RETHINKING SYNTACTOCENTRISM

LESSONS FROM RECENT GENERATIVE APPROACHES TO PRAGMATIC PROPERTIES OF LEFT-PERIPHERY-MOVEMENT IN GERMAN

Inaugural-Dissertation

zur

Erlangung der Doktorwürde

der Philologischen Fakultät

der Albert-Ludwigs-Universität

Freiburg i. Br.

vorgelegt von

Andreas Trotzke

aus Duisburg

SS 2010

Contents

0. Introduction

> [W]hat is called for is an open-mindedness to insights from whatever quarter […] and a joint commitment to fight fair in the interests of deeper understanding. To my mind, that's what the game of science is about. (Jackendoff 2002: xiii)
>
> [T]he syntactocentric model […] was explicitly only an assumption, which quickly hardened into dogma and then became part of the unstated background. (Jackendoff 2003: 659)

The linguist Ray Jackendoff never tires to call for open-mindedness and fairness in the heterogeneous area of research on the mental foundations of language. One of the approaches within this field claims that syntax, as regarded in linguistics à la Noam Chomsky, plays the central role in modeling the mental architecture of the human language faculty. According to Jackendoff, this conception has 'hardened into dogma' and thus is not amenable to any insights from other 'quarters.' To sharpen this claim, he invented the term 'syntactocentrism' and thereby suggests that this concept of generative grammar is more of an ideology, an 'ism,' than an approach that lends itself to participate in the open-minded 'fights' taking place in science.

In this thesis, I will explore to what extent this characterization is justified both by evaluating Jackendoff's notion of syntactocentrism in light of recent models of mainstream generative grammar and by asking what lessons can be learned from applying these recent conceptions to a specific phenomenon of German.

In chapter 1, in order to shed some light on the 'unstated background' Jackendoff refers to, I will ground the concept of syntactocentrism by outlining basic beliefs concerning the mental architecture of the language faculty held at the time the conception of syntactocentrism was introduced. In doing so, I will present the general idea of this model and also clarify why many scholars, including Jackendoff, regard the early version of this conception as a reasonable view due to the historical context it emerged in.

In chapter 2, I will turn to Jackendoff's claim that the syntactocentric view of grammar, although quite reasonable in the 1960s, is now obsolete and can only be adhered to by ignoring progress in both phonology and semantics. To examine this impression, I will survey the format of recent syntactocentrism by first looking at the changed conception of syntax and then turning to current approaches to phonology and semantics that are based on this concept.

Chapter 3 deals with the different grammar models of syntactocentrism that Jackendoff discusses. It focuses on the recent shift from 'representational' to 'derivational' syntactocentrism. To outline this change, I will first illustrate the view of syntax as a mental entity enriched with different levels of representation. Then, in order to arrive at an overall picture of recent syntactocentrism, I will sketch how this model has been successively abandoned and replaced by a fairly-reduced conception of syntactic operations.

In chapter 4, I will turn to prominent theoretical alternatives to syntactocentrism and first concentrate on some basic ideas within the general movement of Cognitive Linguistics. Having illustrated this view, which contradicts mainstream generative grammar in many ways, I will look at Jackendoff's own approach, the 'Parallel Architecture,' which can be regarded as an intermediate position between the two extremes of Cognitive Linguistics and syntactocentrism. In the final section of this chapter, I will reflect on the question whether there is any perspective of convergence between syntactocentrism and its theoretical alternatives.

In chapter 5, based on this reflection, I will explore the conjecture that, once the consequences of recent derivational syntactocentrism are taken seriously, some of Jackendoff's objections to this perspective on language disappear. To investigate this hypothesis, I will reduce the comparison of recent syntactocentrism and the Parallel Architecture to tractable size and thus focus on the analysis of one specific phenomenon, namely the pragmatics of left-periphery-movement in German. In doing so, I will contrast the representational analysis of this phenomenon with a strong derivational account and then, based on the differences that emerge, compare this more recent derivational analysis with the conceptual underpinnings of accounting for this phenomenon within the framework of the Parallel Architecture.

Having thus exemplified recent syntactocentrism by outlining a concrete derivational analysis, chapter 6 addresses the remaining question whether there are points of convergence even between recent syntactocentrism and Cognitive Linguistics. Since in this case, in contrast to comparing syntactocentrism with the Parallel Architecture, there is less potential for convergence at the 'microscopic,' descriptive level, I will turn to more high-level issues and compare both approaches regarding the issue of language evolution. To undertake this comparison, I will first sketch an approach to the evolutionary origins on language that is associated with the strong derivational view on syntactic computations exemplified in the context of left-

periphery-movement in German. After that, I will illustrate an approach that concentrates on language as a communicative system and crucially rests on concepts of Cognitive Linguistics. Finally, I will compare these two approaches and look for points of convergence.

In chapter 7, I conclude by summarizing the main results of this thesis and by turning to the question whether these results vindicate the notion of syntactocentrism as used by Jackendoff.

Before I start with grounding the notion of syntactocentrism, let me add a cautionary note. This thesis rests on the assumption that both syntactocentrism and its theoretical alternatives – be it the Parallel Architecture or Cognitive Linguistics – belong to one single paradigm within linguistics that investigates language as a mental entity. Since the issues addressed in this thesis inherently require an extensive discussion of concepts assumed within mainstream generative grammar, I ask those readers that are more committed to the non-generative alternatives to practice the open-mindedness mentioned at the outset of this introduction and to adopt the view that there is no litmus test to determine membership in the category of mentalist linguistics. Rather, as with other categories, the different approaches, even if disagreeing in various respects, are connected by family resemblance – a concept well known in some branches of non-generative linguistics. Accordingly, looking into generative linguistics may promote an understanding of the category as a whole, even if only in the sense of sharpening one's own account.

1. Grounding Syntactocentrism

Jackendoff has repeatedly pointed out that the "assumption of 'syntactocentrism' […] was never explicitly grounded" (Jackendoff 2003: 655), that is, according to him, this concept has been introduced without serious argument. And indeed, when Chomsky launched this special perspective on the mental architecture of the human language faculty in the late 1950s, it was a new approach and thus, as a matter of fact, explicitly marked as a tentative assumption. However, some crucial concepts underlying the syntactocentric view were anything but new. In order to approach the notion of syntactocentrism, let me first clarify the general framework this conception is situated in.

What was not entirely new and what Jackendoff himself has subscribed to over the years is the mentalist perspective on language resting on a "'capital of ideas' accumulated in the premodern period" (Chomsky 1966: 3). In his attempt to trace back the historical roots of this perspective, Chomsky especially refers to Descartes, who denied that the soul of animals were of the same kind as ours. According to Descartes, this crucial difference between man and animal manifests itself most clearly in the fact that an animal "never […] arranges its speech in various ways […] in order to reply appropriately to everything that may be said in its presence, as even the lowest type of man can do" (Descartes 2003 [1637]: 38). Like Descartes in his reflection on human uniqueness, Chomsky places a premium on this capacity, to which he refers as the "'creative aspect' of ordinary language use" (Chomsky 1966: 4-5). To explore this aspect, adopting the mentalist view of Descartes, Chomsky assumes that a language user must have a mental capacity that enables this 'creative aspect' of language use. So, in contrast to the actual use of language in concrete situations, dubbed 'performance,' this mental capacity was referred to as 'competence,' as "the speaker-hearer's knowledge of his language" (Chomsky 1965: 4). Whatever the precise nature and format of the distinction between competence and performance, many approaches since the 'cognitive revolution,' which was inaugurated in the late 1950s, including Jackendoff's own theory, are committed to the view "that it is essential to consider language as a cognitive (mental) system" (Goldberg 2006: 4). However, controversies emerge with the exact formulation of this 'grammatical knowledge.' And here is where the notion of syntactocentrism comes into play.

Chomsky conceived of the speaker's knowledge as containing both a finite set of symbols, out of which sentences can be constructed, and a finite amount of combinatorial operations, a "system of rules that we can call the *grammar* of his language" (Chomsky 1964: 9, emphasis in the original). Assuming this general view of grammar, Chomsky formulates the syntactocentric claim that "a grammar contains a syntactic component, a semantic component, and a phonological component. The latter two are purely interpretive" (Chomsky 1965: 141). In retrospective, Jackendoff (2003: 655) concedes that "[i]n 1965 this was a perfectly reasonable view" but only, as he goes on, in absence of any in-depth analysis of phonology and semantics. However, pace Jackendoff's remarkable knowledge of this period, to me, it seems overdone to state that "[a]s for semantics, virtually nothing was known […] and […] the sound system of language had been regarded essentially as a sequence of speech sounds" (Jackendoff 2003: 655).

In this chapter, I will slightly amend Jackendoff's statement by grounding the assumption of syntactocentrism, not in the sense of providing the forceful arguments Jackendoff misses, but more in the sense of outlining basic beliefs, concerning the syntactic, the phonological, and the semantic components, held at the time when the conception of syntactocentrism has been introduced. Following Chomsky's (1965) threefold definition of grammar given above, after introducing basic aspects of the syntactic component in section 1.1, I will turn to the phonological component and show in what sense it is regarded as 'purely interpretive.' Finally, in section 1.3, I will briefly sketch the early generative conception of the semantic component.

1.1 The Syntactic Component

In the early days of generative grammar, the focus was on the development of rule systems that possess the appropriate computational properties to account for the 'creative aspect' of language use. These rule systems had to be finite, since the mental resources of humans are limited, but they still should capture the faculty of producing and understanding an indefinite number of sentences in an indefinite range of new situations, and thus, they should provide a formal account of the human capacity to "make infinite employment of finite means" (Humboldt 1999 [1836]: 91). To formulate such rule systems, Chomsky followed Bar-Hillel (1953: 165), who

argues for „evaluation of [...] recent investigations on recursive definitions" within empirical sciences such as linguistics. Accordingly, Chomsky (1957) discusses different finite rule systems that operate with recursive procedures, that is, with loop-like devices that allow rules to apply to their own output (for an overview of Chomsky's early discussion of adequate rule systems, see Lasnik 2000: 12-23).

The first computational device discussed by Chomsky is a finite-state machine. To understand this type of grammar, consider, for example, how this device accounts for the following sentence (cf. Chomsky 1957: 19-20):

(1)   *The man comes.*

Using a finite-state machine to model the generation of sentences like (1), we can represent the grammar graphically in the form of a so-called 'state diagram.' In such a diagram, or 'graph,' the generation of a structure proceeds from an initial state to a final state in the direction of the arrows, where "[t]he 'states' are the junction points in the graph and the […] letters produced for a transition are given beside the corresponding line" (Shannon & Weaver 1949: 15):

(2)   *the   man   comes*
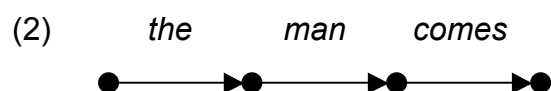
In order to generate an infinite number of sentences, Chomsky extends this grammar by adding closed loops, as shown in (3):

(3)   *old*
       the   man   comes
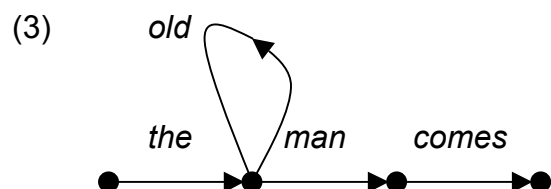
Due to this loop-like device that gives rise to recursion, the grammar can generate an infinite number of expressions like (4):

(4)   *The (old, old, old…) man comes.*

However, Chomsky (1957: 21-25) argues that finite-state grammars are formally incapable of modeling natural languages like English because such rule systems cannot account for so-called non-local dependencies. As a simple example of this crucial aspect of natural languages, consider cases in which the subject of a sentence (indicated with 'a') and the verb (marked with 'b'), though morpho-syntactically connected by agreement, can be far away from each other, as in the case of *the man* and *comes* in (5):

(5)   [*The man*]a [*the dog*]a [*bit*]b [*comes*]b.

