus. All further distinc-

7.5 Appendix: Defining Givenness

We still need to define what it means to be 'given'. As is noted already in Rochemont (1986, 46) and Schwarzschild (1999), it is somewhat easier to define what is 'given' than what is 'new', and thus they define 'focus' as that which does not constitute given information. A precise formal definition of what it means to be given, however, is not trivial. So far the notion 'given' was used at an intuitive level. For example, in which sense is 'convertible' in the following example given?

(75) John drove Mary's red convertible. What did he drive before that?
He drove [her [BLUE] convertible].

Williams (1980) proposes to reduce the notion 'given' to entailment. Intuitively, something counts as given if it is entailed by previous discourse. The trouble in using entailment to test whether a constituent counts as 'given' is that not every constituent has a truth value.

The proposal in Williams (1980) is to address this problem by existentially quantifying free variables of a constituent, and thus lifting it to type t, and then checking whether this new existential proposition is entailed by the discourse. In (75), the constituent 'convertible' would count as given since its existential closure '$\exists$ x. x is a convertible' is entailed by the previous utterance. But in the following case, it is not clear that the context licenses the inference to the existence of a convertible:

(76) There are no red convertibles.
Then I want [a BLUE convertible.]]

Schwarzschild (1999) make the entailment-based notion of givenness formally explicit. He follows Rochemont (1986, 47) that a given constituent has to have an antecedent that entails it. An operation of type shifting is introduced that raises constituents to type t by existentially binding unfilled arguments, following Williams (1980). This can be done both with the constituent under evaluation and with the antecedent in the context. In the case of 'convertible':

tion, under the view presented here, must be due to overt or covert operators (such as exclusives, exhaustive operators, etc.) that are represented in the syntax.

(77) $\exists$ x.red-convertible (x) *entails* $\exists$ x. convertible (x)

This notion of entailment is able to explain why 'convertible' in (75) counts as given if we assume the following notion of givenness (Schwarzschild, 1999, 147) (cf. Rochemont (1986, 47) who employs the notion 'c-construable' instead of entailment):

(78) Givenness (first version)

An utterance U is *Given* iff it has an antecedent A and A entails U, modulo $\exists$-type shifting.

Existential type shifting can be made more explicit as follows:

(79) *Existential Type Shift: ExClo* (Schwarzschild, 1999, 152)

a. If $\omega \in D_t$, then $ExClo(\omega) = \omega$

b. For any conjoinable type $< a, b >$:
If $\omega \in D_{<a,b>}$, then $ExClo(\omega) = \lambda w.\exists u \in D_a[Exclo(\omega(u))(w)]$

c. t is a conjoinable type.
if b is a conjoinable type, then so is $< a, b >$, for any type a.

The definition of givenness so far does not cover constituent of type e. Individuals do not have unfilled arguments, and thus cannot be type-shifted to a constituent of type t by existential closure.[19]

(80) Givenness (final version) (adapted from Schwarzschild, 1999, 151)

An utterance U counts as *Given* iff it has a salient antecedent A and

a. if U is type e, then A and U corefer

b. otherwise: $ExClo(A)$ entails $ExClo(U)$.

This definition of givenness invokes entailment relations that are only computed for the sake of evaluating givenness. The presupposition of relative giveness uses this

[19]Schwarzschild (1999) assumes a second step in the definition of givenness, the existential F-closure. This step consists replacing any F-marked terminal node with a variable, that is then existentially bound.

notion of givenness, but the requirement makes reference to both the constituent that is marked as given and its sister.[20]

[20]Schwarzschild (1999)'s proposal always evaluates givenness of utterances, while the proposal in this chaptercyclically evaluates givenness between two sister constituents. Otherwise, the evaluation of givenness is the same. What is different is not the definition of givenness but the presupposition that is introduced by prosodically marking something as given.

Chapter 8

Conditions on Accent Omission in Alternative Approaches

Abstract

The proposals on prosodic subordination presented in chapter 6 and chapter 7 are compared to alternative approaches: The F-Projection account in Selkirk (1984, et. seq.) and the reference set approach in Reinhart (1995, et. seq.). This chapterdiscusses some of the problems of the alternatives. While the present proposal incorporates important features of both theories, but it covers a wider range of data.

8.1 Conditions on Accent Ommission

The distribution of accents in a sentence is sensitive to grammatical notions such as 'head' vs. 'complement', but also to information-structural notions such as 'new' vs. 'given', and to question-answer congruence.

In the approach to sentence stress in SPE there is a stress pattern that is derived that is used in 'neutral' circumstances. Any diversion from the neutral pattern, according to Chomsky (1971) and Jackendoff (1972), requires that there must be narrow focus on the constituent carrying main stress. The approach presented in chapter 6 and chapter 7 makes similar assumptions although the details of the implementation differ.

One aspect that is quite different is that under the approach taken here a diversion from the neutral stress pattern can still be compatible with wide focus, while in Chomsky (1971) and Jackendoff (1972) non-neutral patterns are not compatible with wide-focus. One motivation for this difference is that it is often sufficient for material to be 'given' or 'thematic' in order for it to be destressed, while the resulting prosody is still compatible with a wide-focus question. This observation is discussed in Kiparsky (1966), Halliday (1967), Ladd (1080), Höhle (1982), Selkirk (1984), Gussenhoven (1984), and many other studies.

The main issues according to the recent literature on the issue (e.g. Selkirk, 1995b, Schwarzschild, 1999, Williams, 1997) are the following: (i) under which conditions are heads non-accented ('Head-Effect'); (ii) under which conditions is given material non-accented ('Givenness-Effect'); (iii) under which conditions is material that is not focused non-accented ('Focus-Effect').[1]

In chapter 7, I discussed the approach to deaccentuation in terms of givenness-marking (Sauerland, 2005, Schwarzschild, 1999). In the following, I will discuss two alternative influential approaches on the conditions on deaccenting: First, the approach in terms of focus projection Selkirk (1984, et. seq.), and second the reference-

[1]Along with Rooth (1992b), Williams (1997), and Schwarzschild (1999) I assume that (ii) and (iii) are two aspects of the same effect.

set approach proposed in Reinhart (1995, et. seq.). I will present evidence that these approaches do not account for the same range of cases as the approach presented in chapter 6 and chapter 7.

8.2 F-Marking

Accenting certain constituents allows for placing focus on much bigger constituents that contain them. This phenomenon was named 'Focus Projection' in Höhle (1982, 99). Selkirk (1984) developed a formal approach to model this phenomenon, which was updated in Selkirk (1995b). The theory uses the tool of F-marking to state conditions on accent placement and capture F-projection. F-markers were originally introduced to explain focus related phenomena such as association with focus (e.g. in Jackendoff, 1972, Rooth, 1985). F-marking was also used to explain why placing accents in a certain place (e.g. the direct object) is often compatible with many different foci (e.g. object focus or VP-focus) (Chomsky, 1971, Jackendoff, 1972).

The theory of F-marking as it is outlined in Selkirk (1995b) tries to account for the distribution of accents and seeks to explain the Head-Effect, the Givenness-Effect, and the Focus-Effect at the same time. F-markers are taken to be diacritics that are part of the syntactic phrase marker. They make it possible to state generalizations about accent distribution. The focus of a sentence is defined as an F-marked node that is not dominated by another F-mark:

(81) FOCUS (Selkirk, 1995b, 555)
The FOCUS of a sentence is defined as an F-marked constituent not dominated by any other F-marked constituent.

One assumption is that the constituent in the answer corresponding to the wh-part of the question that is at issue has to be a FOCUS, that it has to carry an undominated F-marker.[2]

[2]In the exposition in Schwarzschild (1999) and Büring (to appear) both state that F-marking would be sufficient, but it seems clear from the discussion in Selkirk (1995b) that a FOCUS is required.

(82) Focus Condition

An appropriate answer to wh-question must have an undominated F-mark on the constituent corresponding to the wh-phrase.

In an answer to a question such as 'What happened next?', the root of the sentence has to be a FOCUS, hence it has to be F-marked.

F-markers cannot be freely assigned to syntactic nodes. There are entailment relations: placing an F-marker on an unaccented head entails that an F-marker is placed on its internal argument. This is called 'licensing' in the terminology of F-marking.

Every accented lexical item carries an F-mark. This is the 'Basic Focus Rule'. These F-markers can then be projected to higher nodes in the tree. The conventions for F-projection are outlined below (I adopt the terminology 'vertical' and 'horizontal' F-projection' from Büring (to appear)):

(83) F-Marking (First Version) (cf. Selkirk, 1995b, 555)
Basic Focus Rule: An accented word is F-marked.

F-Projection:

a. Vertical F-Projection: F-marking of the head of a phase licenses the F-marking of the phrase.

b. Horizontal F-Projection: F-marking of an internal argument of a head licenses the F-marking of the head.

These two types of projection will be illustrated below. F-marking was employed in Selkirk (1995b) to explain the Head-Effect and the Givenness-Effect.

8.2.1 The Head-Effect

The example in (84) involves only one accent, the one on the direct object. The first question that arises is why the verb can be non-accented, while this is impossible for the direct object:[3]

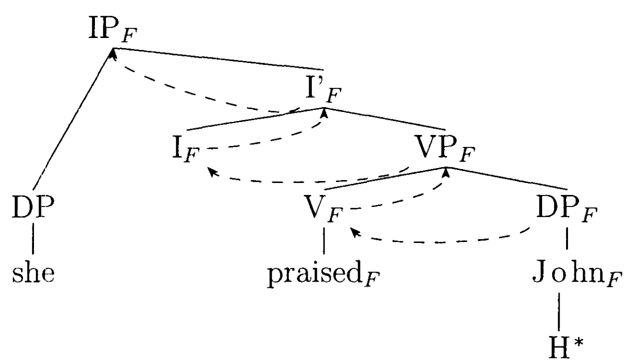
(84) After Mary arrived, what happened next?

a. [She [praised$_F$ JOHN$_F$]$_F$]$_F$.

b. # [She [PRAISED$_F$ John]$_F$]$_F$.

The F-marking of the examples in (84) show an F-marker on the root node, as is required in a context with wide-focus, and obeys the projection conventions. To illustrate how F-marking explains the omissibility of the accent on the verb, consider the way F-markers project (84a):

(85) F-Projection



The root of the tree can be F-marked, without placing an accent on the predicate, as is required by the Focus-Condition. But based on the Focus-Condition, the example in (84b) should also be acceptable, since placing an accent on the verb allows projection to the root node. The example is infelicitous because one constituent has no F-marker, despite the fact that it does not encode given information. This problem will be discussed in the next subsection.

The approach to accent distribution using F-marking is able to explain some generalizations concerning nuclear stress. Nuclear stress is final in English. In fact, in most phrases main prominence falls on the last constituent, an observation that prompted Chomsky and Halle (1968) to posit a nuclear stress rule that assigns highest

[3]I assume the same is true for a subject that is not given. The pronoun used here of course does not get an accent, but also it need to be co-referent with an individual that has been introduced into the conversation.

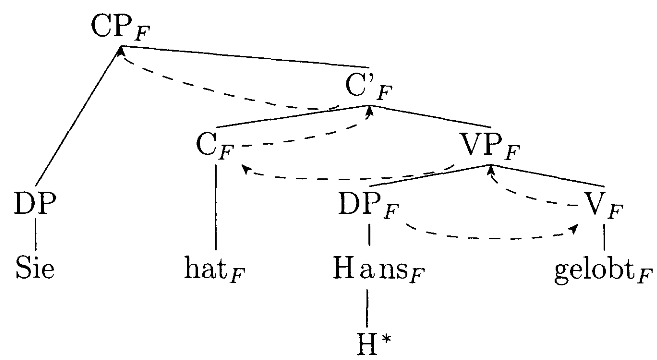
prominence to the rightmost element in a phrase. But this is not the case in German in OV environments: nuclear stress falls on the direct object (based on example (86)):

(86) After Mary arrived, what happened next?