In addition to examples in which two *a*s follow two *b*s, consider now the fact that natural languages allow for infinite embedding. That is, we can extend our structure (5) further by adding another sentence inside the sentence *the dog bit*, as shown in (6):

(6)   [*The man*]a [*the dog*]a [*the girl*]a [*loves*]b [*bit*]b [*comes*]b.

As should be clear from the illustration so far, center-embedded structures like (5) and (6) show the general property of "*n* occurrences of *a* followed by *n* occurrences of *b* and only these" (Chomsky 1957: 21). Crucially, a finite-state grammar cannot correspond to that property because it computes a sequence strictly local, that is, it only 'knows' what state it is currently in and what to do next. Consequently, it does not 'know' what states it has been in, let alone how many times it has been in some particular state. In other words, the finite-state grammar is ignorant of the number of *a*s and *b*s it has already generated, and, therefore, it cannot ensure an equal number of *a*s and *b*s.

To account for cases that cannot be described in terms of a finite-state grammar, Chomsky (1957: 26-33) discusses a rewrite-rule system. Adopting Chomsky's notation, we can formulate the following rules to capture our example (1):

(7) S → NP + VP   (S = Sentence; NP = Noun Phrase; VP = Verb Phrase)

NP → Det + N   (Det = Determiner)

VP → V

Det → *the*

N → *man*

V → *comes*

The rules given in (7) consist of one symbol on the left side, followed by an arrow (standing for 'rewrite as'), followed by at least one symbol. After we have applied these rules to generate our sentence (1), the derivation of the sentence proceeds from step (i) to step (vi), as (8) spells out in detail. In particular, every step of the derivation consists of rewriting one symbol by another (sequence of) symbol(s).

(8) S

NP + VP   (i)

Det + N + VP   (ii)

Det + N + V   (iii)

the + N + V   (iv)

the + man + V   (v)

the + man + comes   (vi)

Now, recall the more complex structures in (5) and (6). In order to account for these cases within a system of rewrite rules, we only have to replace the rule for NPs by a rule that reintroduces S, as shown in (9):

(9) NP → Det + N + S

By reintroducing an abstract symbol like 'S' into the derivation, the rewrite-rule system, like the finite-state grammar sketched above, contains a loop-like device and thus allows for recursion. Specifically, according to (9), the NP *the man*, containing a determiner and a noun, can be extended by a whole sentence (*the dog bit*). As this sentence contains its own NP (*the dog*), the rule given in (9) can apply again and thus extend the NP *the dog* by adding another sentence (*the girl loves*). Crucially

now, in contrast to the loop-like device of a finite-state machine mentioned above, the rewrite rules capture the long-distance dependencies of sentences like (5) and (6) by ensuring that the number of *a*s corresponds to the number of *b*s. The essential difference lies in the use of non-terminal symbols like 'S.' A symbol like 'S' introduces *a*s like *the dog* and *b*s like *bit* into the derivation simultaneously. Accordingly, unlike finite-state grammars, the rewrite-rule system, due to this simultaneous insertion, makes sure that it generates an equal number of *a*s and *b*s, that is, in our case, an equal number of NPs and VPs.

Let us stop here. Of course, as noted by Chomsky (1957) himself, the rule system presented so far still cannot account for significant structural aspects of natural languages.[1] However, this short illustration suffices to point out that rewrite rules, since they both allow for recursive embedding and capture aspects like non-local dependencies, can be regarded as the first formal approximation to a description of the 'creative aspect' of human language within generative grammar. Having sketched this early conception of the syntactic component, we now possess the conceptual basis to turn to the syntactocentric claim that the semantic and phonological components "play no part in the recursive generation of sentence structures" (Chomsky 1965: 141), and, accordingly, they can be conceived of as merely interpreting the structures generated by the syntactic component. How exactly these two 'purely interpretive' components were thought of at the time when Chomsky (1965) formulated this syntactocentric conception, is a topic I will address in the next two sections. Let me begin with the phonological component.

1.2 The Phonological Component

As we saw in the preceding section, the syntactic component, as conceived of in early generative linguistics, contains rewrite rules that both allow for an infinite range of expressions and account for structural relations, like discontinuous dependencies. Now, as in the case of the syntactic component, the generative perspective on the phonological component is also concerned with the 'creative aspect' of language

---

[1] Specifically, although rewrite rules can account for discontinuous dependencies, they do not capture so-called cross-serial dependencies. Discussing one example of this type of structural configurations, Chomsky (1957: 61-84) argues that an adequate grammar must not only contain rewrite rules but also certain transformational operations.

use. Specifically, just as the syntactic component aims at explaining the fact that speakers can form (an infinite number of) new sentences out of a finite amount of symbols and rules, theoretical accounts of the phonological component, as Chomsky et al. (1956: 79-80) point out, should provide a "simple explanation for the fact that native speakers can assign stress patterns to new utterances in a fairly consistent and uniform manner." Chomsky et al. claim that the constituent organization phonological rules operate on correlates in many ways with the constituent structure generated by the syntactic component. By assuming this correlation, Chomsky et al. (1956: 78) emphasize that "[t]his correspondence leads to an overall simplification of the grammar of the language, since the constituent structure once stated can be made to serve a variety of functions." To make this seminal formulation of a syntactocentric perspective on the relation between syntax and phonology concrete, let me illustrate how this simplification, favored by generative linguistics, finds expression in the analysis of prosodic properties of sentences.

In their voluminous study of English sound structure, Chomsky & Halle (1968) are concerned with formulating specific rules of the phonological component and with the crucial question of how these rules are organized with respect to each other.[2] In particular, following first suggestions of Chomsky et al. (1956), they propose a general principle for applying phonological rules: the "transformational cycle" (Chomsky & Halle 1968: 15). Before I will give an illustrative example of how exactly this principle applies, let me first clear up the terms 'cycle' and 'transformational.' The rules of the phonological component are understood to apply cyclically because they proceed step by step to different domains of applicability. In particular, a domain containing a subdomain can only be addressed by rules, when the rules have already applied to the subdomain. This 'cycle' is 'transformational,' inasmuch as the domains of applicability are determined by the hierarchical phrase structure of a string rather than by the linear sequence of elementary symbols constituting the string. Accordingly, the attribute 'transformational' refers to Chomsky's (1965: 89) notion of transformations according to which a "grammatical transformation is […] a rule that applies to Phrase-markers rather than to strings in the terminal and nonterminal vocabulary of

---

[2] It should be mentioned that Chomsky & Halle (1968) are primarily concerned with stress, merely one aspect of the phonological component. So, for instance, they omitted pitch from consideration, since, at that time, they were faced "with the still open question of the systematic role of pitch contours or levels within the general framework of syntactic and phonological theory as we […] understand it" (Chomsky & Halle 1968: ix).

the grammar." So, to put it simply, phonological rules apply step by step to units of different levels of complexity, and the units they apply to are determined by the hierarchy of phrase structure rather than by the linearity of symbols.

After having clarified terminology, let us now turn to some examples of phonological rule application given by Chomsky & Halle (1968: 16-23). As a first step, assume that the vowel receives primary stress in all monosyllables, as shown in (10), where '1' stands for primary stress and the label 'A' for the lexical category 'adjective:'

(10) $[_A \overset{1}{black}]$

As a next step, consider the following examples, in which the monosyllabic *black* is part of a more complex constituent. More specifically, in (11) it is part of the compound noun (N) *blackboard* and in (12) it is part of the complex noun phrase (NP) *black board*. In a first step, the monosyllables in (11) and in (12), according to the rule mentioned above, receive primary stress on their vowels:

(11) $[_N [_A \overset{1}{black}] [_N \overset{1}{board}]]$

(12) $[_{NP} [_A \overset{1}{black}] [_N \overset{1}{board}]]$

As (11) and (12) indicate, rules of the phonological component first apply, in accordance with the principle of the 'transformational cycle,' to the smallest subdomains of a given string, in our case to the monosyllables *black* and *board*. However, since the theory proposed by Chomsky & Halle (1968) aims at a prosodic description of the whole compound *blackboard* and the whole NP *black board*, phonological rules do not stop here. Rather, Chomsky & Halle claim that the phonological rules to follow should eliminate syntactic structure that is not relevant for the linear dimension of phonology. In other words, the syntactic bracketing, accounting for the hierarchical dimension of sentences, must be deleted step by step. Therefore, after rule application to the subdomains *black* and *board*, the innermost brackets are erased, as shown in (13) and (14):

(13) $[_N$ *black board*$]$
$\quad\quad\quad\;\;$ 1 $\quad\;$ 1

(14) $[_{NP}$ *black board*$]$
$\quad\quad\quad\quad$ 1 $\quad\;$ 1

Chomsky & Halle postulate two different rules applying to cases like (13) and (14). In case of (13), the 'Compound Rule' "assigns primary stress to a primary-stressed vowel which is *followed* by another primary-stressed vowel in a noun" (Chomsky & Halle 1968: 17, emphasis in the original). On the other hand, in case of complex noun phrases like (14), the 'Nuclear Stress Rule' "assigns primary stress to a pri-mary-stressed vowel which is *preceded* by another primary-stressed vowel in a noun phrase" (Chomsky & Halle 1968: 17, emphasis in the original). Applying these rules to the strings given in (13) and (14) results in the following representations, where '2' indicates weakened stress:

(15) $[_N$ *black board*$]$
$\quad\quad\quad\;\;$ 1 $\quad\;$ 2

(16) $[_{NP}$ *black board*$]$
$\quad\quad\quad\quad$ 2 $\quad\;$ 1

As these examples show, the compound *blackboard*, with a falling stress contour (1-2), is prosodically distinguished from the noun phrase *black board*, which shows a rising contour (2-1). Now, assuming the Compound Rule, the Nuclear Stress Rule, and the basic rule of assigning primary stress to vowels within monosyllables, the different stress contours of the even more complex constituents *black board-eraser* ('board eraser that is black') and *blackboard eraser* ('eraser for a blackboard') can be accounted for by the following derivations, in which rules are applied from step (i) to step (iii) in accordance with the transformational cycle. Let us first examine the deri-vation of *black board-eraser*:

(17) $[_{NP}$ $[_A$ *black*$]$ $[_N$ $[_N$ *board*$]$ $[_N$ *eraser*$]]]$

1	1	1	(i)
	1	2	(ii)
2	1	3	(iii)

According to this representation, the rule application in the first cycle assigns primary stress ('1') to *black*, *board*, and *eraser*. While *black* and *board* only contain one vowel and primary stress can only go to that vowel, the word *eraser* is itself a complex form containing three vowels. To capture such cases, Chomsky & Halle (1968) modify the rule for monosyllables mentioned above. In particular, they state that rules also apply cyclically within derivatives. In our case, with *eraser* consisting of the verb *erase* and the agentive affix *-r*, the rules first apply to *erase*. In such cases, as Chomsky & Halle (1968: 28-43) show, primary stress is placed on the final vowel. Since in monosyllables like *black* the only vowel is, of course, also the 'final' vowel, this rule holds true for both, for cases like *black* and for cases like *eraser*. After rule application has assigned primary stress to the final vowels of *black*, *board*, and *eraser*, the syntactic structure is reduced by erasing the innermost brackets, as shown in (18):

(18)   [$_{NP}$ [$_A$ *black*] [$_N$ *board eraser*]]

Consequently, the string under consideration in the second cycle, which is the next subdomain rules apply to, can be represented as follows:

$$\overset{1}{}\qquad\overset{1}{}$$

(19)   [$_N$ *board eraser*]

In the second cycle, the Compound Rule assigns primary stress on *board*, since *board* is followed by another primary-stressed vowel in a noun. After the application of this rule, the erasing of innermost brackets concludes the second cycle, resulting in the string (20), which corresponds to step (ii) in (17):

(20)   [$_{NP}$ $\overset{1}{black}$ $\overset{1}{board}$ $\overset{2}{eraser}$]

Given this string, the Nuclear Stress Rule applies and assigns primary stress on *board*, since board is preceded by another primary-stressed vowel in a noun phrase. As a result, all other stresses are weakened by one and the last syntactic brackets are erased, as the final representation, correlating with step (iii) in (17), shows:

$$\overset{2\quad\;\;1\quad\;\;3}{}$$

(21)   *black board-eraser*

After having considered the derivation of the noun phrase *black board-eraser*, let us finally, for the sake of contrast, turn to the derivation of stress contour for the noun *blackboard eraser*, summarized in (22):

(22)   [N [N [A *black*] [N *board*]] [N *eraser*]]

1	1	1	(i)
1	2		(ii)
1	3	2	(iii)

While this derivation has the same first cycle as the derivation given in (17), the string under consideration in the second cycle – due to differences in syntactic bracketing – is the noun *blackboard* rather than the noun *board-eraser*. So, the Compound Rule applies, assigning primary stress on *black*, since *black* is followed by another primary-stressed vowel in a noun. Thus, stress on *board* is weakened by one. After the deletion of the innermost brackets, the string in the next cycle is (23):

$$\overset{1\quad\;\;2\quad\;\;1}{}$$

(23)   [N *black board eraser*]

As this string is a noun rather than a complex noun phrase, the Compound Rule applies again, placing primary stress on *black*, since it is followed by another primary-stressed vowel in a noun. Consequently, all other stresses are weakened by one, resulting in the final contour (24), which corresponds to step (iii) given in (22) above:

$$\overset{1\quad\;\;3\quad\;\;2}{}$$

(24)   *blackboard eraser*

From the illustration so far, it should become clear that the phonological component of the grammar, as conceived of in the early days of generative linguistics, assigns a phonetic interpretation (in the cases discussed above: a specific stress contour) to syntactic structures by referring to properties of syntactic bracketing. In particular, the derivations given above show that the correct cyclic application of phonological

rules (like the Compound Rule or the Nuclear Stress Rule), and thus the resulting stress contours are determined by syntactic bracketing of strings and by specific labels on the brackets (e.g. N or NP). So, under this hypothesis, the phonological component merely interprets the structures generated by the syntactic component. This results in the general syntactocentric assumption that "[o]nce the speaker has selected a sentence with a particular syntactic structure and certain lexical items […], the choice of stress contour is not a matter subject to further independent decision" (Chomsky & Halle 1968: 25). In sum, even though I have only scratched the surface of the detailed processes observed within early generative approaches to phonology, the presentation so far may help to bring out the early perspective on phonology that served as the theoretical background of Chomsky's (1965) claim that the phonological component of the grammar can be conceived of as 'purely interpretive.' Having shown the syntactocentric relation between syntax and phonology by having explicated some concepts Chomsky (1965) tacitly assumes, let me now, in a similar vein, turn to the semantic component.

1.3 The Semantic Component

In his discussion of how sentences are interpreted semantically, Chomsky (1965: 136) refers to Katz & Fodor (1963) as the "first relatively clear formulation […] of the theory of semantic interpretation." Much like Chomsky, Katz & Fodor (1963) are primarily concerned with the 'creative aspect' of language use. So, they suppose that, since

> the set of sentences is infinite and each sentence is a different concatenation of morphemes, the fact that a speaker can understand any sentence must mean that the way he understands sentences which he has never previously encountered is compositional: on the basis of his knowledge of the grammatical properties and the meanings of the morphemes of the language, the rules which the speaker knows enable him to determine the meaning of a novel sentence in terms of the manner in which the parts of the sentence are composed to form the whole. (Katz & Fodor 1963: 171-172)

The central assumption behind the view expressed in the preceding quote is that the speaker obtains the meaning of any sentence from the meanings of its lexical items and from the way in which they are combined. The opinion that sentence meaning is

a compositional function of the meanings of the lexical items is shared by most formal logicians and goes back at least to Frege, who speculates that "the possibility of our understanding sentences which we have never heard before rests evidently on this, that we can construct the sense of a sentence out of parts that correspond to words" (Frege 1980 [1914]: 79).[3] Crucially now, since Katz & Fodor point out that a speaker must be able to understand, that is, he must be capable of semantically interpreting sentences he has never encountered before, compositionality seems to be a conceptual necessity for humans with finite mental resources because semantic interpretation must, as a consequence, be based on certain fixed combinatorial rules, allowing for the interpretation of an, in principle, infinite range of sentences. Again, as we saw in the context of the phonological component, the most desirable result in view of early generative linguistics is to make the recursive operation and the resulting structures that are postulated for the syntactic component serve as a description and, ultimately, as an explanation of the infinity also involved in semantic interpretation. Let us see how this works in detail.

As a first step, Katz & Fodor (1963: 181-183) point out several cases that demonstrate that the structural descriptions assigned by the syntactic component do not suffice to specify the meaning of sentences. As a first approximation to identify what additional components are necessary, consider the (admittedly trivial) fact that the syntactic component provides identical structures for sentences that are different in meaning, as shown in (25):

(25) a. *The dog bit the man.*

     b. *The cat bit the man.*

As (25) makes clear, one component of semantic interpretation has to be a dictionary of a single language because the syntactic component cannot account for the fact that sentences like (25), which differ only morphemically, are interpreted as different in meaning. However, as Katz & Fodor show, the syntactic structure and the

---

[3] Note, however, that some qualification is in order, if one adopts the common usage of referring to compositionality as 'Frege's principle,' inasmuch as Frege never mentioned compositionality as a principle and, accordingly, "[c]ompositionality is not Frege's, but it might be called 'Fregean' because it is in the spirit of his later writings" (Janssen 1997: 421).

dictionary are not enough to interpret sentences accurately. Consider the next examples (cf. Katz & Fodor 1963: 182):

(26)   a.   *Seal the letter.*

       b.   *The seal is on the letter.*

       c.   *One of the oil seals in my car is leaking.*

In case of (26a), the constituent *seal* is marked syntactically as a verb, whereas in (26b) it is marked syntactically as a noun. Accordingly, in this case, the dictionary together with the syntactic structure suffice to distinguish the meaning of *seal*, insofar as in (26a) it refers to an 'action' and in (26b) it denotes an 'object.' However, since the dictionary supplies more senses of *seal* used as a noun than it bears in an occurrence like (26b), the meaning in (26b) cannot be inferred from the dictionary and the syntactic structure alone. This becomes clear in light of cases like (26c), where a 'less prominent' sense of *seal* is used ('object used to prevent liquid substance from escaping'). Since this choice between the senses in (26b) and (26c) can only be derived from the semantic relations between lexical items in a sentence, Katz & Fodor claim that the speaker needs rules that take into account the semantic relations between morphemes, when he selects readings provided by the dictionary. Katz & Fodor (1963: 183) call these rules "projection rules." Assuming that the meaning of a sentence is a function of the meanings of its parts, Katz & Fodor state that these projection rules explicate just this function. So, in addition to the dictionary component, these rules constitute another component of semantic interpretation: the projection-rule component. As I did for the phonological component, let me give a short illustration of how the semantic component operates with these projection rules.

   As already mentioned, the structural description generated by the syntactic component is, according to a syntactocentric view, hypothesized to provide the input for the semantic component. So, for illustration purposes, let us assume the structural description given in (28) for the following sentence (cf. Katz & Fodor 1963: 197):

(27)   *The man hits the colorful ball.*

(28)

```
                          S
                 ┌────────┴────────┐
               NPc                 VP
             ┌──┴──┐          ┌─────┴─────┐
            Det   Nc          V          NPc
             │     │          │      ┌────┴────┐
             │     │          │     Det       Nc
             │     │          │      │      ┌──┴──┐
             │     │          │      │      A    Nc
             │     │          │      │      │     │
            the   man       hits   the  colorful ball
```
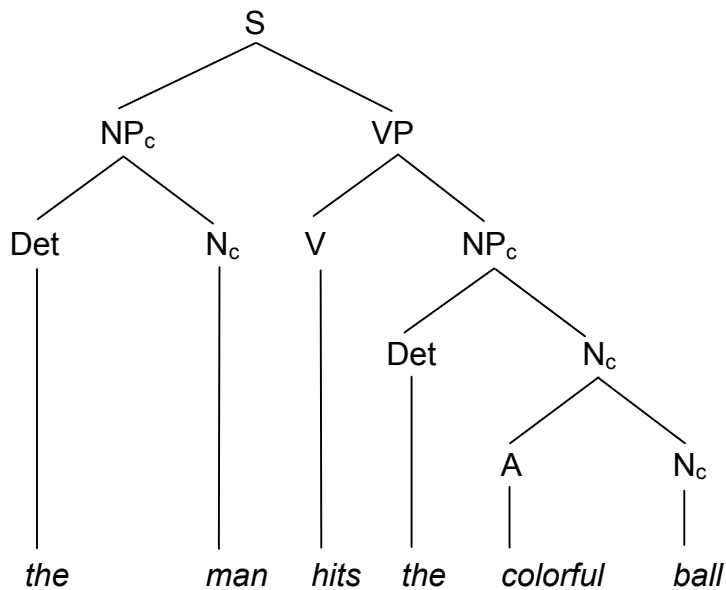
As a starting point, Katz & Fodor assume that the dictionary component chooses only those readings of the lexical items that correspond to the information already given in the structural description (28). So, for instance, the syntactic marking of the lexical item *ball* as a concrete noun (N_c) excludes both the association with readings of *ball* as a verb and the sense of *ball* in, for instance, *He plays better ball than me*, where *ball* is not a concrete noun but refers to some abstract entity ('the game'). After we narrowed down the options of interpretation of the lexical items to certain fixed sets of readings, the input to the projection rule component contains the syntactically-generated structural description and certain sets of readings, called 'paths' (P) of interpretation, as schematized in (29), where the two occurrences of *the* in (28) correspond to two instantiations of P_1 (cf. Katz & Fodor 1963: 197):