[Sie [hat_F [HANS_F gelobt_F $]_F$ $]_F$ $]_F$.
she has Hans praised

F-projections works as follows:

(87) F-Projection: The case of German



The F-marking approach applies to *both* English and German, without reference to generalizations based on linear order such as 'leftmost' or 'rightmost' in placing the accent. Further evidence that this is superior to a treatment based on a phonological rule as in SPE ('assign main prominence to the rightmost constituent') is that English behaves just like German when a predicate follows its internal argument. In the following example, 'painted' does not receive an accent unless 'walls' constitutes given information:[4]

(88) She wanted to have [the WALLS_F painted_F $]_F$.

Horizontal F-projection (83b) is crucially involved in explaining the non-accentuation of predicates in general, and in explaining the distribution of nuclear stress in English and German in particular.

F-marking is assumed to be part of the syntactic representation in Selkirk (1984, 1995b). The conventions of F-projections make reference to syntactic concepts. Consider the convention of horizontal F-projection, which is used to explain the deaccentuation of predicates when their complement is accented.

[4] I only mark the F-marker of the predicate and its complement for simplicity.

The concept ‘Internal argument’ used in (83b) has a semantic component, since it excludes non-arguments, and a syntactic component, ‘internal’, imposing a locality restriction on F-projection. Consider the case of depictive predicates. Secondary predicates need an accent even when adjacent to their argument.

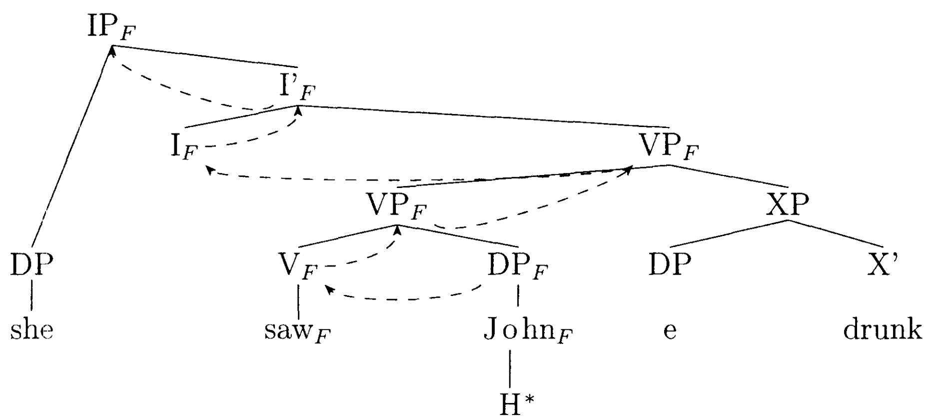
(89) Why is Mary laughing?

a. She saw JOHN DRUNK.

b. # She saw JOHN drunk.

It seems that the argument ‘John’ can license F-marking on the preceding verb but not on the following predicate, although it is arguably an argument of both of them. It has been argued that the argument of depictive predicates is not in the complement position of the predicate:

(90) No F-Projection to an Argument that is not in Complement Position



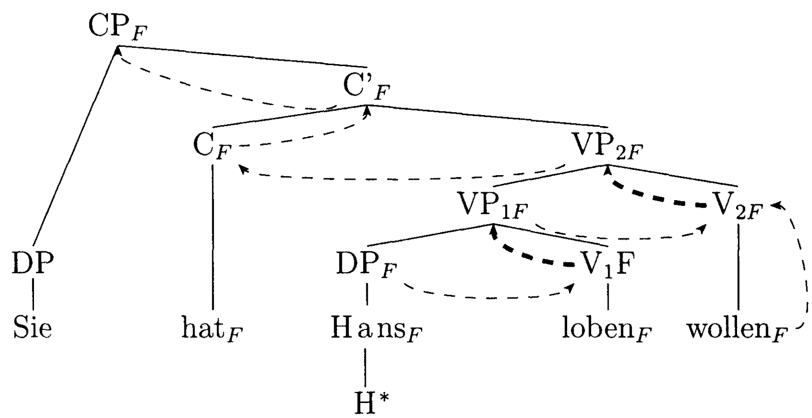
The correct prediction is that (89b) is only felicitous if the predicate ‘drunk’ is given. Otherwise, the predicate has to bear an accent as in (89a), such that the projection that it heads can be F-marked. Horizontal F-projection is thus possible only very locally.

Vertical F-projection (110) serves two purposes: First, it is used to meet the Focus-Condition in (82), since it allows to project F-markers up to higher constituents, e.g., the root node in wide-focus contexts. Second, vertical projection assures that predicates can be F-marked by arguments that are not *directly* their internal argument. An illustration is a German sentence with two sentence final predicates:

(91) Sie hat HANS loben wollen.
she has john praise wanted
'She wanted to praise Hans.'

F-projection works as follows:

(92) Vertical Projection: Multiple Heads



The outer predicate 'wollen' can only be F-marked if its complement 'Hans loben' is F-marked. This is possible only if 'Hans' first F-projects its F-markers to the predicate 'loben' by horizontal F-projection. This then explains why 'loben' does not have to carry an accent. The same reasoning applies to the auxiliary 'hat': in order for it to be F-marked, the predicate 'loben' first needs to project up the F-marker to the complement of 'hat', which then licenses F-marking by horizontal F-projection.

The bold-faced instances of vertical projection in (92) are thus necessary for two reasons: to facilitate further vertical projection to the root node; and to facilitate horizontal projection to further heads.

As we will see later, F-marking is minimized, that is, as few F-markers as possible have to be used. The number of accents however is a factor in choosing between structures. The prediction is that additional optional accents can be placed, as long as the F-marking does not change. This can be illustrated by the following example, which involves the same number of F-markers as (84a), but has one more accent.

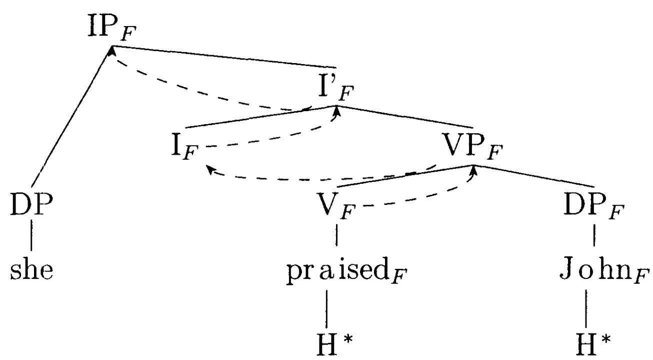
(93) After Mary came, what happened next?

a. [She [praised$_F$ JOHN$_F$]$_F$]$_F$ (example (84a)

b. [She [PRAISED$_F$ JOHN$_F$]$_F$]$_F$

In this case, the verb is F-marked just based on the Basic F-marking rule which states that accented words are F-marked, and no horizontal F-projection from the complement of the verb is necessary:

(94) Optional Accent Placement



In (85), in contrast, the verb was F-marked by horizontal F-projection. F-projection is generally not obligatory, it is only invoked if necessary to license F-markers. The additional accent is compatible with the principle Avoid-F, since the number of F-markers has not been increased by placing the additional accent (compare (94) and (85)).

Unexpectedly, German sharply contrasts with English with respect to this optional accent-placement. In German, it is impossible to place an accent on the verb in a wide-focus context:

(95) After Mary arrived, what happened next?

a. [Sie [hat$_F$ [HANS$_F$ gelobt$_F$]$_F$]$_F$]$_F$. (example (86a))

b. # [She [hat$_F$ [HANS$_F$ geLOBT$_F$]$_F$]$_F$]$_F$.

But the two representations do not differ with respect the number of F-markers, only with respect to the number of accents, just as in the case of English (93) repeated above.

This asymmetry does not follow from any of the conditions on accent-placement within the F-marking approach. The crucial difference between English and German seems to be the position of the head. This asymmetry generalizes to other heads than the verb within both languages:

(96) Prepositions vs. Postpositions in German

a. Sie fuhr entlang des FLUSSES.

b. Sie fuhr ENTLANG des FLUSSES.
she drove along the river

c. Sie fuhr den FLUSS entlang.

d. # Sie fuhr den FLUSS ENTLANG.
she drove the river along

Why can a head bear an optional accent when it precedes its complement, but not when it follows it? One difference between head-initial and head-final structures is that placing an accent on a final head changes the main phrasal stress, while placing an additional accent on an initial head does not. This is due to the fact that the last H* accent within a phrase is perceived as its nuclear stress. The asymmetry points to the following generalization:

(97) Asymmetry stated in terms of F-Projection

An optional accent, i.e., an accent that is not necessary to obtain the required F-marking, may not be the nuclear accent.

But why would this generalization hold? Maybe we can exploit a modified version of Schwarzschild (1999)'s principle of relative prominence. The principle is the following, which is reminiscent of observations from Cinque (1993):

(98) A head is less prominent than its internal argument(s).

This principle could be argued to tolerate an accent on a head that precedes its complement, since it is not perceived as being more prominent; but it penalizes an accent on a head in a head-final construction, since in this case, the head would be perceived as being more prominent than the complement.

The mapping-principle between phonology and syntax (8.2.4) has then taken over part of the job that horizontal F-projection was designed to capture. Even if there was no F-projection from an argument to its head, (98) alone could be blamed for the lack of an accent in head-final structures.

8.2.2 The Givenness-Effect

In the F-marking system in Selkirk (1995b) and Schwarzschild (1999), F-marking interacts with givenness based on the following principle:

(99) Givenness Condition

The absence of F-marking indicates givenness in the discourse.

The only constituent not F-marked in (84a), repeated below, is the subject 'she':

(84) After Mary arrived, what happened next?

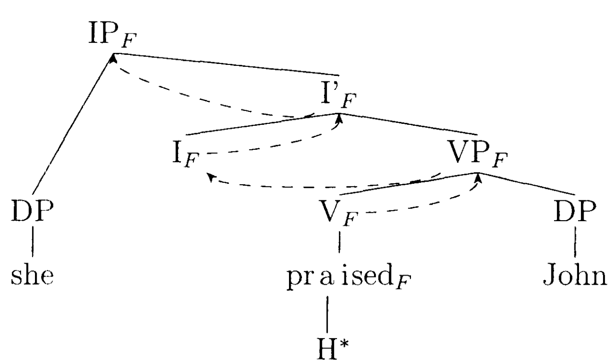
a. [She [praised$_F$ JOHN$_F$]$_F$]$_F$.

b. # [She [PRAISED$_F$ John]$_F$]$_F$.

Since it is not accented, it is not F-marked by the basic focus rule, and none of the F-projection rules license an F-mark on the subject. It remains without marking, therefore it has to be given. This does not cause a problem, since the pronoun indeed refers to discourse referent that was mentioned in the question already, so the Givenness-Condition is not violated.

The Givenness-Condition explains that the non-accentuation of the direct object is infelicitous in a context in which it is not given as in (84b). Consider the pattern of F-projection in the case of an unaccented direct object:

(100) F-Projection and Givenness



The direct object and the subject remain without F-marking, so both have to be given. The direct object is not given, however, in the context in (84b), hence the structure is infelicitous. Consider now the case where the direct object *is* given:

(101) After John and Mary arrived, what happened next?

a. # [She [praised$_F$ JOHN$_F$]$_F$]$_F$.

b. [She [PRAISED$_F$ John/him] $_F$]$_F$.

Here, the lack of the accent on 'John' is acceptable, since the constituent is given. The conditions on givenness do not yet explain why (101a) is infelicitous. Nothing requires an F-marked constituent NOT to be given.

The addition proposed in Schwarzschild (1999) is that F-markers are minimized, that is, if a structure with fewer F-markers is felicitous in a context, an alternative with more F-marking is ungrammatical.[5]

(102) *Avoid-F* (Schwarzschild, 1999, 156)

F-mark as little as possible, without violating Givenness.

Both example (101a), with the F-marking in (85)) and example (101b, with the F-marking in (100)) do not violate the Givenness-Condition. But (101a) involves more F-markers, hence it is infelicitous and only (101b) is acceptable.