(29)

```
                          S
                 ┌────────┴────────┐
               NPc                 VP
             ┌──┴──┐          ┌─────┴─────┐
            P1     P2        P3          NPc
                                    ┌─────┴─────┐
                                   P1           Nc
                                            ┌───┴───┐
                                           P4       P5
```
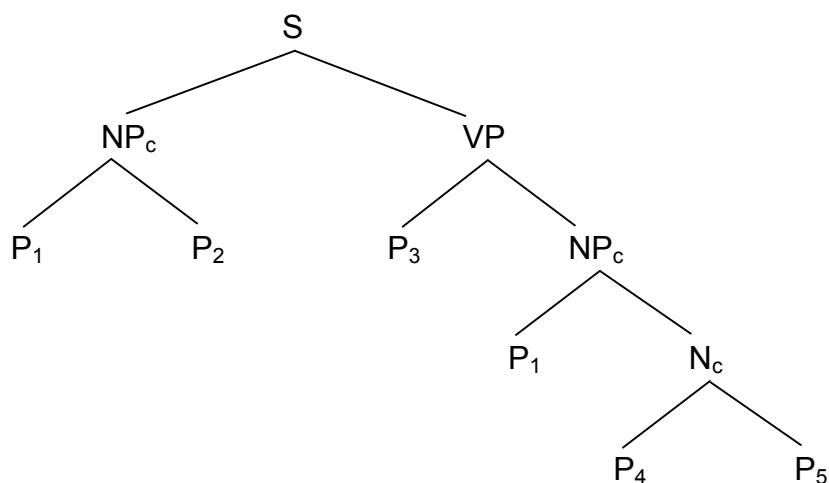
While we assume that the structure in (29) serves as the input to the projection-rule component, the projection rules proceed from the bottom to the top of this structure. In particular, by combining the sets of readings (the 'paths') of low-order constituents of the tree, the projection rules derive readings for high-order constituents. Accordingly, as in the case of the phonological component, the derivation of meaning is conceived of as proceeding cyclically. To make this more concrete, let us start at the bottom of the tree, with the set of readings for *colorful* (P_4) and *ball* (P_5).[4] Instead of spelling out these sets in terms of their dictionary entries, I will use some of Katz & Fodor's examples that convey the distinct senses of these lexical items in order to avoid introducing the formalism applied by Katz & Fodor. So, as the following examples show, P_4 contains two readings of *colorful* and P_5 implies three senses of *ball* (cf. Katz & Fodor 1963: 198):

(30) P_4

    a. *The gift came in a colorful wrapper.*

    b. *No novel is less colorful than Middlemarch, excepting Silas Marner.*

(31) P_5

    a. *Tennis is played with a ball.*

    b. *The balls whistle free over the bright blue sea.*

    c. *The queen danced at the French ambassador's ball.*

As the sentences in (30) indicate, *colorful* can refer, on the one hand, to the property of showing an abounding variety of colors, as (30a) exemplifies. On the other hand, as (30b) demonstrates, it can be used in an evaluative sense, referring to the property of having distinctive character. Turning now to the concrete noun *ball*, the sentences in (31) indicate three senses of this lexical item. First, *ball* can refer to any physical object with globular shape, as (31a) shall indicate. Second, and more specifically, it can be used to denote a physical object used as a solid missile, as (31b) points out. Finally, as demonstrated in (31c), *ball* can also refer to the social activity

---

[4] Note that we could also have started with combining the readings of *the* and *man*, since they also show up at the bottom of the tree. However, according to Katz & Fodor (1963: 198), "the order is immaterial" within this conception of semantic interpretation.

for which people meet to dance with each other. Now, given the 'paths' P_4 and P_5, each consisting of the distinct readings for the lexical items *colorful* and *ball*, the projection rules project certain readings to the next higher level in the tree by amalgamating P_4 and P_5 to form the new path P_6. This derivational step results in the tree given in (32), in which P_4 and P_5 are replaced by the new path P_6:

(32)

```
                         S
            ┌────────────┴────────────┐
          NP_c                        VP
        ┌──┴──┐                   ┌────┴────┐
       P_1    P_2                P_3       NP_c
                                        ┌───┴───┐
                                       P_1     P_6
```
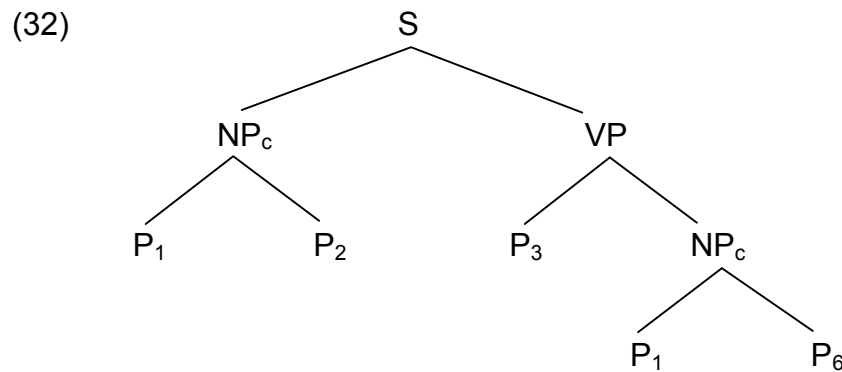
Returning to the concrete readings illustrated above, note that there are, in principle, six possible amalgamations with respect to combining P_4 and P_5. However, as Katz & Fodor (1963: 199) argue, P_6 contains only four derived readings because the combination of *colorful* used in an evaluative sense, as in (30b), with *ball* referring to the two senses of physical objects, indicated in (31a) and (31b), results in a semantic anomaly. So, while the sentence in (33) receives an unproblematic interpretation, the adjective *colorful* in the sentences given in (34) cannot be understood as referring to a 'distinctive character' in any meaningful sense:

(33)   *The queen danced at the French ambassador's colorful ball.*

(34)   a.   *The baby is playing with a colorful ball.*

      b.   *The colorful balls whistle free over the bright blue sea.*

Since a detailed exposition of the subsequent derivational details concerning the whole sentence would take me too far afield, let me summarize briefly, based on the illustration so far, how the derivation proceeds. After having amalgamated P_4 with P_5, resulting in the derived set P_6, the projection-rule component goes on to proceed from the bottom to the top of the constituent-structure tree given in (32). In doing so,

the projection rules yield a series of further amalgamations until they reach the highest constituent level 'S.' This derivational history, ending with the set of readings for the whole sentence (P_{10}), can be summarized by the following schema (cf. Katz & Fodor 1963: 205):

(35)

```
                        S: P₁₀
                  ╱              ╲
           NPc: P₇              VP: P₉
          ╱      ╲             ╱       ╲
    Det: P₁    Nc: P₂     V: P₃      NPc: P₈
                                    ╱       ╲
                               Det: P₁     Nc: P₆
                                          ╱       ╲
                                     A: P₄       Nc: P₅
```
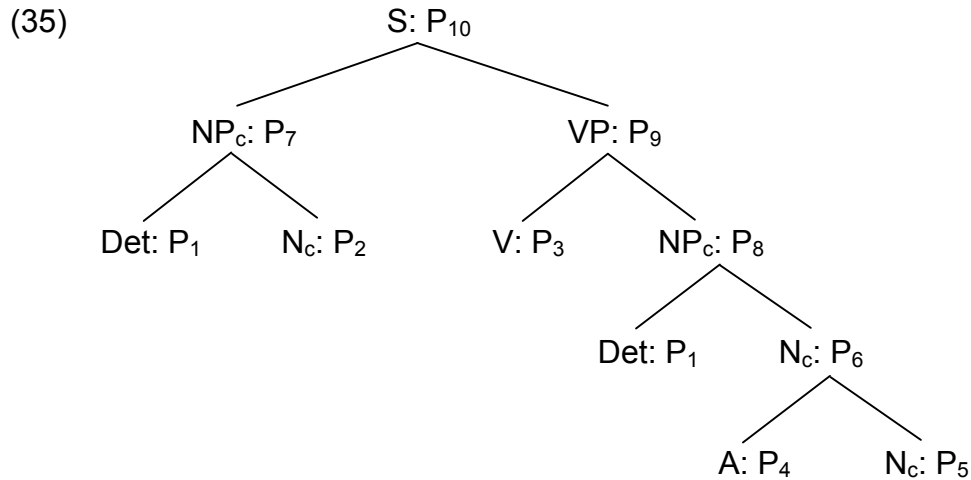
As shown in (35), every derivational step towards the final output (P_{10}) of the semantic component of the grammar consists of a complex amalgamation of sets of distinct readings, along the lines I indicated in case of the amalgamation of P_4 and P_5. That is, as in the case of amalgamating P_4 and P_5, only those readings are derived and projected to the next higher level in the tree whose subsets can combine without resulting in a semantic anomaly. Crucially, the derivation as a whole, similar to the derivation of stress, sketched in the previous section, proceeds in a cyclic, step by step manner, and the cyclic steps are determined by the structural description generated by the syntactic component. So, for instance, the semantic interpretation of our sentence (27) implies the sets P_1-P_{10}, which contain the respective readings of the constituents *the* (P_1), *man* (P_2), *hits* (P_3), *colorful* (P_4), *ball* (P_5), *colorful ball* (P_6), *the man* (P_7), *the colorful ball* (P_8), *hits the colorful ball* (P_9), and *the man hits the colorful ball* (P_{10}). Yet, since the semantic rules only operate on structures provided by the syntactic component, the semantic component does not derive any readings of substrings like *hits the*, which do not show up as a constituent in the syntactic structure of the sentence.

Let me conclude. By sketching an early concept of viewing semantic interpretation of sentences as derived compositionally from distinct senses of their subconstituents, I illustrated the theoretical background Chomsky's (1965) conception of

semantics as being 'purely interpretive' rests on. With this illustration in place, together with the brief sketch of the phonological and the syntactic component in the preceding sections, we have arrived at a general picture of the syntactocentric perspective on grammatical knowledge, as conceived of in the 1960s. From the presentation so far, it should become clear why even Jackendoff concedes this nascent perspective on grammatical knowledge to be a quite reasonable view, given the by then available approaches to phonology and semantics we hinted at in the preceding sections. Controversies revolve around the fact that this syntactocentric view has been "preserved in every subsequent version of Chomskyan theory" (Jackendoff 2002: 108) – and that, as Jackendoff argues, at the expense of taking into account the progress made both in phonology and semantics. Let us now turn to these more controversial issues by looking at the format of recent syntactocentrism.

2. The Format of Recent Syntactocentrism

As we saw in the previous chapter, the main motivation for assuming a syntactocentric conception of grammar was to describe a single system of generating rules that can both account for the ability to produce (phonetically) and to understand (semantically) an infinite range of sentences. In the context of matching this rule system to the special needs of producing and understanding expressions, I illustrated that both the phonological and the semantic component need some inventory of interpretive rules added to operations that generate the syntactic structure. Since Chomsky's (1965) seminal formulation of syntactocentrism, many approaches both to the phonological and to the semantic component have been developed. Referring to crucial insights of these various accounts, Jackendoff argues that the syntactocentric view of grammar, although it was reasonable in the 1960s, is obsolete and can only be adhered to by ignoring lessons learned from progress in both phonology and semantics. Moreover, he goes so far as to claim that the most prominent approach within recent mainstream generative grammar, "the Minimalist Program[,] offers no formal account of either phonology or semantics […], so it is no wonder that it needs only a syntactic engine" (Jackendoff 2010: 4, n. 1).

In this chapter, I will attempt to amend Jackendoff's impression by illustrating recent approaches to phonology and semantics he appears to be unaware of. These approaches, although explicitly situated within the framework of syntactocentrism, both offer detailed formal accounts and cover crucial insights from research in phonology and semantics gained since the 1960s. To do this, I will first outline the changed view of the syntactic component current approaches to phonology and semantics rest on. Having thus sketched the recent perspective on syntax, I will then turn to approaches regarding the phonological and the semantic components that are the most suitable for presenting recent syntactocentric views of these components.

2.1 The Syntactic Component

As we saw in section 1.1, one significant part of the syntactic component is the system of rewrite rules, accounting for the 'creative aspect' of language use by allowing for recursive application. Turning now to the conception of the syntactic component in recent mainstream generative linguistics, we will see that rewrite rules are dispensed with in favor of more basic combinatorial operations. To capture this change, let us briefly recall the crucial operations of the syntactic component within early generative grammar.

Returning to our sentence (1) in the last chapter, repeated here for convenience, we will see that the syntactic component must contain rewrite rules like (2) in order to generate phrase markers such as (3):

(1)   *The man comes.*

(2)   S → NP + VP

NP → Det + N

VP → V

(3)

```
                      S
                    /   \
                  NP     VP
                 /  \     |
               Det   N    V
```
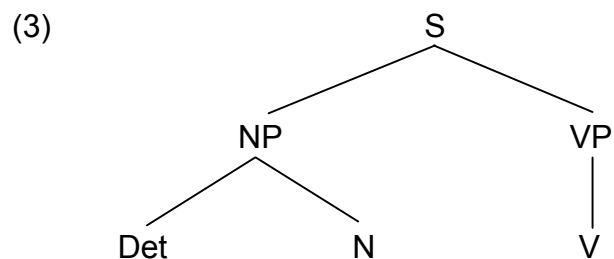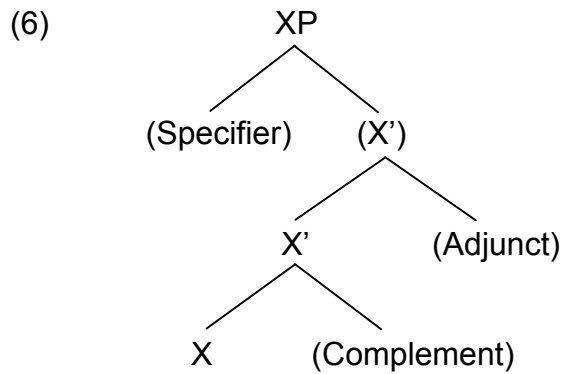
Then, lexical items are inserted into the terminal positions of the phrase marker, as the rules in (4) dictate, resulting in the tree given in (5):

(4)   Det → *the*

N → *man*

V → *comes*

(5)

```
                        S
                  ┌─────┴─────┐
                 NP           VP
              ┌───┴───┐        │
             Det      N        V
              │       │        │
             the     man     comes
```
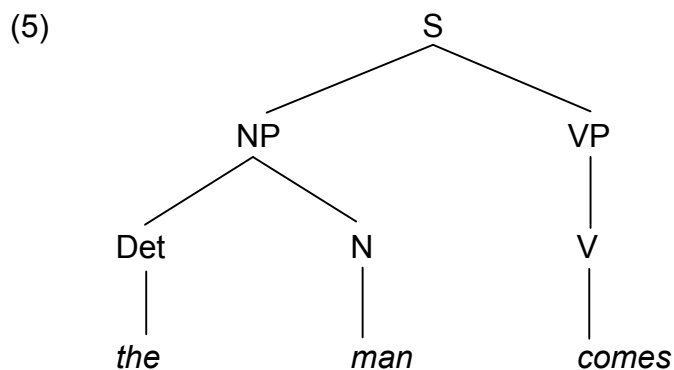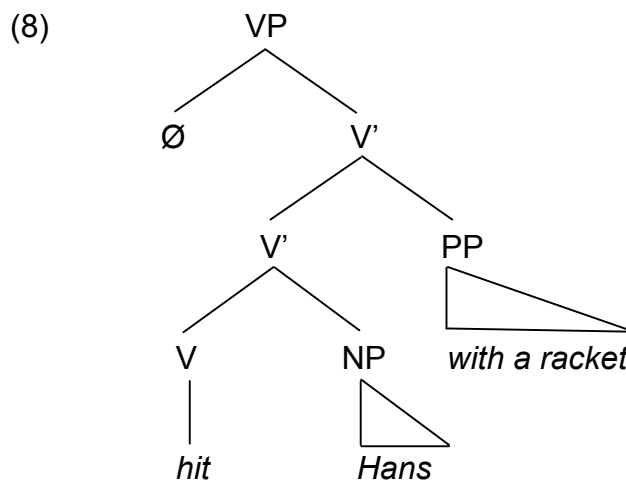
With this conception in mind, let us now turn to the problems this conception was facing in the context of the subsequent development of generative grammar.

In his discussion of the relation between verbs and their nominalizations, Chomsky (1970b) points out that theories that assume a rule component like (2) and lexical insertion such as (4) in order to generate structural descriptions like (5) imply a crucial redundancy. That is, the lexical information of elements is coded twice, both "as categories of the base […] and as features in the lexicon" (Chomsky 1970b: 208). Turning to our example, look at the structure generated by the rewrite rules in (2), which are understood as part of the 'base' component of the grammar. The VP contains only a verb, so the process of lexical insertion can, in this case, only insert intransitive verbs like *come*, *sleep*, etc. – that is, verbs that do not require another phrase as their argument. Accordingly, the rewrite rules must match the lexical entries of the verbs that will be inserted in the generated structure. Since the information about argument structure of a verb in a given derivation is technically encoded both in the lexical entries and in the rewrite rules, Chomsky (1970b: 208) points out that this "resulting 'mixed theory' had a certain technical artificiality." As a consequence, the rewrite rules of early generative linguistics were abandoned, and the structural description that serves as input to the semantic and the phonological component was henceforth viewed as the syntactic projection of the argument structure of a lexical item. With this new perspective, the general goal was to "replace the rules […] by a single schema, with a variable standing for the lexical categories N, A, V" (Chomsky 1970b: 210). This schema has come to be known as the 'X-bar schema,' whose common formulation is represented in (6):

(6)
```
                    XP
                   /  \
         (Specifier)   (X')
                      /   \
                    X'    (Adjunct)
                   /  \
                  X   (Complement)
```

According to this general format of projections, a lexical head (X) projects a maximal constituent (XP) and, depending on the information encoded in the lexicon, can take a complement. In addition, the projection can involve some elements modifying the head semantically (termed 'adjuncts') and a specifier. Let me clarify this concept with the help of the projection of the transitive verb *hit* in (7), which can be depicted as shown in (8), where some detail is omitted:

(7) *hit Hans with a racket.*

(8)
```
                VP
               /  \
              Ø    V'
                  /  \
                V'    PP
               / \    /◣
              V   NP  with a racket
              |   /◣
             hit  Hans
```

The X-bar schema implies several crucial claims about how a projection from the argument structure of a lexical item takes place. However, they are claims that I cannot substantiate here, as it would require an in-depth discussion of phrase structure (for an overview of these basic claims concerning the projection of lexical categories like *hit*, see Haegeman 1991: 78-96). Instead, let us now concentrate on critical aspects of this model, which, at the same time, will lead to a revised view of the syntactic component as a whole.

According to the X-bar schema, it is possible for a projection to contain no complement, no adjunct, and no specifier at all, as already indicated in (6) with the help of parentheses. So, if none of these positions is occupied, one question is whether the different levels of X-bar theory are motivated entirely on theory-internal grounds, as the VP of an intransitive verb like *sleep* suggests:

(9)   VP
          |
          V'
          |
   V
          |
        sleep
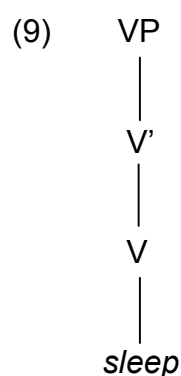
To abandon structures with such theory-internal levels, Chomsky (1995) reflects on alternative options of how a theory can conceptualize the difference between XP, X', and X. Note that in X-bar theory, the difference is accounted for by claiming that the different levels of structure should be regarded as different categories, just like nouns and verbs. Alternatively, Chomsky claims that the difference between levels like XP and X is due to their local relation with other elements. To make this point, Chomsky (1995: 242) states that "[a] category that does not project any further is a maximal projection XP, and one that is not a projection at all is a minimal projection X^{min}; any other is an X'." With this definition in mind, take a look at the following representation:

(10)   V
            /   \
           V   N
 | |
         hit *Hans*
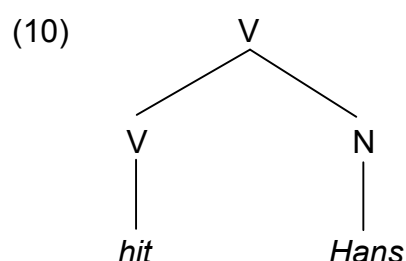
According to the quotation provided above, the elements *hit* and *Hans* are each a minimal projection ('X^{min}' or 'X') because they are not a projection at all. Both the N-projection dominating *Hans* and the topmost V-projection are maximal projections

('XPs'), since they do not project any further. The lower V-projection, which only dominates *hit*, is a projection, thus it is not an X, but, in contrast to both the N-projection and the topmost V-projection, it projects further. Under the terms of the definitions given above, it must be an X', in this case. Notice that Chomsky adopts a view of phrase structure according to which vacuous projections such as X', meaning the "intermediate nodes [, … are] left unspecified by the theory" (Muysken 1982: 61). In other words, structures like (9) are ruled out because, due to relational identification of the different levels, the representation can dispense with the X-bar level (X'), which is suspicious of being motivated merely on theory-internal grounds.

On the basis of this relational approach to phrase structure, however, one can 'minimalize' phrase structure representation even further. In his attempt to "keep to the minimalist assumption that phrase structure representation is 'bare,' excluding anything beyond lexical features and objects constructed from them" (Chomsky 1995: 245), Chomsky introduces a crucial condition of simplicity, the 'Inclusiveness Condition.' According to this condition, "any structure formed by the computation […] is constituted of elements already present in the lexical items" (Chomsky 1995: 228). Since the lexical entry for *Hans*, for instance, already provides the information that *Hans* is a noun, the representation can be reduced further by eliminating the categorial labels that refer to the terminal nodes, as shown in (11):

(11)
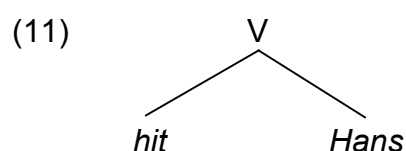
Now, consider the remaining categorial label, which marks the topmost V-projection. The raison d'être of this label is to encode that [*hit Hans*] is a verbal rather than a nominal constituent. However, the need for a mechanism to label maximal projections does not necessarily imply the use of categorial nodes like 'V,' as the following representation demonstrates:

(12)
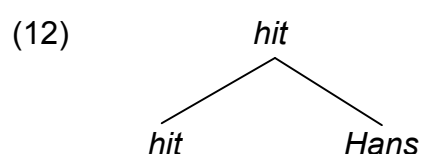
Although the 'bare' representation in (12) is "just one way to encode the 'projection' of the head" (Hornstein et al. 2005: 206), it demonstrates that the phrase structure component of the grammar can dispense with theory-internal categories because it merely needs an operation that should do both: combining lexical items and labeling the resulting product of combinatorial operations. In recent mainstream generative grammar, this operation is called 'Merge.' As already hinted at, this operation does not only consist of combining elements to form a set. Instead, since the resulting structure requires a label, "[t]he operation Merge(α, β) is asymmetric, projecting either α or β, [and] the head of the object that projects becoming the label of the complex formed" (Chomsky 1995: 246). In set-theoretic format, the asymmetric product of Merge can be depicted as follows:

(13)   {α {α, β}}

When we assume this basic operation and adopt the set notation of (13), the steps that build the structure of our example can be represented as follows (cf. Hornstein et al. 2005: 201-204 for representations of derivations using this notation):

(14)   {*hit* {*hit*, *Hans*}}

hit ←——→ Merge *Hans*

With this brief illustration of the operation Merge, we have reached the basics of the current perspective on the syntactic component within mainstream generative linguistics. Due to a "recursive step in the definition […] of admissible objects" (Chomsky 1995: 270), an aspect I do not want to delve into here, no upper bound is imposed on the number of applications of Merge. Consequently, like the syntactic component in early generative linguistics, Merge allows for recursive application and thus accounts for the 'creative aspect' of language use – but, as I attempted to point out, without postulating theory-internal categories, several rewrite rules, etc. In the light of this recent conception of the syntactic component, let us now turn to approaches concerning the phonological component that explicitly presuppose this recent view of syntactic operations within mainstream generative grammar. In this way,

I will disprove Jackendoff's (2003: 657) claim that, within mainstream generative grammar, "the relation between syntax and phonology became a no-man's-land."