The reason that the Givenness condition and Avoid-F cannot be conflated into a bi-conditional (a constituent is given if and only if it is not f-marked) are cases in which a constituent has to be the FOCUS, but all lexical items inside are given. In this case, one of those items has to be accented and hence F-marked in order to be able to project the F-mark. Hence, given material is sometimes F-marked.

A simple case is that of a given constituent that corresponds to the wh-part of the question that is at issue. In this case, the Focus Condition requires this condition to be a FOCUS, hence to contain an F-marker. Then the only candidate for providing

[5]Jacobs (1991a, 5) invokes a similar condition on the minimization of F-marking. F-marking plays out in a different way in the overall theory presented there, although it ultimately serves similar purposes, namely explaining the relation between syntax, focus, and phonology.

the accent that licenses the F-marker is the given item itself (Similar cases with sentence-wide focus will be discussed in section 8.2.4):

(103) Who did John's mother praise?

She praised HIM_F.

Head-initial structures differ from head-final structures with respect to the placement of optional accents. A quite parallel asymmetry can be observed in the placement of optional accents in cases of given material. Placing an additional accent on a given direct object ((104b) is not felicitous in English):

(104) Mary met John. What happened then?

a. [She [PRAISED_F John. $]_F$ $]_F$

b. # [She [PRAISED_F JOHN_F $]_F$ $]_F$.

This is as expected assuming the principle that assures minimizing of F-marking, Avoid-F. The accent on the direct object induces an F-marker, based on the Basic Focus Rule. But the example (104a) involves fewer F-markers, hence (104) is infelicitous. In German, however, placing an additional accent on the direct object is perfectly acceptable:

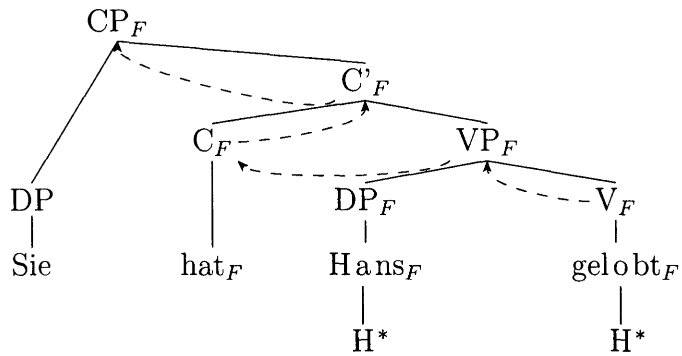
(105) Mary met Hans. What happened then?

a. [Sie [hat [Hans GELOBT_F $]_F$ $]_F$ $]_F$.

b. [Sie [hat [HANS_F GELOBT_F $]_F$ $]_F$ $]_F$.
she has Hans praised

This is unexpected given Avoid-F, since the Basic Focus Rule (83) forces accented material to be F-marked. The additional accent on 'Hans' induces an additional F-marker which is not necessary, hence (105b) should by all means be infelicitous:

(106)



Why is it legitimate to place an additional accent in German, but not in English? One factor that distinguishes the two examples is that only in English does the addition of the optional accent change the location of nuclear stress.

The F-Projection approach allows for a unified statement of the asymmetry observed here and the asymmetry concerning the deaccentuation of predicates, repeated below:

(97) Asymmetry stated in terms of F-Projection

An optional accent, i.e. an accent that is not necessary to obtain the required F-marking, may not be the nuclear accent.

This covers both asymmetries discussed so far descriptively, but it hardly explains them. Why would this generalization hold? If we don't want (97) to be part of our grammar, we have to find an alternative explanation for the givennness asymmetry, along the lines of the principle in (98). The following is a candidate:

(107) Newness-Prominence

New material is more prominent than given material.

This principle may be used to force deaccentuation *without any reference to F-marking.* The logic goes as follows: the last of a series of H_* accents is perceived as most prominent. The principle in (107) thus rules out placing an accent on given material that follows new material (the English example in (104)), but it does not rule out placing an accent on given material that precedes new material (the German example in (105)).

Newness-Prominence introduces a relative notion of givenness: it affects given information only in the presence of new information. This is exactly the notion of givenness that the data from all-given contexts discussed above suggested. If the grammatical marking (e.g. deaccenting) of givenness was *always* the relative marking of given information in the presence of new information, then the observation that the neutral pattern emerges in all-given contexts is no longer surprising.

The principle of Newness-Prominence does part of the job that F-marking was intended to do, similar to the case of the principle of complement-prominence (8.2.4). It can explain the lack of an accent on given material that follows new material. If given material preceding new material was always accented, as in the German example in (105a), then the principle of Newness-Prominence would already account for the data, without any reference to F-marking at all. A version of such a theory was proposed above, invoking a prominence-based account of givenness-marking.

Avoid-F, on the other hand, faces a severe problem: an optional accent on given material should under all circumstances be illicit, since, based on the Basic Focus Rule, each accented word has to be F-marked.

8.2.3 A Revision of Vertical Projection

The F-projection rules in Selkirk (1995b) make a wrong prediction, as was observed in Schwarzschild (1999) and Büring (to appear). Consider a case where the entire VP encodes given information, a type of example also discussed in Jacobs (1988):

(108) Mary praised John. What happened next?
[ANNA$_F$ praised John.]

The problem that arises is that the F-marking on the subject does not license F-marking of the entire sentence. The most conservative fix for this problem is to revise the projection conventions, and allow vertical projection from *any* constituent, not just from heads, following a suggestions from Büring (to appear). Then the facts fall out correctly. The subject is accented, and the material inside of the VP is not accented, since neither it nor the VP-node itself need to be F-marked. The most parsimonious F-marking in tune with the givenness condition in (99) and the focus condition in (82) is the one below:

(109) Mary praised John. What happened next?
[ANNA$_F$ praised John.]$_F$

The following summarizes the revised convention for vertical projection (Büring, to appear):[6]

(110) F-marking on a constituent licenses F-marking on every dominating node.

This improved version of F-marking will now be looked at, and two more problems are pointed out: F-marking is unable to account for the accent-distribution in all-given contexts, and F-marking cannot account for the fact that givenness marking requires a condition on the *sister* of a givenness-marked constituent.

8.2.4 The 'All-Given' Problem

In 'all-given' contexts, the accent placement that emerges sometimes is the same as in the in 'all-new' context. This has already been discussed in Schwarzschild (1999) and Büring (to appear). But sometimes it is actually not the 'neutral' pattern. This section adds to the discussion some more observations related to this problem, and shows why it argues for relatived givenness.

The problem can be illustrated by the following example (based on a similar example suggested by Jon Gajewski, p.c.):[7]

[6]It may seem that we can get rid of vertical F-projection altogether if we restate the givenness condition to the following:

(1) Constituents not containing an F-marker have to be given.

But this is only possible if we find a way to also get rid of horizontal F-projection, since vertical F-projection also helps facilitating horizontal F-projection in the case of multiple predicates, as illustrated in (92).

[7]The problem was first discussed in Schwarzschild (1999, 171–172). In a context where every constituent within a phrase is given, the theory of F-marking remains mute with respect to predictions

(111) Last week after the game all the coach did was praise John. I wonder what she did after this week's game.

a. $[$ She $[$ praised JOHN$_F$ $]_F$ $]_F$.

b. # $[$ She $[$ PRAISED$_F$ John $]_F$ $]_F$.

It is impossible to place the accent anywhere but on the direct object.[8] The fact that in an all-given context the neutral stress pattern emerges is not explained in Selkirk (1995b). Schwarzschild (1999) suggests that the explanation resides in principles

---

regarding the distribution of accents. Consider the following example (Adapted from Schwarzschild changing the F-marking according to the assumption made here.):

(1) {The rising of the TIDES depends on the MOON being full, and}
$[$ the MOON$_F$ being full $]_F$ depends upon $[$ the POSITION$_F$ of the SUN$_F$.$]_F$

The question is how accent placement is arbitrated within a constituent in which all subconstituents are given, as in 'the MOON being full' in (1). The emerging accent pattern according to Schwarzschild in this example is the one that is attested also when [the moon being full] is all new.

[8]Based on the original definition of F-projection and the principle of Avoid-F, the facts should be exactly the reverse:

(1) F-marking based on original convention of vertical projection in (83):

a. $[$ She $[$ praised$_F$ JOHN$_F$ $]_F$ $]_F$.

b. $[$ She $[$ PRAISED$_F$ John $]_F$ $]_F$.

Avoid-F would choose (1b) over (1a). This is another reason not to adopt the original convention but to adopt (110).

guiding the mapping between syntax and phonology.[9]

This is somewhat of a disappointment, since the F-marking approach promised to explain the distribution of accents under all information-strutural circumstances. The principle that Schwarzschild (1999, 170) suggests is the following, repeated from (98):

(98) Complement Prominence

A head is less prominent than its internal argument(s).

The principle (8.2.4) is a principle that relates directly to the mapping between syntax and prosody, it does not use F-markers or relate to conditions of givenness in any way. This principle kicks in exactly when the account in terms of F-marking does not decide between possible outputs.

The problem for the F-marking approach is more serious than Schwarzschild (1999) suggests. There must be further principles apart from (8.2.4) at play, since in all-given cases where the constituent consists of more than just heads and internal arguments the all-new pattern also emerges. Consider the prosody in a context with a subject that is a full noun phrase and not a pronoun:[10]

(112) The only memorable thing at the last party was that Mary praised John. What do you think was most surprising this time?

a. [MARY_F [praised JOHN_F $]_F$ $]_F$.

[9]In the account in Selkirk (1995b) and Schwarzschild (1999), syntax-phonology mapping principles might also account for the fact that there is no utterance with no accent whatsoever, a state of affairs that one might have expected in the Avoid-F approach to accent distribution:

(1) * She praised John.

(with no accent whatsoever)

However, in the answer to the question in (111), it may only appear as if all material is given. The particular time and occasion of the event reported in the answer is not, and this might be enough to force F-marking on the some part of the representation and hence explain the accent.

[10]For each example, I put the minimum number of F-markers possible in accordance with the Basic Focus Rule, the projection principles, and the Focus Condition.

b. # [MARY_F [PRAISED_F John $]_F$ $]_F$

c. # [MARY_F praised John. $]_F$

The subject gets accented and hence F-marked just like in 'all-new' contexts. The very principles used to explain F-marking in all-new contexts, Avoid-F and the Givenness Condition, do not seem to be at work in all-given contexts. F-marking does not seem to be minimized. Note that (112b) has just as many F-markers as (112a), and that (112c) has fewer F-markers than either (112a) or (112b). Any of the structures is compatible with the Givenness condition and the Focus Principle, so there is no reason why (112a) of all cases should be the felicitous one. The accent on 'Mary' in (112) should be ruled out by Avoid-F, since it is not needed to satisfy either the Givenness-Condition or the Focus-Condition.

There must be an entire system of accent rules, only there in order to operate in all-given situations. This comes as a surprise for the F-marking approach. Even conceding these additions to the F-marking story, a fundamental property of all-given contexts still remains unexplained. *The all-given pattern is the same as that observed in all-new contexts, that is the pattern in a context where no information is given.* But this pattern emerges for entirely different reasons that do not relate to F-marking. Why would it be precisely the 'all-new' pattern that emerges? This does not follow from the approaches taken in Selkirk (1995b) or Schwarzschild (1999).

We can conclude with Schwarzschild (1999) that principles of the syntax-phonology mapping such as the Complement-Prominence principle and more are needed to account for all-given contexts. The additional problems pointed out in this context are that principles such as Avoid-F seem to be suspended in 'all-given' contexts, and that 'all-given' contexts strangely show the same prosodic pattern as 'all-new' contexts.

The problem is worse, however. In the following example, again all constituents are given (repeated from (113):

(113) Who of Mary and Ann praised John?

a. [MARY_F praised John. $]_F$

b. # [Mary praised JOHN_F $]_F$

The theory of F-marking fails to decide between (113a and b), and moreover it fails to distinguish between cases with 'neutral stress' (112) and cases with 'marked stress' (113). Schwarzschild (1999, 166) discusses a related example.

(114) Did Karen get the money or did Marc get the money?

a. KAREN got the money.

b. * Karen got the MONEY.

Here, marking [got the money] as given as in (114a) is possible, and hence obligatory. There is an alternative y for 'KAREN' such that 'y ate got the money' is given, the presupposition of relative givenness is fulfilled. A more detailed discussion of the problem can be found in chapter 7.

8.2.5 The Locality Problem

The condition on givenness (99) together with the principle of Avoid-F (102) is able to account for many cases of non-accentuation of given material. The following is an example adapted from Schwarzschild (1999, 146):

(115) Mary's old convertible is no longer available. What is John going to do?

a. He'll [[rent_F [her NEW_F convertible. $]_F$ $]_F$

b. # He'll [[rent_F [her new CONVERTIBLE_F. $]_F$ $]_F$

c. # He'll [[rent_F [her NEW_F CONVERTIBLE_F. $]_F$ $]_F$

The non-accentuation of 'convertible' in (115a) is licensed by the fact that the constituent 'x convertible' is given in the context. 'New' needs to carry an accent, since it constituents new information, so (115b) is not felicitous due to the Givenness Condition. Accenting BOTH 'new' and 'convertible' as in (115c) is ruled out by Avoid-F. In the following example, 'x convertible' is again given in the context, yet the judgements are different:

(116) Mary's uncle, who produces high-end convertibles, is coming to her wedding. I wonder what he brought as a present.

a. # He brought [a RED$_F$ convertible]$_F$.

b. He brought [a red CONVERTIBLE$_F$]$_F$

c. He'll [[brought$_F$ [a RED$_F$ CONVERTIBLE$_F$.]$_F$]$_F$

The change in judgement is unexpected for the theory of F-marking. 'Convertible' in (116) should be just as given as in (115), and thus (115b) should be the only felicitous candidate, but in fact it is the only *infelicitous* one in (116).

It seems that what is wrong with (116b) is that 'high-end' is not a relevant alternative to 'red', and therefore a contrast does not make a lot of sense. Of course, given the right context that provides a partition of convertibles into 'high-end' ones and 'red'-ones, the two predicts would be alternatives and givenness marking would be possible.

Similar facts can be observed for DPs co-referent with a salient discourse antecedent:

(117) Do you think Mary had anything to do with the burglary?

a. Yes. [They [arrested JOHN$_F$ and MARY$_F$.]$_F$]$_F$

b. ?? Yes. [They [arrested JOHN$_F$ AND$_F$ Mary.]$_F$]$_F$

c. # Yes. [They arrested JOHN and Mary.]

The acceptability of (117a) is unexpected, since 'Mary' is given in the discourse and hence should not need be able to receive an F-marker due to Avoid-F. Hence, it should not be able to carry an accent.

The intuition about what is going on here is that when deciding whether or not a constituent is marked as 'given', the immediate context of a constituent somehow is taken into account. This intuition is captured with the approach based on relative givennness presented in chapter 7.

8.3 Reference Set Approach

In the reference set approach (Szendröi, 2001, Reinhart, 2004a,b) prominence is assigned by the nuclear stress rule, just as in Cinque (1993). Under special circumstances (focus related) stress shift takes places; under certain other circumstances givenness-related destressing applies. These are two different processes, and both apply directly to phonological representations. The main problem of the Reference Set Approach is that it is redundant, and one of the two operations can be eliminated once we use the notion of *relative givenness* from chapter 7.

8.3.1 Destressing and Main Stress Shift

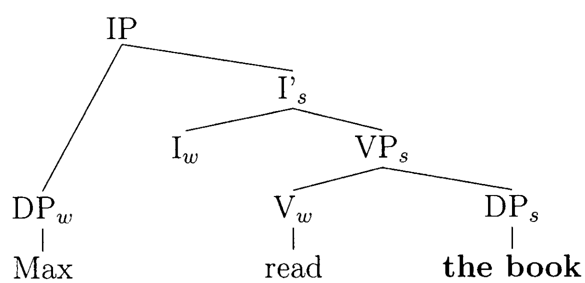
The phonological representation assumed are the annotated syntactic trees from Liberman (1975). The nuclear stress rule assumed in this approach is as follows:

(118) Nuclear Stress Rule

Assign a strong label to the node that is syntactically more embedded at every level of the metrical tree. Assign Weak to its sister node.

Every annotated syntactic tree defines a 'focus set', a set of constituents that can be in focus given the annotation of the representation. The definition of the focus set is based on where main stress falls.

(119)



Focus Set: { DP$_{object}$, VP, IP}

Conventions about Main and Secondary Stress from Szendröi (2001) are as follows (two special cases from the tree to number conversion in Libermann and Prince (1977, 259):

(120) a. Main Stress

...falls on the terminal node that is connected to the root node by a path that does not contain any weak nodes, including the root note itself and the terminal node. [...]

b. Secondary Stress

...falls on the terminal node whose path to the root node contains only S nodes, except for exactly one W label on the node immediately dominated by the root node. [...]

The Focus set is now defined as follows:

(121) Focus Set

The focus set of a derivation D comprises all and only subtrees (constituents) which contain the main stress of D.

There are two operations in addition to the nuclear stress rule that manipulate stress. They do so after the NSR has applied, and, similar to the approach in Chomsky (1971), they overwrite the original output:

(122) a. Stress Shift (global)

Relocate the main stress on a constituent you want to focus.

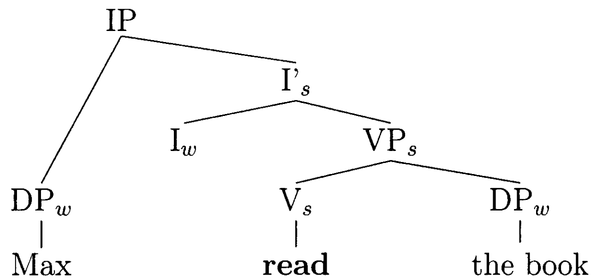
b. Destressing (local)

Mark a constituent that is given with 'w'

Destressing changes the label 's' to 'w' is a node is given. This operation is adopted form Williams (1997). It is applied locally, and only requires evaluating whether or not a node is given.

(123) 'Destressing' following Szendröi (2001) (cf. Williams (1997))

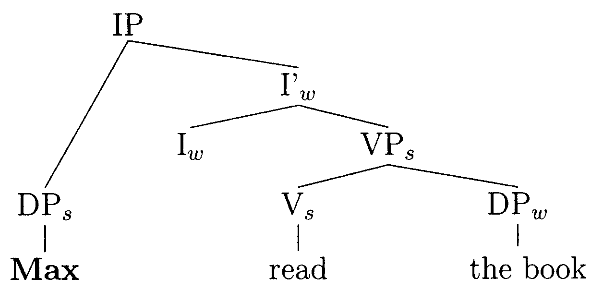
Assign W to an anaphoric node.



There is a second operation, main stress shift. This operation makes sure that the focus set includes the constituent that is in focus. It operates globally, and can change every label of the tree if necessary.

(124) Main Stress Shift

Assign S to a node α and every node dominating α



8.3.2 Global Economy

Why is it that the sentence in (124)—according to the Reference Set account—is not compatible with a question with wide focus? Note that the focus set includes IP. The reason is that this sentence is ruled out by Economy. Economy compares derivations, and selects the most economical derivation compatible with the focus in the context:

(125) Global Economy: Marked and Unmarked derivations

a. For a sentence meaning with a certain focus there may be marked and umarked derivations that express the meaning and whose focus set includes the required focus marking.

b. Only the most economical element of this set is grammatical.
Stress shift is costly (in fact, each single s—w change in the case of main stress shift is costly).

The converging derivations that have the correct focus set form the reference set:

(126) a. d: converging derivations based on numeration (including PF)

b. i: interpretation

c. reference set: $< d, i >$ (derivations yielding correct interpretation)

For the example in question:

(127) Example

d: Max read a BOOK $\{DP_{object}$, VP, IP$\} \rightarrow$

MAX read a book $\{DP_{subject}$, IP$\}$

i: Focus: IP

'Max read a BOOK' blocks 'MAX read a book', since it does not involve stress shift.

One problem with the Reference Set approach is that Detressing and Main Stress Shift overlap. In fact, once 'destressing' is redefined in terms of *relative giveness*, 'Main Stress Shift' becomes unnecessary.

8.3.3 Overlap

Main stress shift and destressing often have similar results. Consider the following example:

(128) a. IP-Focus: What's this noise?

[$_F$ My neighbor is building a **desk**].

#[$_F$ **My neighbor** is building a desk].

b. VP Focus: What's your neighbor doing these days?

My neighbor [$_F$ is building a **desk**].

My neighbor [$_F$ is **building** a desk.]

According to Reinhart (2004a, 28), the example in (128a) cannot be used in contexts with wide focus since main stress shift has occurred: 'As we saw, stress obtained by the default (nuclear, neutral) stress rule allows any projection containing it to serve as focus, e.g., the whole IP [...]. The shifted case by contrast, cannot be used in the same context [...].' The claim is that VPs and indefinites cannot lose their stress by destressing, but only by main stress shift.

Only stress shift can give the marked prosodies in (128), reference set comparison rules out wide focus. But sometimes (128b) is possible, and in fact obligatory, in contexts that require with wide focus (cf. Jacobs (1988)):

(129) Last week, there was a lot of noise because Bill was building a desk. What's the noise today?

#[$_F$ My neighbor is building a **desk**].

[$_F$ **My neighbor** is building a desk].

Another claim in Reinhart (2004a) is that indefinites cannot undergo destressing, but can only lose stress by main stress shift, requiring narrow focus. But in the appropriate context, they can destress, and yet the sentence occurs in a wide-focus context:

(130) Last time I saw her, your neighbor was looking for a desk at a yardsale. What's she doing these days?

\# My neighbor [$_F$ is building a **desk**].

My neighbor [$_F$ is **building** a desk.]

Destressing of indefinites and entire VPs exists, and is compatible with wide focus, contrary to Reinhart's claims. The data is of course still compatible with the Reference set approach, in which there is no reason to stipulate that VPs and indefinites cannot destress.

But once this stipulations are abandoned, there are two different ways to account for the stress pattern in (128b), main stress shift and destressing. Destressing of given material alone is sufficient as an explanation why the VP is not stressed:

(131) Who build a desk?

#[$_F$ My neighbor] is building a **desk**.

[$_F$ **My neighbor**] is building a desk.

Similarly, a question that places focus on the verb makes the indefinite object given, and hence the lack of stress on it can be explained by destressing and without main stress shift:

(132) What did you say your neighbor is doing with a desk?

My neighbor is [$_F$ building] a **desk**.

My neighbor is [$_F$ **building**] a desk.

In other words, main stress shift is unnessary to account for the data. A remaining problem is the contrast between (129) and (133). Here, there is an antecedent 'building a desk' for destressing. But the subsequent answer has focus on the VP.

(133) Last week, there was a lot of noise because Bill was building a desk. Now I think your neighbor is making a lot of noise. What is he doing?

[$_F$ My neighbor is building a **desk**].

#[$_F$ **My neighbor** is building a desk].

Why is destressing not possible, or at least not favored here? The presupposition of relative givenness would require there to be an alternative x for the subject, such that 'x was building a desk' is given. This seems to be the case, since 'Bill was building a desk' is given in the context. This looks like an effect of the requirement that the focus of an answer has to include the main word stress.

But note that this antecedent is not an element of the question under discussion: 'What is he doing?'. If we require antecedents for givenness accenting to be part of the question under discussion as proposed in chapter 7, then we can explain this last effect as well, without invoking the focus set.

To summarize, 'Destressing' can apply to constituents of any size, still allowing for wide focus. Once we adopted 'destressing', using the version requiring *relative givenness* for it to apply, main stress shift becomes unnecessary.