2.2 The Phonological Component

To demonstrate how the syntactic operations sketched above are related to the phonological component, I will again focus on the prosodic structure that is assigned to a syntactic structure by the phonological component, as I already did in section 1.2 in order to show the conception of the phonological component within early generative linguistics. Therefore, I will abstract away from other prominent phonological aspects, like the process of linearization "that brings hierarchical structure and linear order together" (Kayne 1994: 131). As we saw in section 1.2, the early syntactocentric view of the relation between prosody and syntax implies that prosodic structure fundamentally relies on the hierarchical representation generated by the syntactic component. In recent mainstream generative linguistics, the prosodic structure assigned by the phonological component – now termed the '(P)honetic (F)orm-component' – is also considered to be determined by syntactic structure. Yet, since more proposals are on the table than in the 1960s, current approaches differ in answering the question to what extent this determination holds true (for a recent overview of different approaches to the syntax-phonology interface, see Elordieta 2007: 126-157). One recent approach that, according to Boeckx (2008: 70), can be conceived of as an exemplary instance of accounts that "argue strongly for a syntactocentric organization of the PF-side of the grammar" is Wagner's (2005) 'cyclic prosodic mapping,' which, as we will see, is essentially a particular version of the transformational cycle illustrated in section 1.2.

Wagner (2005) is primarily concerned with the prosodic property of boundary strength, which divides an utterance into different prosodic groups. Boundary strength is encoded by several suprasegmental cues, such as position, length of pause, and preboundary lengthening (for a detailed discussion of some of these cues, see Ladd 1996: 239-251), and it has been attested in various psycholinguistic studies, which demonstrate "that naive and untrained listeners can indeed give reliable and meaningful judgements […] when asked simply to indicate the strength of word boundaries in spoken utterances" (Pijper & Sanderman 1994: 2045). Wagner's

(2005) theory is grounded in these insights of both theoretical and experimental pho-
nology and especially focuses on prosodic boundaries within coordinate structures.
To gain more concreteness, look at the examples given in (15), where prosodic
boundaries are indicated by the pipe symbol '|,' and the relative rank of these
boundaries is represented by the number of pipes (cf. Wagner 2005: 40):

(15)   Who went to the forest?
    a.   *Lysander | and Demetrius | and Hermia.*
    b.   *Lysander | and Demetrius || and Hermia.*
    c.   *Lysander || and Demetrius | and Hermia.*

In contrast to the 'natural' answer given in (15a), the prosodic structures in (15b) and
(15c) suggest that some contextual aspect motivates the grouping-together of two
individuals. In particular, in (15b) the listener can assume, for instance, that Lysander
and Demetrius went to the forest together, and Hermia went on her own because
and Hermia is prosodically set apart from the rest of the utterance. By contrast, the
prosodic phrasing in (15c) can introduce the presupposition that Demetrius and
Hermia went together and Lysander went on his own (for the precise relation be-
tween presupposition and prosody, see Wagner 2005: 245-281).

Wagner (2005) is particularly interested in differences concerning 'flat' prosody
and hierarchically-organized prosodic structure. That means that he focuses on the
circumstances under which coordinate structures show each conjunct to be prosodi-
cally on a par with every other conjunct, like in (15a), and he concentrates on the
circumstances under which the conjuncts are hierarchically organized, like in (15b)
and (15c). So, in an attempt to force this difference, he alternates the functors *and*
and *or*, as shown in the following examples (cf. Wagner 2005: 43):

(16)   a.? *Lysander | and Demetrius | or Hermia.*
    b.   *Lysander || or Demetrius | and Hermia.*
    c.   *Lysander | or Demetrius || and Hermia.*

Due to the alternation of *and* and *or*, the flat prosody in (16a) seems inappropriate,
while the articulated (i.e. hierarchically organized) prosody in (16b) and (16c) is fine.
As a next step, Wagner (2005) formalizes such clear cases of articulated prosodic

structure, which contain alternating functors, by replacing concrete examples, such as (16b) and (16c), with the following more abstract structures, where, using a simplified notation of Wagner (2010: 186), (17a) corresponds to (16b) and (17b) to (16c):

(17)   a.   A || or B | and C   Interpreted as: A or (B and C)

       b.   A | or B || and C   Interpreted as: (A or B) and C

As a next step, Wagner (2005) extends this representation of prosodic properties by adding a metrical grid to account for both the prosodic grouping illustrated so far and aspects of prosodic prominence. In particular, he uses a version of the 'bracketed metrical grid,' as proposed by Halle & Vergnaud (1987). To capture Wagner's argument about the relation between syntax and prosody, a little digression to this notation is in order.

   Look at the following representations, which apply the notational variant of the 'bracketed metrical grid' to the cases presented above (cf. Wagner 2005: 67):

(18)   a.   A or (B and C)   b.   (A or B) and C
            |✕| ✕   ✕| |✕   ✕| ✕|
            |✕| ✕| ✕| |✕| ✕| ✕|
             A   B   C   A   B   C

The representations in (18) are to be read as follows: Like in (17), the 'pipes' indicate boundaries, that is, they represent aspects of prosodic grouping. In contrast to the representation given in (17), the notation in (18) encodes boundary strength by using the height of the column of pipes. In other words, indicated foot boundaries at higher grid lines are considered to be stronger than boundaries marked at lower lines. Crucially now, the representation in (18) allows for the encoding of pitch accents. In coordinate structures such as (18a) and (18b), each conjunct receives an accent, so each conjunct projects top-line grid marks. Consequently, the accent of each element is on a par with the other elements. Note that, due to the use of this notation, the fact that one element receives the 'nuclear stress' of the structure does not have to be represented as a further projection to a higher line. When we assume, with respect to English, that "the last heavy stress in an intonational unit takes the nuclear

34

heavy stress" (Newman 1946: 176), main prominence can be derived from representations such as (18) because all elements are on a par concerning heavy stress. Therefore, nuclear stress is placed on the final element of the intonational unit. So, at first sight, a hierarchical notation seems to be superfluous, inasmuch as a more 'flat' notation, as shown in (19), would work as well to derive nuclear stress and encode prosodic grouping:

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

        || × || × | × || || × | × || × ||

          A   B   C   A   B   C

However, in other cases, some conjuncts are deaccented, that is, they do not receive the same accent, the same 'heavy stress,' as the other conjuncts. Consider the following example (cf. Wagner 2005: 71):

(20)   a.   *Demetrius and Hermia?*

     b.   *No, Lysander and Hermia.*

In the answer (20b), the material following the first conjunct *Lysander* is viewed as deaccented. In such cases, the final element of the structure does not receive nuclear stress and, accordingly, does not project to the top line, as shown in (21):

(21) | × |

       | × | × |

        Lysander   Hermia

To summarize, the hierarchical notation of the metrical grid introduced above can be regarded as a necessary instrument to account for a broader range of prosodic structures, including cases like (21). After having dealt with the notation Wagner (2005) uses to represent the prosodic structure of coordinate structures,[5] let us now

---

[5] I should mention that this representation deviates from more common representations used in phonological theory by not labeling the lines in the metrical grid with respect to the categorial status of being, for instance, either an intonational or a phonological phrase – distinctions that are of high relevance in prosodic phonology (for a seminal schematic overview of distinct categories in prosodic phonology, see Selkirk 1986: 384).

turn to the question of how this prosodic structure can be derived from syntax.

Recall that, in early syntactocentric concepts, the prosody of an expression is determined compositionally by the transformational cycle, as illustrated in section 1.2. Turning now to the derivation of prosodic properties of coordinate structures, we will see that Wagner (2005) assumes a particular version of the transformational cycle in order to account for mapping syntactic structure to prosody. Let us illustrate this approach, starting with the 'flat' coordinate structure given in (22):

(22)   A | and B | and C

Concerning this structure, Wagner (2005: 74-75) postulates four derivational steps, as shown in the following representations, in which the syntactic derivation is represented schematically on the left side and in which the corresponding metrical grid is notated on the right side. At this point of visualization, Wagner abstracts away from functors such as *and*, treating the structure given in (22) as if it consists only of the conjoined elements.

(23)   a.   Start Point
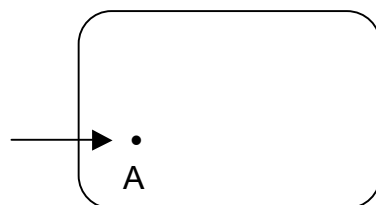
| ✕ |
A

b.   First Merge
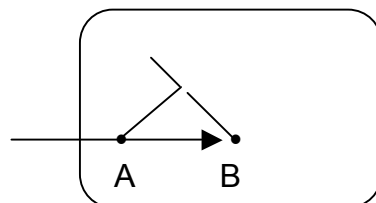
| ✕ | ✕ |
A   B

c.   Second Merge
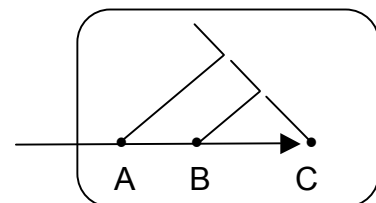
| ✕ | ✕ | ✕ |
A   B   C

As mentioned in section 2.1, elements are combined by the operation Merge in recent versions of mainstream generative linguistics. This operation builds the syntactic tree step by step, as (23a)-(23c) illustrate. When we look at the phonological status of the elements combined by Merge, each conjunct is regarded as already spelled out, that is, as already mapped to the phonological component. This mapping is indicated by black dots standing for a 'Spell-Out domain.' Consequently, the merging of elements on the left side of the representation corresponds to the addition of these elements to the prosodic structure, given on the right side of the schema. In addition to these small domains of Spell-Out, at some points in the syntactic derivation, la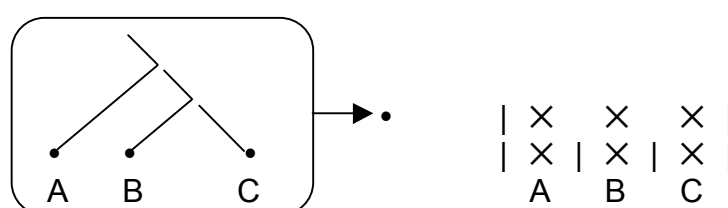rger domains are completed and mapped to the phonological component as a whole, as the following representation depicts:

(24)   Spell-Out

Wagner (2005) represents this mapping of the whole structure by adding a new top-line to the metrical grid and by projecting all grid marks on the old top-line up to the new top grid line, as shown on the right side of step (24). The result is the prosodic grouping we already indicated in (22).