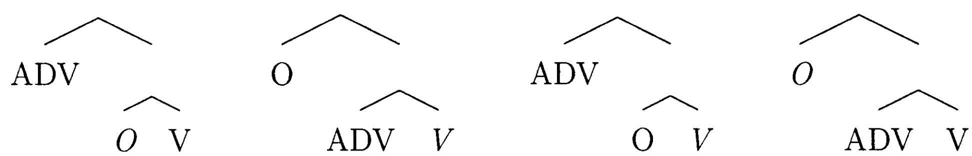
8.3.4 Stress and Movement

Neeleman and Reinhart (1998) relate stress and movement in a direct way. The basic idea is that main stress shift is dispreferred, and in cases where the direct should not receive nuclear stress because it is given, it is better if this is achieved by scrambling the object and applying the default stress rule rather than applying main stress shift.

The four relevant combinations of scrambling and nuclear stress (scrambling yes/no, nuclear stress on verb/direct object) are summarized below. The element carrying main stress is marked with italics:

(134) Expected Focus Sets, Stress * Scrambling (stressed phrase slanted)

a. Base w.o. b. Scrambled c. Base w.o., SS d. Scr., SS



{IP,VP,O} {IP, VP, V} {IP, VP, V} {IP, O, VP}

Basic Idea: Stress assigned to object in base position, to verb when object is scrambled. When having the option, scramble rather than shift stress.

Derivations (c) and (d) involve stress shift and should hence be dispreferred. It is correct that given material tends to scramble and not be stressed. A typical example involves shifing a *given* direct object across 'probably' (135a). Not shifting and receiving nuclear stress under these circumstances is marked, as indicated in (b):

(135) What did she do with the book?

a. Sie hat wahrscheinlich das Búch gelesen. (134a)
she has the book probably read

b. 9

Sie hát das Buch/es wahrscheinlich gelésen.(134b)
she has the book/it probably read

However, it also seems possible to simply stay to the right of the adverb and stress the predicate:

(136) What did she do with the book?

a. Sie hat wahrscheinlich das Buch gelésen.(134c)
she has the book probably read

b. # Sie hát das Búch/es wahrscheinlich gelesen. (134d)
she has the book/it probably read

The question is then whether it is not sufficient to say that main prominence goes to the predicate when the object is given—overt scrambling across the adverb does not seem to be a relevant factor here. Of course, the direct object might still undergo short scrambling even when it is to the right of the adverb—this can not be ruled out.

A more severe problem relates to the case where the object did scramble away. What the approach in Neeleman and Reinhart (1998) captures correctly is that scrambling in cases in which the direct object is not given is not felicitous under either nuclear stress assignment, at least in the following context:[11]

(137) Why was there an uproar in the palace?

a. # Kóenig Ubú hat die Króne wahrscheinlich verlóren.(134b)
king Ubu has probably the crown lost
'King Ubu probably lost the crown.'

b. # Kóenig Ubú hat die Króne wahrscheinlich verloren.

(134d)

Sentence (a) would require 'the crown' to be given information, while (b) would require an even more specific context. What this means is that scrambling is not optional and free but has to be motivated by discourse reasons.

Where the approach in Neeleman and Reinhart (1998) goes wrong is in tieing destressing to scrambling. If the direct object moves as in (137), one has in principle both options—either nuclear stress on the verb or the direct object.

(138) a. # Kóenig Ubú hat die Króne wahrscheinlich verlóren.(134b)
king Ubu has probably the crown lost

[11]Note that definiteness here is not a cue to givenness since there is usually only one crown per king. See chapter 7for discussion, and also Wagner (2002b). The direct object is discourse new in this context.

'King Ubu probably lost the crown.'

Required context: The crown is salient.

b. # Kóenig Ubú hat die Króne wahrscheinlich verloren. (134d)

Required context: There is an alternative y to the crown such that [y probably lost] is salient.

We can make sense of these presuppositions by applying the reasoning of relative givenness from chapter 7. By moving the direct object, we create the following representation:

(139) [The crown] [λx. probably lost x]]

Marking 'the crown' as given relative to its sister introduces the presupposition that there is some other proposition about the crown that is salient—in other words, the crown has to be salient; marking [λx. probably lost x] as given relative to the crown introduces the presupposition that there is some alternative object y such that [probably lost y] is salient.

The latter presupposition is stronger since the part presupposed to be part of the antecedent is bigger. This is the reason for the intuition that this stress pattern is 'more marked'. But the formal ingredients of both stress patterns are identical.

In both cases, the movement is motivated by givenness marking—what differs is which element is given relative to which.

There is thus no automatism of the form: if you scramble, the nuclear stress fall on the verb. The prosody depends on the operator that is involved in the movement. For more problems for stress-motivated reasons see Horvath (2004).

To conclude, the reference set approach is redundant, and it seems that destressing and main stress shift overlap and often serve to account for the same prosodic facts. The motivation for the reference set approach in terms of an interaction between movement and stress seems not to rest on the wrong assumption that there indeed is a straightforward trade-off between scrambling and losing nuclear stress.

Summary

This chapter looked at proposals that try to account for the conditions of accent omission. The two alternative approaches discussed here cannot account for the same range of data as the proposal outlined in chapter 6 and chapter 7. There are many other proposals that would require a closer look at this point, e.g. the approaches in Gussenhoven (1984, et. seq.) and Jacobs (1991b, 1999). Some aspects of those proposals are discussed in chapter 6 and chapter 7, but a thorough review has to be postponed to some other occasion.

Chapter 9

Conclusion

The Relation between Syntax and Prosody

The main tool in capturing prosodic patterns used in this thesis is recursion. Every output of the grammar is assigned a meaning and a phonological representation. The output can then enter the computation of a new and bigger structure as an atom. The assigned semantic and prosodic representations are not changed anymore. They will enter and form part of the meaning and phonological representation of the new structure. Both the computation of meaning and the computation of prosody is compositional.

The particular compositional theory proposed here is tested against a limited set of data from a very limited set of languages (English, Dutch, German). As any other theory, it will have to undergo changes as more facts are looked at. There are reasons for optimism that the proposed theory of how syntax relates to prosody is a step in the right direction.

First, the theory is very simple and maximizes the use of the independently motivated recursive steps in syntax. There are only two mapping principles. The first is PROSODIC MATCHING (1):

(1) PROSODIC MATCHING

a. Concatenate

Concatenate the prosodic representation of the elements in the domain aligning their top lines and filling the columns where necessary.

b. Project

Create a new top-line grid line n by projecting all grid marks on line n-1, and mapping them into a single foot on line n.

The second principle is PROSODIC SUBORDINATION:

(2) PROSODIC SUBORDINATION

A is subordinated to B by concatenating A and B such that if A projects to line i, B projects to line i+1.

Which principle applies and when depends on syntactic and semantic factors, e.g. on argument structure and information structure. This choices of which principle to apply are the *only* way in which these factors can affect prosody, and they can do so only at points which where independently identified as syntactic cycles.

Second, some well-known facts that have earlier been taken as mismatches between prosody and syntax turn out not to be mismatches. In the system proposed here, they fall out from the very properties of the mapping principle in (1), and the syntactic derivation.

Many questions remain open, and many expectations untested. Two of the major gaps in the picture are first, an account for the precise alignment of pitch accents and boundary tones with the metrical grid; and second, the source of word order variation (e.g., the difference between VO and OV) remains unexplored.

Third, the close link to recursion in syntax makes the proposed theory vulnerable to counter-arguments not only from phonology but also from syntax and semantics. This vulnerability is a good thing. It means that the theory relates independent sets of facts and enables to ask questions that to the core of the architecture of grammar.

If the theory presented here is on the right track, then it provides an argument that syntax is derivation and computed in cycles. Prosody reflects information both about the syntactic bracketing of phrase marking and about the derivational history of how they were assembled.

References

Abney, Steven. 1987. The engilsh np in its sentential aspect. Doctoral Dissertation, MIT, Cambridge, Ma.

Adger, David. 2001. Syntax and prosody in Scottish Gaelic. VSO Workshop, Stuttgart.

Adger, David. 2003a. *Core syntax*. Oxford Universtiy Press.

Adger, David. 2003b. Phonology and phasal syntax. Paper presented at GLOW.

Alexiadou, Artemis. 1997. *Aderb placement: A case study in antisymmetric syntax*. Amsterdam: Benjamins.

Allbritton, D.W., G. McKoon, and R. Ratcliff. 1996. Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of experimental psychology: Learning, Memory, and Cognition* 22:714–735.

Alonso-Ovalle, Luis. 2005. Distributing the disjuncts over modal space. In *Proceedings of NELS 34 at Stony Brook*, ed. Mathew Wolf and Keir Moulton.

Altmann, Hans, ed. 1988. *Intonationsforschungen*. Tübingen: Niemeyer.

Anagnastopoulou, Elena, and Artemis Alexiadou. 2001. The subject in situ generalization, and the role of case in driving computations. *Linguistic Inquiry* 32:193–231.

Andrews, A. 1983. A note on the constituent structure of modifiers. *Linguistic Inquiry* 14:695–97.

Arregi, Karlos. 2002. Focus on Basque movement. Doctoral Dissertation, MIT, Cambridge, Ma.

Blaszczak, Joanna, and Hand-Martin Gärtner. 2005. Intonational phrasing and the scope of negation. *Syntax* 7:1–22.

Bard, E.G., and M.P. Aylett. 1999. The dissociation of deaccenting, givenness, and syntactic role in spontaneous speech. In *Proceedings of the XIVth International Congress of Phonetic Sciences*, ed. J. Ohala, Y. Hasegawa, D. Ohala, A. Granville, and C. Bailey, 1753–1756.

Bard, Ellen Gurman, Anne H. Anderson, Catherine Sotillo, Matthew Aylett, Gwyneth Doherty-Sneddon, and Alison Newlands. 2000. Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language* 42:1–22.

Bech, Gunnar. 1955/57. *Studien über das deutsche Verbum Infinitum.* Tübingen: Niemeyer.

Beckman, Mary, and Janet Pierrehumbert. 1986. Intonational structure in english and japanese. *Phonology Yearbook* 3:255–310.

Beckman, Mary E., Julia Hirschberg, and Stefanie Shattuck-Hufnagel. 2005. The orginal tobi system and the evolution of the tobi framework. In *Prosodic models and transcription: Towards prosodic typology*, ed. Sun-Ah Jun. Oxford University Press.

van den Berg, Rob, Carlos Gussenhoven, and Toni Rietveld. 1992. Downstep in Dutch: implications for a model. In *Papers in laboratory phonology, vol. ii: Gesture, segment, prosody*, ed. Gerlad Docherty and Robert Ladd, 335–58. Cambridge: Cambridge University Press.

Berman, A, and M. Szamosi. 1972. Observations on sentential stress. *Language* 48:304–325.

Bierwisch, Manfred. 1966a. Regeln für die Intonation deutscher Sätze. In Bierwisch (1966b), 99–201.

Bierwisch, Manfred, ed. 1966b. *Untersuchungen über den deutschen Akzent*, volume 7 of *Studia Grammatika*. Berlin: Akademie Verlag.

Bierwisch, Manfred. 1968. Two cricitcal problems in accent rules. *Linguistics* 4:173–178.

Bierwisch, Manfred. 1983. Semantische und konzeptuelle Repräsentation lexikalischer Einheiten. In *Untersuchungen zur semantik*, ed. R. Ruzicka and W. Motsch, volume 22 of *Studia Grammatika*, 61–99. Berlin: Akademie Verlag.

Bierwisch, Manfred. 2003. Heads, complements, adjuncts: Projection and saturation. In Lang et al. (2003), 113–159.

Boersma, Paul, and David Weenink. 1996. PRAAT, a system for doing phonetics by computer. report 132. Institute of Phonetic Sciences of the University of Amsterdam.

Bolinger, D. 1958. Stress and information. *American Speech* 33:5–20.

Bolinger, D. 1972. Accent is predictable (if you are a mind reader). *Language* 48:633–644.

Bolinger, Dwight. 1977. *Meaning and form*. New York: Longman.

Bresnan, Joan. 1971. Sentence stress and syntactic tranformations. *Language* 47:257–81.

Bresnan, Joan W. 1972. Stress and syntax: A reply. *Language* 48:326–342.

Büring, Daniel. 2005. *Binding theory*. Cambridge: Cambridge University Press.