Let us now look at the process of deriving a metrical grid for an expression that shows a more articulated – meaning hierarchical – prosodic structure, just like (17a), repeated here for convenience as (25):

(25)   A || or B | and C   Interpreted as: A or (B and C)

According to Wagner (2005: 75-76), this structure is the result of the spelling-out of two separate cycles. First, the constituents *B* and *C* are combined in a first cycle, forming the subconstituent '*B and C*.' Since this combination, according to Wagner's approach, constitutes a cycle, the product of Merging *B* and *C* is spelled out and

mapped on a phonological representation, thus projecting a new top grid line, as shown in (26):

(26)　　　　　　Spell-Out of the First Cycle


```
| ✕   ✕ |
| ✕ | ✕ |
```
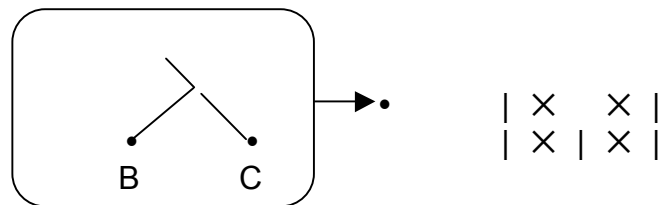
Then, the output of this first cycle is inserted into a second cycle, as shown in (27):

(27)　　　　　　Second Cycle
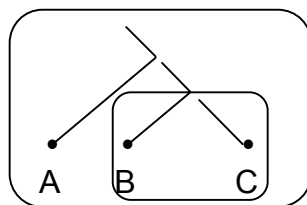
As a first step of spelling out this second cycle, the metrical elements are concatenated at the top grid line, as shown in (28):

(28)　| ✕ | ✕　✕ |
　　　　　| ✕ | ✕ |
　　　　A　B　C

After the concatenation, a new grid line is projected and the grid column belonging to the lowest grid line is filled for notational reasons, as shown in (29):

(29)　| ✕　✕　✕ |
　　　　| ✕ | ✕　✕ |
　　　　| ✕ | ✕ | ✕ |
　　　　　A　B　C
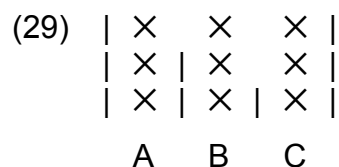
In sum, the Spell-Out of the second cycle can be represented as follows:

(30)   Spell-Out of the Second Cycle
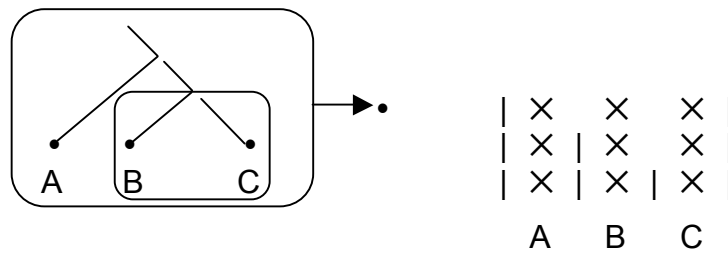
By contrast, consider the derivational steps that yield (17b), repeated here as (31):

(31)   A | or B || and C   Interpreted as: (A or B) and C

According to Wagner (2005: 77), the prosodic structure indicated in (31) is due to a left-branching syntactic structure, which is again associated with spelling out two cycles. In a first cycle, the constituents *A* and *B* are combined, forming the sub-constituent '*A or B*.' This constituent is spelled out, that means, it is mapped to phonology, and thus, a new top grid line is projected, as (32) shows:

(32)   Spell-Out of the First Cycle
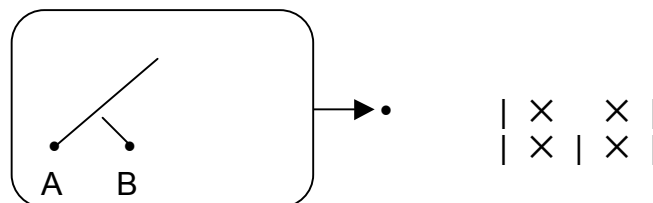
Then, the output of this first cycle is inserted into a second cycle, as shown in (33):

(33)   Second Cycle
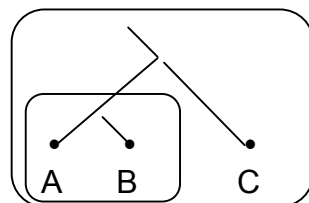
Like in the derivation of the prosodic structure based on a right-branching structure, the next step for deriving a metrical grid is concatenating the metrical elements of the first and the second cycle, as (34) illustrates:

(34) | ✕   ✕ | ✕ |
      | ✕ | ✕ |
        A   B   C

Finally, projecting a new grid line and filling up the rightmost grid column results in the phonological representation (35), which corresponds to the structure indicated in (31):

(35) | ✕   ✕   ✕ |
      | ✕   ✕ | ✕ |
      | ✕ | ✕ | ✕ |
        A   B   C
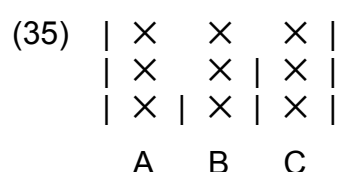
So, in sum, the Spell-Out of the second cycle of this left-branching structure can be depicted as follows:

(36)   Spell-Out of the Second Cycle

| ✕   ✕   ✕ |
| ✕   ✕ | ✕ |
| ✕ | ✕ | ✕ |
  A   B   C
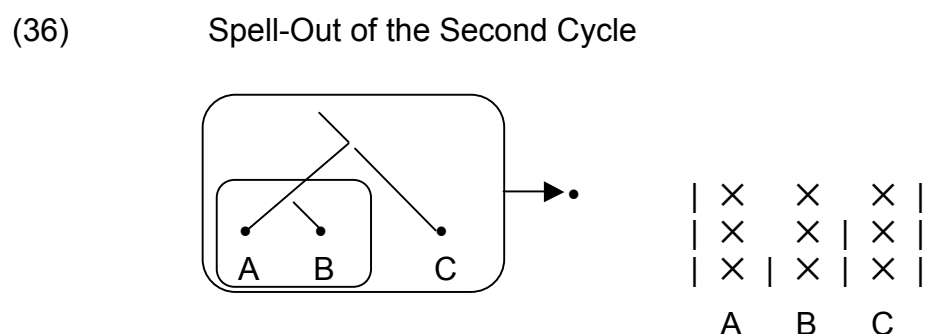
Let us take stock. As Wagner (2005) argues in light of data from the prosody of co-ordinate structures, the prosody of an expression is determined compositionally, that is, by the corresponding syntactic structure and by the way of how the structure is assembled. Hence, in general accordance with the early conception of the phonological component we sketched in section 1.2, Wagner's (2005) approach assumes a "one-to-one mapping between syntactic derivations and grid representations" (Wagner 2005: 80). His conception, unlike the approach sketched in section 1.2., does not require the postulation of specific categorial entities such as N or NP,

which were questioned in recent conceptions of the syntactic component, as we saw in section 2.1. Rather, as Boeckx (2008: 67) points out, Wagner's account can serve as an illustration that "prosodic structure building requires nothing more than basic operations like Concatenate, Project, and Embed – all of which are part of the standard definition of Merge."

To conclude, Wagner's approach both maintains the syntactocentric view of the relation between syntax and phonology and, whether one believes his syntactocentric perspective or not, presents a detailed formal account of the phonological component. Moreover, this account draws on many insights of the broad field of phonology, as can be witnessed by his extensive discussion of both theoretical and empirical evidences, which was spared out in this section for the sake of illustration. Accordingly, the presentation so far supports the general claim that "Jackendoff's assertion that minimalism lacks serious study of the interfaces is unacceptable in light of detailed works on the syntax-phonology interface" (Boeckx & Piattelli-Palmarini 2007: 409). Having thus amended Jackendoff's argument that the relation between syntax and phonology is a 'no-man's-land' within mainstream generative grammar, let us now turn to recent conceptions of the semantic component.

2.3 The Semantic Component

As I already mentioned with respect to the phonological component, there are various approaches to the semantic component within recent mainstream generative linguistics. So, while there is the prominent view that formal devices like "[f]unctional application, predicate modification and lambda abstraction are probably the minimal inventory that is needed for the interpretation of all hypothesized syntactic structures" (Sauerland & von Stechow 2001: 15415), several approaches assuming a clearly higher number of combinatorial modes exist (cf., e.g., Chung & Ladusaw 2004). Yet, also recent accounts assume semantic interpretation to be rather simple and uniform, thus, they represent an "illustration of Minimalist thinking in semantics" (Pietroski 2008a: 19). Since this minimalist perspective on semantics dovetails in many respects with the recent conception of the syntactic component sketched in section 2.1, in the following, I will limit myself to a brief sketch of this particular perspective on semantics.

Remember from section 2.1 that the only combinatorial operation of the syntactic component is the operation Merge. According to Pietroski (2008b), this operation can be decomposed into two suboperations, 'Concatenate' and 'Label.' In order to avoid terminological confusion, Pietroski substitutes 'Merge,' which is originally conceived of as a primitive operation, with 'Combine,' understood as a non-primitive operation, as the following representation clarifies (cf. Pietroski 2008b: 326):

(37)   COMBINE(α, β) = LABEL[CONCATENATE(α, β)]
       CONCATENATE(α, β) = α^β
       LABEL[α^β] = [α β]$_{\alpha/\beta}$

This decomposition merely spells out underlying assumptions that I have already touched on in section 2.1, when I illustrated the syntactic operation Merge. In particular, merging two elements, for example *hit* and *Hans*, does not only involve concatenating these items, or, to put it more technically, it does not merely result in an ordered pair like hit, Hans. Rather, this operation also implies the labeling of the concatenated structure, as already shown in (14), repeated here for convenience as (38):

(38)   {*hit* {*hit*, *Hans*}}

       hit ←——→ $_{Merge}$ *Hans*

Let us turn to semantics now: Hornstein & Pietroski (2009), committed to the syntactocentric perspective, hypothesize that the meaning of a product of Combine is compositional, that is, the meaning crucially depends on its subconstituents and how they are combined. Importantly, in contrast to most accounts within Formal Semantics, Hornstein & Pietroski do not assume several modes of combination but claim that the semantic contribution of combining elements can be accounted for by merely referring to the two suboperations of Combine. In more technical terms, in their view, combining two expressions *A* and *B* results in the semantic instruction '**SEM**COMBINE(A, B),' which can be decomposed into suboperations, as the following notation makes clear (cf. Hornstein & Pietroski 2009: 116):

(39) **SEM**COMBINE(A, B) = **SEM**LABEL[CONCATENATE(A, B)]

So, by assuming that the semantic instruction to interpret expressions looks roughly like (39), Hornstein & Pietroski (2009) correlate the basic syntactic operations with basic semantic operations in an isomorphic way. Turning to the suboperations in detail, they propose "that concatenation is an instruction to *conjoin monadic concepts*, and that labeling provides a vehicle for *invoking thematic relations* when (and only when) the labels of concatenates *conflict*" (Hornstein and Pietroski 2009: 116, emphasis in the original). Let us exemplify this abstract claim.