Büring, Daniel. to appear. Focus projection and default prominence. In *Proceedings from the symposium 'Informationsstruktur —Kontrastiv' (working title)*, ed. Valéria Molnár and Susanne Winkler. Mouton De Gruyter.

Byrd, Dani, and Elliot Saltzman. 1998. Intragestural dynamics of multiple phrasal boundaries. *Journal of Phonetics* 26:173–199.

Byrd, Dani, and Elliot Saltzman. 2003. The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31:149–180.

Calhoun, Sasha. 2004. Phonetic dimensions of intonational categories—the case of L+H* and H*. In *Speech Prosody 2004*. Nara, Japan.

Calhoun, Sasha. 2005. It's the difference that matters: An argument for contextually-grounded acoustic intonational phonology. Paper presented at the Annual Meeting of the LSA in Oakland, Ca.

Cardinalettie, Anna, and Michal Starke. 1994. The typology of structural deficiency. Ms. University of Venice, University of Geneva.

Carlson, Katy, Jr. Charles Clifton, and Lyn Frazier. 2001. Prosodic boundaries in adjunct attachment. *Journal of Memory and Language* 45:58–81.

Chen, M. Y. 1987. The sytax of Xiamen tone sandhi. *Phonology Yearbook* 4:109–49.

Cho, T. 2002. *The effects of prosody on articulation in English*. New York: Routledge.

Cho, T., and P. Keating. 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29:155–190.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of a theory of grammar*. Cambridge, Ma: MIT Press.

Chomsky, Noam. 1971. Deep structure, surface structure, and semantic interpretation. In *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology*, ed. D.D. Steinberg and L.A. Jakobovits. Cambridge: Cambridge University Press.

Chomsky, Noam. 1994. Bare phrase structure. In *Government binding theory and the minimalist program*, ed. Gerd Webelhuth, 382–439. Oxford: Oxford University Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale. a life in language.*, ed. Michael Kenstowicz, 1–52. MIT Press.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English.* New York: Harper & Row.

Chomsky, Noam, Morris Halle, and Fred Lukoff. 1957. On accent and juncture in English. In *For Roman Jacobson*, ed. Morris Halle, Horace Lunt, and Hugh MacLean. The Hague: Mouton.

Chomsky, Noam, and M. P. Schützenberger. 1963. The algebraic theory of context-free languages. In *Computer programming and formal systems*, ed. P. Braffort and D. Hirschberg, 119–161. Amsterdam: North Holland.

Church, K., and R. Patil. 1982. How to put the blocks in the box on the table. *American Journal of Computational Linguistics* 8:139–149.

Cinque, Guglielmo. 1993. A null-theory of phrase and compound stress. *Linguistic Inquiry* 24:239–298.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective.* Oxford: Oxford University Press.

Clark, Herbert H. 1971. The importance of linguistics for the study of speech hesitations. In *The perception of language*, ed. D. Horton and J. Jenkins, 69–78. Columbus, OH: C.E. Merril.

Clements, G. N. 1978. Tone and syntax in Ewe. In *Elements of tone, stress, and intonation*, ed. D. J. Napoli, 21–99. Washington, D.C.: Georgetown University Press.

Cohn, Abigail. 1989. Stress in Indonesian and bracketing paradoxes. *Natural Language and Linguistic Theory* 7:167–216.

Collins, Chris. 1988. Part I: Conjunction adverbs. Ms. MIT.

Cooper, William E., and Jeanne Paccia-Cooper. 1980. *Syntax and speech.* Cambridge, Mass.: Harvard University Press.

Csirmaz, Aniko, Zhiqiang Li, Andrew Nevins, Olga Vaysman, and Michael Wagner. 2002. *Phonological answers (and their corresponding questions)*, volume 42 of *MIT Working Papers in Linguistics*. Cambridge, Ma.

Davidson, Donald. 1967. The logical form of action sentences. In *The logic of decision and action*, ed. N. Reschler, 81–95. Pittsburgh: University of Pittsburgh Press.

Diesing, Molly. 1992. *Indefinites*. Cambridge, Ma.: MIT Press.

Dilley, L., S. Shattuck-Hufnagel, and M. Ostendorf. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24:423–444.

Dilley, Laura. 2005. The phonetics and phonology of tonal systems. Doctoral Dissertation, MIT.

Dobashi, Yoshihito. 2003. Phonological phrasing and syntacic derivation. Doctoral Dissertation, Cornell University.

Donegan, P. J., and D. Stampe. 1983. Rhythm and the holistic organization of language structure. In *Proceedings of the Parasession on the Interplay of Phonology, Morphology and Syntax at CLS 19*.

Dowty, Richard. 2003. What is 'Direct Compositionality'? Talk presented at the workshop on direct compositionality. Brown University.

Dresher, Elan B. 1994. The prosodic basis of the Tiberian Hebrew system of accents. *Languaage* 70:1–52.

Elordieta, G., S. Frota, P. Prieto, and Vigário M. 2003. Effects of constituent weight and syntactic branching on phrasing decisions in Romance. In *Proceedings of the XVth International Congress of Phoentic Sciences*.

Emonds, J. E. 1976. *A transformational approach to English syntax*. New York: Academic Press.

Ernst, Thomas. 2001. *The syntax of adjuncts*. Cambridge: Cambridge University Press.

Esposito, A., and H. Truckenbrodt. 1998. A note on raddoppiamento sintattico and the phonological phrase in italian. IIASS and Rutgers University. MS.

Evers, Arnold. 2001. Verb clusters and cluster creepers. Utrecht.

Ferreira, F. 1988. Planning and timing in sentence production: The syntax-to-phonology conversion. Doctoral Dissertation, Umass Amherst.

Ferreira, F. 1993. Creation of prosody during sentence production. *Psychological Review* 100:233–253.

Fery, Caroline, and Hubert Truckenbrodt. 2004. Tonal scaling and the sisterhood principle. Universität Potsdam and Universität Tübingen.

Fougeron, C., and P. Keating. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101:3728–3740.

Fox, Danny. 2002. Antecedent contained deletion and the copy theory of movement. *Linguistic Inquiry* 33:63–96.

Fox, Danny, and Jon Nissenbaum. 1999. Extraposition and scope: a scope for overt QR. In *Proceedings of the 18th West Coast Conference on Formal Linguistics (WCCFL 18)*, ed. S. Bird, A. Carnie, J. D. Haugen, and P. Norquest. Somerville, Ma.: Cascadilla Press.

Fox, Danny, and Jon Nissenbaum. 2000. Extraposition and covert movement. Talk presented at NYU.

Fox, Danny, and David Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31:1–45.

Frazier, Lyn, Alan Munn, and Chalres Jr. Cliften. 2000. Processing coordinate structures. *Journal of Psycholinguistic Research* 29:343–370.

Freeze, Ray. 1992. Existentials and other locatives. *Language* 68:553–95.

Gazdar, G., E. Klei, G. Pullum, and I Sag. 1985. *Generalized phrase-structure grammar*. Cambridge, Ma.: Harvard University Press.

Gee, J., and F. Grosjean. 1983. Performance structures: A psycholinguistic appraisal. *Cognitive Psychology* 15:411–458.

Geuder, W. 2000. Oriented adverbs. issues in the lexical semantics of event adverbs. Doctoral Dissertation, Universität Tübingen.

Ghini, Mirco. 1993a. ϕ-formation in Italian: A new proposal. In *Toronto working papers in linguistic*, ed. Kerry Dick, volume 4, 41–78. University of Toronto.

Ghini, Mirco. 1993b. Phonological phrase formation in Italian. M.A. Thesis. University of Toronto.

Gleitman, Lila R. 1965. Coordinating conjunction in English. *Language* 41:155–84.

Goldsmith, John A., ed. 1995. *The handbook of phonological theory*. London: Blackwell.

Grice, H.P. 2005. *Studies in the way of words*. Cambridge, Ma.: Harvard University Press.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, Ma.: MIT Press.

Grosjean, F., and M. Collins. 1979. Breathing, pausing, and reading. *Phonetica* 36:98–114.

Gussenhoven, Carlos. 1983. Testing the reality of focus domains. *Language and Speech* 26:61–80.

Gussenhoven, Carlos. 1984. *On the grammar and semantics of sentence accents*. Dordrecht: Foris.

Gussenhoven, Carlos. 1986. English plosive allophones and ambisyllabicity. *Gramma* 10:119–141.

Gussenhoven, Carlos. 1990. Tonal association domains and the prosodic hierarchy in English. In *Studies in the pronuncation of english. a commemorative volume in honour of a.c. gimson*, ed. S. Ramsaran, 27–37. London: Routledge.

Gussenhoven, Carlos, and A. C. M Rietveld. 1992. Intonation contours, prosodic structure, and preboundary lengthening. *Journal of Phonetics* 20:282–303.

Habsieger, L., M. Kazarlan, and S Lando. 1998. On the second number of plutarch. *American Monthly* 105:446.

Haider, Hubert. 1993. *Deutsche syntax - generativ*. Tübingen: Gunter Narr.

Haider, Hubert. 1994. Fakultativ kohärente infinitivkonstruktionen im Deutschen. In *Zur satzwertigkeit von Infinitiven und Small Clauses*, ed. Anita Steube and Gerhild Zybatow. Tübingen: Niemeyer.

Haider, Hubert. 2000. Branching and discharge. In *Lexical specifciation and insertion*, ed. Peter Coopmans, Martin Everaert, and Jane Grimshaw. Amsterdam, Philadelphia: John Benjamins.

Hale, Kenneth, and Elizabeth O. Selkrik. 1987. Government and tonal phrasing in Papago. *Phonology Yearbook* 4:185–194.

Hall, Daniel Currie. 2000. Prosodic representation and lexical stress. In *Proceedings of the 2000 Annual Conference of the Canadian Linguistic Association*, ed. J.T. Jensen and G. van Herk, 49–60. Ottawa: Cahier Linguistique d'Ottawa.

Halle, Morris, and William Idsardi. 1995. General properties of stress and metrical structure. In Goldsmith (1995), 403–443.

Halle, Morris, and Jean-Roger Vergnaud. 1987. *An essary on stress*. Cambridge: MIT Press.

Halliday, M.A.K. 1967. Notes on transitivity and theme in English, part 2. *Journal of Linguistics* 3:199–244.

Harris, Zellig S. 1957. Co-occurrence and transformation in linguistic structure. *Language* 33:283–340.

Hartmann, Katharina. 2001. *Right node raising and gapping: Interface conditions on prosodic deletion*. Amsterdam: John Benjamins.

Haspelmath, Martin. 2004. Coordinating constructions: An overview. In *Coordinating constructions*, ed. Martin Haspelmath. Amsterdam: Benjamins.

Hayes, Bruce. 1984. The phonology of rhythm in English. *Linguistic Inquiry* 15:33–74.

Hayes, Bruce. 1989. The prosodic hierarchy in meter. In *Rhythm and meter*, ed. Paul Kiparsky and Gilbert Youmans, 201–260. Orlando, Fl: Academic Press.

Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Doctoral Dissertation, University of Massachussetts, Amherst.

Heim, Irene. 1991. Artikel und Definitheit. In von Stechow and Wunderlich (1991), 487–535.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Hoeksema, J. 1983. Plurality and conjunction. In *Studies in model-theoretic semantics*, ed. A. ter Meulen, 63–83. Dordrecht: Foris.

Höhle, Tilman. 1982. Explikation für 'Normale Betonung' und 'Normale Wortstellung'. In *Satzglieder des Deutschen*, ed. Werner Abaraham, 75–153. Narr.

Horn, Larry R. 1996. Exclusive company: *Only* and the dynamics of vertical inference. *Journal of Semantics* 13:1–40.

Horvath, Julia. 2004. Is "focus movement" driven by stress. *Approaches to Hungarian*

.

Hoskins, Steve. 1996. A phonetic study of focus in instransitive verb sentences. In *Proceedings of the ICSLP 96*.

Hulsey, Sarah, and Uli Sauerland. 2005. Sorting out relative clauses. Ms. MIT and ZAS, Berlin.

Hurford, James R. 1974. Exclusive or inclusive disjunction. *Foundations of Language* 410–411.

Idsardi, William. 1992. The computation of prososdy. Doctoral Dissertation, MIT.

Inkelas, Sharon. 1990. *Prosodic constituency in the lexicon*. New York: Garland.

Inkelas, Sharon, and Draga Zec, ed. 1990. *The phonology-syntax connection*. CSLI and CUP.

Ishihara, Shinishiro. 2001. Stress, focus, and scrambling in japanese. In *A few from building e39*. MITWPL.

Ishihara, Shinishiro. 2003. Intonation and interface conditions. Doctoral Dissertation, MIT.

Ito, K., S.R. Speer, and M.E. Beckman. 2004. Informational status and pitch accent distribution in spontaneous dialogues in English. In *Speech Prosody 2004*. Nara, Japan.

Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, Ma.: MIT Press.

Jacobs, Joachim. 1988. Fokus-Hintergrund-Gliederung und Grammatik. In Altmann (1988), 89–134.

Jacobs, Joachim. 1991a. Focus ambiguities. *Journal of Semantics* 8:1–36.

Jacobs, Joachim. 1991b. Negation. In von Stechow and Wunderlich (1991).

Jacobs, Joachim. 1991c. On the semantics of modal particles. In *Discourse particles*, ed. Werner Abraham, 141–162. Amsterdam/Philadelphia: Benjamins.

Jacobs, Joachim. 1992. Neutral stress and the position of heads. In *Informationsstrukur und Grammatik*, ed. Joachim Jacobs, Linguistische Berichte Sonderheft 4, 220–244. Opladen: Westdeutscher Verlag.

Jacobs, Joachim. 1999. Informational autonomy. In *Focus. linguistic, cognitive, and computational perspectives*, ed. Peter Bosch, 56–81. CUP.

Jacobs, Joachim. 2001. The dimension of topic-comment. *Linguistics* 39:641–681.

Jacobson, Pauline. 1999. Toward a variable free semantics. *Linguistics and Philosophy* 22:117–184.

Jaeger, Florian, and Elisabeth Norcliffe. 2005. Post-nuclear phrasing. Presented at the LSA Meeting, Oakland, CA.

Johnson, Kyle. 2002. Towards an etiology of adjunct islands. Umass Amherst.

Jun, S.-A. 1993. The phonetics and phonology of Korean prosody. Doctoral Dissertation, Ohio State University.

Jun, S.-A. 1998. The accentual phrase in the Korean prosodic hierarchy. *Phonology* 15:189–226.

Jurafsky, Dan, Alan Bell, Michelle Gregory, and William D. Raymond. 2000. Probabilistic relations between words: evidence from reduction in lexical production. In *Frequency in the emergence of linguistic structure*, ed. joan Bybee and Paul Hopper. Amsterdam: John Benjamines.

Kadmon, Nirit. 2003. Some theories of the interpretation of accent placement. Presented at CSSP 2003.

Kahn, David. 1976. *Syllable-based generalizations in english phonology*. New York: Garland.

Kahnemuyipour, Arsalan. 2003. Toward a phase-based theory of phrasal stress. Workshop on Phases. MIT 2003.

Kahnemuyipour, Arsalan. 2004. The syntax of sentential stress. Doctoral Dissertation, University of Toronto.

Kayne, Richard. 1983. *Connectedness and binary branching*. Dordrecht: Foris.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Ma.: MIT Press.

Kayne, Richard S. 1998. Overt vs. covert movement. *Syntax* 2:128–191.

Keating, P. to appear. Phonetic encoding of prosodic structure. In *Speech production: Models, phonetic processes and techniques*, ed. J. Harrington and M. Tabain. New York: Psychology Press.

Keating, P., T. Cho, C. Fougeron, and C-S. Hsu. 2003. Domain-initial strengthening in four languages. In Local et al. (2003), 143–161.

Kenesei, István, and Irene Vogel. 1995. Focus and phonological structure. University of Szeged, University of Delaware.

Kiparsky, Paul. 1966. Über den deutschen Wortakzent. In Bierwisch (1966b), 71–98.

Kiss, Katalin É., and Henk van Riemsdijk, ed. 2004. *Verb clusters. a study of Hungarian, German and Dutch*. Amsterdam/Philadelphia: John Benjamins.

Kisseberth, C., and M. I. Abasheikh. 1974. Vowel length in chimwi:ni—a case study of the role of grammar in phonology. In *Papers from the parasession on natural phonology*, ed. M. M. L. Galy, R. A. Fox, and A. Bruck. Chicago: Chicago Linguistic Society.

Klatt, Dennis H. 1975. Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* 3:129–140.

Ko, Heejeong. 2005. Syntactic edges and linearization. Doctoral Dissertation, MIT.

Krajic, Tanya, and Susan E. Brennan. 2005. Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology* 50:194–231.

Kratzer, Angelika. 1991. The representation of focus. In von Stechow and Wunderlich (1991), 825–834.

Kratzer, Angelika. 1995. Stage-level and individual-level predicates. In *The generic book*, ed. G.N. Carlson and F.J. Pelletier, 125–175. Chicago: Chicago University Press.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase structure in the lexicon*, ed. Johan Ryck and Laurie Zaring, 109–137. Dordrecht: Kluwer Academic Publishers.

Kratzer, Angelika. forthcoming. *The event argument and the semantics of verbs*. Cambridge, Ma.: MIT Press.

Krifka, Manfred. 1984. Fokus, Topik, syntaktische Struktur und semantische Interpretation. Universität Tübingen.

Kroch, Anthony, and Beatrice Santorini. 1987. The derived constituent structure of the west germanic verb raising constrution. In *Proceedings of the Princeton Workshop on Comparative Grammar*. Cambridge, Ma.: MIT.

Kubozono, Haruo. 1989. Syntactic and rhythmic effects on downstep in japanese. *Phonology* 6:39–67.

Kubozono, Haruo. 1992. Modeling syntactic effects on downstep in japanese. In *Papers in laboratory phonology, vol. ii: Gesture, segment, prosody*, ed. Gerald Docherty and Robert Ladd, 368–287. Cambridge: Cambridge University Press.

Ladd, Robert. 1080. *The structure of intonational meaning*. Bloomington: Indiana University Press.

Ladd, Robert. 1986. Intonational phrasing: the case for recursive prosodic structure. *Phonology Yearbook* 3:311–340.

Ladd, Robert. 1988. Declination and 'reset' and the hierarchical organziation of utterances. *JASA* 84.

Ladd, Robert D. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

Ladusaw, William A. 1994. Thetic and categorical, stage and individual, weak and strong. In *Proceedings of SALT IV*.

Lakoff, George. 1972. The global nature of the nuclear stress rule. *Language* 48:285–303.

Lambrecht, Knud. 1994. *Information structure and sentence form. topic, focus and the mental rrepresentations of discourse referents*. Cambridge University Press.

Lang, Ewald, Claudia Maienborn, and Cathrine Fabricius-Hansen, ed. 2003. *Modifying adjuncts*. Berlin: Mouton de Gruyter.

Langendoen, Terence. 1998. Limitations on embedding in coordinate structures. *Journal of Psycholinguistic Research* 27:235–259.

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19:335–391.

Larson, Richard K. 2005. Sentence final adverbs and scope. In *Proceedings of NELS 34*, ed. Keir Moulton and Matt Wolff. Stony Brook.

Larson, Richard K., and Sungeon Cho. 1999. Temporal adjectives and structure of possessive dps. In *Proceedings of WCCFL 18*. Somerville, Ma.: Cascadilla Press.

Lavoie, Lisa. 2001. *Consonant strength: Phonological patterns and phonetic manifestations*. New York: Garland.

Lehiste, Ilse. 1973. Phonetic disambigation of syntactic ambiguity. *Glossa* 7:107–122.

Lehiste, Ilse, Joseph P. Olive, and Lynn A. Streeter. 1976. Role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America* 60:1199–1202.

Liberman, M. 1975. The intonational system of English. Doctoral Dissertation, MIT.

Libermann, M., and A. Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8:249–336.

Lieberman, P. 1967. *Intonation, perception, and language*. Cambridge, Ma.: MIT Press.

Linebarger, Marcia. 1980. The grammar of negative polarity. Doctoral Dissertation, MIT.

Local, J., R. Ogden, and R. Temple, ed. 2003. *Phonetic interpretation: Papers in laboratory phonology VI*. Cambridge, UK: Cambridge University Press.

Maienborn, Claudia. 2001. On the position and interpretation of locative modifiers. *Natural Language Semantics* 9:191–240.

Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, Ma.: MIT Press.

Martin, G., James. 1970. On judging pauses in spontaneous speech. *Journal of Verbal Learning and Verbal* 9:75–78.

Martin, J.G. 1971. Some acoustic and grammatical features of sponaneous speech. In *The perception of language*, ed. D. Horton and J. Jenkins, 47–68. Columbus, OH: C.E. Merril.

Marvin, Tatjana. 2002. Topics in the stress and syntax of words. Doctoral Dissertation, Massachusetts Institute of Technology.

McGinnis, Martha. 2002. Object asymmetries in a phase theory of syntax. In *Proceedings of the 2001 CLA Annual Conference*, ed. John T. Jensen and Gerard van Herk, Cahiers Linguistiques d'Ottawa, 133–144.

Merchant, Jason. 2003. Fragments and ellipsis. To appear in Linguistics in Philosophy.

Miller, G.A., and N. Chomsky. 1963. Finitary models of language users. In *Handbook of mathematical psychology*, ed. D.R. Luce, R.R. Bush, and E. Galanter, volume II. New York: John Wiley.

Min, Young-Gie. 1996. Role of punctuation in disambiguation of coordinate compounds. University of Arizona, Tucson.

Munn, Alan. 1987. Coordinate structure and x-bar theory. *McGill Working Papers in Linguistics* 4:121–140.

Munn, Alan. 1993. Topics in the syntax and semantics of coordinate structures. Doctoral Dissertation, University of Maryland.

Munn, Alan. 2000. Three types of coordination asymmetries. In *Ellipsis in conjunction*, ed. Kerstin Schwabe and Ning Zhang, 1–22. Tübingen: Niemeyer.

Neeleman, Ad, and Tania Reinhart. 1998. Scrambling and the PF-interface. In *The projection of arguments*, ed. Miriam Butt and Wilhelm Geuder, 309–353. CSLI.

Nespor, Marina, and Irene Vogel. 1982. Prosodic domains of external sandhi rules. In *The structure phonological representations*, ed. H. van der Hulst and N. Smith, 222–255. Dordrecht: Foris.

Nespor, Marina, and Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.

Newman, S.S. 1946. On the stress system of English. *Word* 2:171–187.

Nissenbaum, Jonathan D. 2000. Investigations of covert phrase movement. Doctoral Dissertation, MIT.

Nunberg, Geoffrey. 1990. *The linguistics of punctuation*. Stanford: CSLI.

Pak, Marjorie. 2005. Explaining branchingness effects in phrasal phonology. Ms. UPenn.

Parsons, T. 1990. *Events in the semantics of english*. Cambridge: MIT Press.

Pesetsky, David. 1995. *Zero syntax. experiencers and cascades*. Cambridge, Ma: MIT Press.

Phillips, Colin. 1996. Order and structure. Doctoral Dissertation, MIT.

Phillips, Colin. 2003. Linear order and constituency. *Linguistic Inquiry* 34:37–90.

Pierrehumbert, Janet. 1980. The phonology and phonetics of english intonation. Doctoral Dissertation, MIT.

Pierrehumbert, Janet. 1999. Prosody and intonation. In *The MIT encyclopedia of the cognitive sciences*, ed. R.A. Wilson and F. Keil, 479–82. Cambridge, Ma.: MIT Press.

Pierrehumbert, Janet, and Mary E. Beckman. 1988. *Japanese tone structure*. Cambridge, Ma.: MIT Press.

Pierrehumbert, Janet, and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In *Intentions in communication*, ed. Philip R. Cohen, Jerry Morgan, and Marth E. Pollack, 271–311. Cambridge, Ma.: MIT Press.

Potts, Christopher. 2005. Lexicalized intonational meaning. In *Papers on prosody*, ed. Shigeto Kawahara, volume 30 of *University of Massachussetts Occasional Papers in Linguistics*, 129–146. Amherst, Ma.: GLSA.

Price, P.J., S. Ostendorf, S. Shattuck-Hufnagel, and C. Fong. 1991. The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America* 9:2956–2970.

Prieto, Pilar. forthcoming. Syntactic and eurythmic constraints on phrasing decisions in catalan. *Studia Linguistica* .

Prince, Alan. 1983. Relating to the grid. *Linguistic Inquiry* 14:19–100.

Progovac, Ljiljana. 2003. Structure for coordination. In *The second glot international state-of-the-article book*, ed. Lisa Cheng and Rint Sybesma, 241–287. Berlin, New York: Mouton de Gruyter.

Purnell, Thomas Clark. 1997. Principles and parameters of phonological rules: Evidence from tone languges. Doctoral Dissertation, University of Delaware.

Pylkkänen, Liina. 2002. Introducing arguments. Doctoral Dissertation, MIT.

Reinhart, Tania. 1976. *Anaphora and semantic interpretation*. London & Canberra: Croom & Helm.

Reinhart, Tania. 1995. Interface economy: Focus and markedness. Ms.

Reinhart, Tanya. 2004a. Focus—the PF-Interface. Ms. University of Utrecht.

Reinhart, Tanya. 2004b. The processing cost of reference-set computation. To appear as chapter 3 of 'Interface Strategies'. MIT Press.

Reiss, Charles. 2003. Stress computation using non-directed brackets. Ms. Concordia University.

Rochemont, Michael. 1986. *Focus in generative grammar*. Amsterdam/Philadelphia: John Benjamins.

Rogers, James. 2003. Syntactic structures as multi-dimensional trees. *Research on Language and Computation* 1:265–305.

Ronat, Mitsou. 1982. Logical form and prosodic islands. *Journal of Linguistic Research* 2:33–48.

Rooth, Mats. 1985. Association with focus. Doctoral Dissertation, University of Massachussetts, Amherst.

Rooth, Mats. 1992a. Ellipsis redundancy and reduction redundancy. In *Proceedings of the Stuttgart Ellipsis Workshop*, ed. S. Berman and A. Hestvik, number 29 in Arbeitspapiere des Sonderforschungsbereiches 340: Grundlangen für de Computerlinguistik. Universität Stuttgart/Tübingen.

Rooth, Mats. 1992b. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.

Rooth, Mats. 1999. Association with focus or association with presupposition? In *Focus*, ed. P. Bosch and R. van der Sandt, 232–244. Cambridge: Cambridge University Press.

Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral Dissertation, MIT.

Sabbagh, Joey. to appear. Ordering and linearizing rightward movement. *Natural Language and Linguistic Theory* .

Sandalo, Filomena, and Hubert Truckenbrodt. 2002. Some notes on phonological phrasing in Brasilian Portuguese. In Csirmaz et al. (2002), 285–210.

Sauerland, Uli. 1999. Why variables? In *Proceedings of nels 29*, ed. P. Tamanji, 323–337. UMass Amherst: GLSA.

Sauerland, Uli. 2005. Recursion in semantics? the case of binding. Interfaces + Recursion = Language. The view from Syntax and Semantics. Symposion at ZAS, Berlin.

Schafer, Amy J., S. Speer, , P. Warren, and S. D. White. 2000. Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research* 29:169–182.

Schein, Barry. 1993. *Plurals and events*. Cambridge, Ma.: MIT Press.

Schein, Barry. 1997. Conjunction reduction redux. chapter 1. Ms. USC.

Schein, Barry. 2001. Conjunction reduction redux. chapter 2. Ms. USC.

Schmerling, S. F. 1976. *Aspects of English sentence stress*. Austin: University of Texas.

Schwarzschild, Roger. 1996. *Pluralities*. Dordrecht: Kluwer.

Schwarzschild, Roger. 1997. Interpreting accent. Ms. Rutgers University.

Schwarzschild, Roger. 1999. Givenness, AVOIDF and other constraints on the placement of accent. *Natural Language Semantics* 7:141–177.

Seidl, Amanda. 2000. Minimal indirect reference: A theory of the syntax-phonology interface. Doctoral Dissertation, University of Pennsylvania.

Selkirk, Elisabeth. 1978. On prosodic structure and its relation to syntactic structure. In *Nordic prosody 11*, ed. T. Fretheim. Trondheim: Tapir.

Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3:371–405.

Selkirk, Elisabeth. 1995a. The prosodic structure of function words. In *Papers in optimality theory. university of massachussetts occasional papers*, ed. Jill N. Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk. Amherst, Ma.: Umass.

Selkirk, Elisabeth. 1995b. Sentence prosody: Intonation, stress, and phrasing. In Goldsmith (1995), 550–569.

Selkirk, Elisabeth. 1996. The prosodic structure of function words. In *Signal to syntax*, ed. J.L. Morgan and K. Demuth, 187–213. Mahwah, NJ: Lawrence Erlbaum Associates.

Selkirk, Elisabeth. 2002. Contrastive FOCUS vs. presentational focus: Prosodic evidence from right node raising in English. In *Speech Prosody 2002: Proceedings of the 1st International Conference on Speech Prosody, Aix-en-Provence*, 643–646.

Selkirk, Elisabeth. to appeara. Comments on intonational phrasing in English. In *Prosodides (selected papers from the conference on phonetics and phonology in Iberia, 2003*, ed. Sonia Frota, Marina Vigario, and Maria Joao Freitas. Berlin/New York: Mouton de Gruyter.

Selkirk, Elisabeth. to appearb. The prosodic phrasing of FOCUS followed by focus: The case of English. Proceedings of the TIE conference on prosody, Santorini, Greece, 2004.

Selkirk, Elisabeth, and Angelika Kratzer. 2005. Focuses, phases and phrase stress. Presented at the Mediterranean Syntax Meeting. Rhodes, June 23rd.

Selkirk, Elisabeth O. 1980. Prosodic domains in phonology: Sanskrik revisited. In *Juncture*, ed. M. Aronoff and M.-L. Keans. Saratoga, Ca.: Anma Libri.

Selkirk, Elisabeth O. 1981. On the nature of phonological representation. In *The cognitive representation of speech*, ed. J. Laver and T. Meyers. Amsterdam: North Holland.

Selkirk, Elisabeth O. 1984. *Phonology and syntax. the relation between sound and structure*. Cambridge, Ma.: MIT Press.

Sells, Peter. 2001. *Structure, alignment and optimality in swedish*. Stanford: CSLI.

Shattuck-Hufnagel, Stefanie. 2005. Final lengthening. Talk at MIT.

Shattuck-Hufnagel, Stefanie, and Alice E. Turk. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25:193–247.

Silverman, Kim, Mary Beckman, Mari Ostendorf, Colin Wightman, Patti Price, Janet Pierrehumbert, and Julia Hirschberg. 1992. ToBI: A standard for labeling English prosody. In *Proceedings of the 1992 international conference conference of spoken language processing*, volume 2, 867–870.

Sloane, N. J. A., and S. Plouffe. 1995. *The encyclopedia of integer sequences*. San Diego: Academic Press.

Snedeker, J., and J. Trueswell. 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language* 48:103–130.

Stalnaker, R. 1978. Assertion. *Syntax and Semantics* 9:315–332.

Stanley, Richard P. 1997. Hipparcus, Plutarch, Schröder, Hough. *American Mathematics Monthly* 104:344–350. Http//www-math.mit.edu/řstan/papers.html.

Starke, Michal. 2001. On the inexistence of specifiers and the nature of heads. NYU.

von Stechow, Arnim, and Dieter Wunderlich, ed. 1991. *Semantik. Ein internationales Handbuch der zeitgenössischen Forschung. (HSK 6)*. Berlin: de Gruyter.

Steedman, Mark. 1985. Dependency and coordination in the grammar of Dutch and English. *Language* 61:523–568.

Steedman, Mark. 2000. Information structure and the phonology-syntax interface. *Linguistic Inquiry* 34.

Steedman, Mark. 2001. *The syntactic process*. Cambridge, Ma.: MIT Press.

Steedman, Mark. 2004. Surface compositional scope-alternation without existential quantifiers. Ms. University of Edinburgh.

Stockwell, R.P. 1960. The place of intonation in a generative grammar. *Language* 36:360–367.

Stockwell, R.P. 1972. The role of intonation: Reconsiderations and other considerations. In *Readings on intonation*, ed. Dwight Bolinger, 87–109. Harmonsworth, England: Penguin.

Szendröi, Krista. 2001. Focus and the syntax-phonology interface. Doctoral Dissertation, University College London.

Taglicht, Josef. 1984. *Message and emphasis. on focus and scope in English*, volume 15 of *English Language Series*. London and New York: Longman.

Taglicht, Josef. 1998. Constraints on intonational phrasing in English. *Journal of Linguistics* 34:181–211.

Takahashi, Shoichi. 2005. Pseudogapping and cyclic linearization. In *Proceedigns of NELS 34*, 571–585. Amherst, Ma.: GLSA.

Truckenbrodt, Hubert. 1993. Syntax vs. phonology: Which gets the stress right? Ms. MIT.

Truckenbrodt, Hubert. 1995. Phonological phrases: Their relation to syntax, focus, and prominence. Doctoral Dissertation, MIT, Cambridge, Mass.

Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30:219–255.

Uhmann, Susanne. 1991. *Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie*. Tübingen: Niemeyer.

Uriagareka, Juan. 1999. Multiple spell-out. In *Working minimalism*, ed. Samual David Epstein and Nobert Hornstein, 251–283. Cambridge, Ma.: MIT Press.

Wagner, Michael. 2002a. Configurational stress in derivatives, compounds, and phrases. Paper presented at the Workshop on Pertinacity, University of Konstanz. Ms. at web.mit.edu/chael.

Wagner, Michael. 2002b. Holmberg's generalization in an OV-language. Poster presented at NELS 33. MIT 2002.

Wagner, Michael. 2004a. Asymmetries in prosodic domain formation. To appear in Proceedings of Workshop on Phases, MIT, January 2003.

Wagner, Michael. 2004b. Asymmetries in the syntax and prosody of verb-intial interpolated clauses in German. In *Proceedings of Console XII, Patras 2003*.

Watson, Duane, and Edward Gibson. 2004a. Intonational phrasing and constituency in the language production and comprehension. *Studia Linguistica* In press.

Watson, Duane, and Edward Gibson. 2004b. The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes* In press.

Welby, Pauline. 2003. Effects of pitch accent position, type, and status on focus projection. *Language and Speech* 46:53–81.

Wightman, C.W., S. Shattuck-Hufnagel, M. Ostendorf, and P.J. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 92:1707–1717.

Williams, Edwin. 1980. Remarks on stress and anaphora. *Journal of Lingusitic Research* 1:1–16.

Williams, Edwin. 1997. Blocking and anaphora. *Linguistic Inquiry* 28:577–628.

Williams, Edwin. 2003. *Representation theory*. Cambridge: MIT Press.

Winter, Yoad. 2001. *Flexibility principles in boolean semantics*. Cambridge, Ma.: MIT Press.

Wurmbrand, Susi. 2003. Verb-clusters in West-Germanic: The emipirical domain. In *The verb-cluster Sprachbund: Verb clusters in Germanic and Hungarian*, ed. Katalin É. Kiss and Henk van Riemsdijk. Amsterdam: John Benhamins.

Zec, Draga. 2002. On the prosodic status of function words. *Working papers of the Cornell Phonetics Laboratory* 14.

Zec, Draga, and Sharon Inkelas. 1990. Prosodically constrained syntax. In Inkelas and Zec (1990).

Zoerner, Edward Cyril. 1995. Coordination: The syntax of &P. Doctoral Dissertation, University of California, Irvine.

Zubizarreta, M. L. 1998. *Prosody, focus, and word order*. Cambridge, Ma.: MIT Press.