Examine first an elementary case of adjunction like the phrase *red ball* – a paradigm case of configurations in which the labels of concatenates do not conflict (for what follows, see Hornstein & Pietroski 2009: 126-127). Hornstein & Pietroski hypothesize that lexical items like *red* are, in the context of semantic interpretation, instructions to 'fetch' monadic concepts like RED (for an elaboration of this explicitly cognitive view on meaning, see Pietroski 2007: 343-347). Adjuncts like *red* are regarded as unlabeled lexical 'roots,' when they combine with a labeled (in this case: nominal) element like *ball*. More specifically, the phrase *red ball* is the result of concatenating the root '√*red*' with the labeled head '[$_N$ √*ball*^N]' (which is itself the result of combining the functional formative 'N' with the unlabeled lexical root '√*ball*'). According to this conception, the combinatorial suboperations concerning *red ball* can be represented as follows:

(40) [$_N$ √*red*^*ball*$_N$] =
    LABEL[CONCATENATE (√*red*, *ball*$_N$)] =
    LABEL[CONCATENATE (√*red*, LABEL[CONCATENATE √*ball*, N)]]

So, in cases of adjunction like *red ball*, Hornstein & Pietroski suppose that there is only one option to label the whole phrase, since, according to their conception, 'N' is the only label in the compositional structure, and thus, only 'N' can project to label the whole phrase. Since there are no conflicting labels in such cases, the semantic interpretation of the syntactic structure [$_N$ √*red*^*ball*$_N$], according to this hypothesis,

merely implies the instruction to conjoin the concepts that were fetched via *red* (RED), *ball* (BALL), and 'N' (INDEXABLE),[6] as (41) illustrates:

(41) RED( ) & BALL( ) & INDEXABLE( )

To summarize, examples like *red ball* are viewed as simple cases of combining predicates; that is, roughly speaking, *red ball* merely corresponds to a concept of things that are red, and a ball, and indexable. With this paradigm case of adjunction in mind, let us now look at cases that contain conflicting labels.

In contrast to adjunction, combining a predicate with an argument, as Hornstein & Pietroski argue, calls for a semantic factor that is absent in cases like *red ball*. Examine the case of the syntactic structure '[$_V$ *stab*$_V$^*Caesar*$_N$]' (for the following aspects, see Hornstein & Pietroski 2009: 127-128). In this case, the concatenates *stab* and *Caesar* have competing labels and thus, according to the definition of Hornstein & Pietroski given above, combining them is not just an instruction to conjoin concepts that are fetched via *stab*$_V$ and *Caesar*$_N$. Instead, conjoining these two constituents, under this hypothesis, must be supplemented with an additional operation. Note that the labeling of the whole phrase as a complex 'V' is regarded as an introduction of grammatical relations. In particular, and in more cognitive terms, the instruction '*Caesar*$_N$-as-V' is considered to be an instruction to fetch a concept of things that have Caesar as their internal participant. As a consequence, Hornstein & Pietroski assume that being an internal argument of a predicate, like '*Caesar*$_N$-as-V' is in our case, can be interpreted semantically as a two-step instruction. First, the argument expression *Caesar* itself is an instruction to fetch a monadic concept CAESAR(X). Second, the marking of *Caesar* as an internal participant of another concept is an instruction to form the corresponding complex concept (42), which amounts to something like 'there was an 'x,' it was Caesar, and it was the internal participant of something:

(42) $\exists$X [CAESAR(X) & INTERNAL (_, X)]

---

[6] Here, Hornstein & Pietroski (2009) draw on accounts that emphasize "noun's being the only lexical category that bears a referential index" (Baker 2003: 104).

Now, this something (a 'thing with participants'), whose internal participant is expressed by *Caesar*, can be spelled out by using event representations like (43), which amounts to something like 'there was an event (E) of stabbing, and the event involved Caesar as an internal participant.'[7]

(43)   ∃E [STAB(E) & INTERNAL(E, CAESAR)]

Hornstein & Pietroski go on to specify the semantic interpretation further by referring to common terms of semantic roles. So, since internal participants of *stab* are 'Patients,' rather than, for instance, unaffected 'Themes,' the concept INTERNAL(E, X) can be replaced by specific thematic content, as shown in (44):

(44)   ∃E [STAB(E) & PATIENT(E, CAESAR)]

To summarize their discussion, the argument *Caesar* is interpreted, according to their approach, as a predicate of an event – the event of 'stabbing' – and thus, its interpretation demonstrates just another case of predicate conjunction, not, in principle, different to the conjunctive operation involved in cases like *red ball*. However, in contrast to cases of adjunction like *red ball*, complex predicate-argument constructions can only be analyzed in terms of predicate conjunction because syntactic "[l]abels […] provide vehicles for introducing dyadic concepts like INTERNAL(E, X)" (Hornstein & Pietroski 2009: 130). In other words, and with respect to our case *stab Caesar*, the label 'V,' marking the whole phrase '[$_V$ stab$_V$^Caesar$_N$]' as verbal, yields the instruction '*Caesar*$_N$-as-V,' which introduces a dyadic relation, as illustrated above. At this point, it may be worth to mention the fact that this invoking of dyadic concepts does not necessarily rely on categorial labels such as V or N. As we mentioned in section 2.1, any other mechanism, even postulating no syntactic categories at all, would work here, as long as it labels the product of Merge or, in terms of Hornstein & Pietroski's approach, Combine.

All in all, by reducing semantic interpretation to the operation Concatenate – an instruction to conjoin monadic concepts – and to the operation Label – a vehicle for

---

[7] This conception of Hornstein & Pietroski (2009) crucially relies on the assumption that "[t]he words 'is an event that consists in the fact that' are to be viewed as an operator which […] forms a predicate of events" (Davidson 1967: 90).

invoking thematic concepts – Hornstein & Pietroski (2009) claim that basic semantic operations can be derived from the basic operations that are associated with Merge in an isomorphic way. Therefore, their approach fits well into the syntactocentric view of language within mainstream generative linguistics. So, in contrast to Jackendoff's claim that there is no formal account of semantics within minimalism, this section illustrated that approaches exist within recent mainstream generative grammar that do not only provide a detailed formal account of the semantic component but also adopt a cognitive view on meaning by referring to processes like 'fetching' concepts. Thus, these accounts "do not suffer from the shortcomings, justly stressed by Jackendoff, of works in semantics within the purely logical tradition" (Boeckx & Piattelli-Palmarini 2007: 409). So, after having amended Jackendoff's impression that both phonology and semantics have become a 'no-man's-land' in mainstream generative grammar, let us now turn to an even more striking omission in Jackendoff's discussions of syntactocentrism: the recent shift from 'representational' to 'derivational' syntactocentrism.

3. The Shift from Representational to Derivational Syntacto-centrism

In the preceding chapter, I indicated that recent versions of syntactocentrism, contrary to Jackendoff's assertions, offer detailed formal accounts of both the phonological and the semantic component, taking into account crucial insights from research in these areas carried out since the 1960s. So, while Jackendoff's (2002: 108) claim that the syntactocentric view "is preserved in every subsequent version of Chomskyan theory" may be right, his observation that serious study of the interfaces in mainstream generative grammar is a 'no-mans-land' misses important developments within mainstream generative grammar. In this chapter, I will argue, in a similar vein, that Jackendoff misses another crucial point in his discussions of syntactocentrism, namely the recent shift from 'representational' to 'derivational' syntactocentrism. Specifically, even in his recent publications (cf. Jackendoff 2007: 37), he merely cites a grammar model of the early 1990s that served, at that time, as the starting point for a more derivational thinking. In this chapter, since recent derivational syntactocentrism defines itself via modifying particular aspects of preceding models, I will first illustrate the development of a representational view of syntax. In doing so, to ensure commensurability, I will orient my presentation to Jackendoff's short overview of this development (cf. Jackendoff 2002: 107-111). Then, in section 3.2, I will concentrate on the view of derivational syntactocentrism, which, as I will show, qualifies the claim that "[t]he Minimalist Program […] assumes that the structures and derivations of Principles and Parameters Theory are essentially correct" (Culicover & Jackendoff 2005: 88). However, before doing so, let us first consider the view of mainstream generative grammar that still underlies most of Jackendoff's assertions: the concept of representional syntactocentrism.

3.1 Representational Syntactocentrism

At the beginning of generative linguistics, Chomsky points out that language is a highly complex system. Famous examples like the following illustrate that an analysis of a sentence into a linear string of words or morphemes cannot account for all the properties of an utterance (cf. Chomsky 1975 [1955]: 215):

(1)   *Flying planes can be dangerous.*

While, in this case, there is one linear string of words, this string can receive two different interpretations, paraphrased in (2) and (3):

(2)   'Flying planes (as objects) can be dangerous.'

(3)   'Flying planes (as an activity) can be dangerous.'

So, this ambiguous case illustrates that an analysis in terms of linear order cannot capture the entire complexity of a sentence. Consequently, as Chomsky (1975 [1955]: 63, emphasis in the original) points out,

> [l]inguistic theory attempts to reduce this immense complexity to manageable proportions by the construction of a system of *linguistic levels*, each of which makes a certain descriptive apparatus available for the characterization of linguistic structure.

In addition to different levels identified outside of syntax, including phonetics, phonemics, morphemics, and morphophonemics (for a short overview, see Lasnik 2005: 63-66), Chomsky used examples like (1) to motivate syntax-internal "levels of representation" (Chomsky 1975 [1955]: 6). In early conceptions of generative grammar, it was already argued for the distinction of at least two syntax-internal levels of structure by assuming "[a] grammar containing a level **P** and a level **T**" (Chomsky 1975 [1955]: 380, emphasis in the original), that is, a level providing the structural information about a sentence in terms of phrase structure (P) and a level giving information concerning transformational steps that affect this structure (T). However, it was not before Chomsky (1965) that the famous distinction between 'Deep Structure' and 'Surface Structure' was drawn, which, as Chomsky (1966: 31-51) argues, is, like the 'creative aspect' of language, again rooted in philosophical concepts of the premodern period.[8] According to this distinction, Deep Structure is the level of representation interfacing with semantics and underlying a sentence prior to any transformations. So, for instance, while there are two different Deep Structures and one Surface

---

[8]   For a recent reflection that traces the historical roots of this distinction by referring to philosophical antecedents not touched on by Chomsky (1966), see Raible (2009).

Structure in cases like (1), there are, under this hypothesis, only one Deep Structure and two Surface Structures in cases like active and passive sentences. In other words, the active and the corresponding passive roughly mean the same but sound different due to transformational operations, and thus, they have different Surface Structures. The sentences given in (4) illustrate this long-held assumption that "an active sentence and the corresponding passive are synonymous" (Chomsky 1957: 94):

(4)   a.   *Hans arrested Peter.*

       b.   *Peter was arrested by Hans.*

So, putting these two levels of representation together with the rewrite rules, illustrated in section 1.1, and with the lexicon – together: the 'base component' – and adding certain transformational operations, we arrive at the overall picture that

> the syntactic component consists of a base that generates deep structures and a transformational part that maps them into surface structures. The deep structure of a sentence is submitted to the semantic component for semantic interpretation, and its surface structure enters the phonological component and undergoes phonetic interpretation. (Chomsky 1965: 135)

Graphically, this model can be represented as follows (cf. Jackendoff 2002: 109):

(5)   Phrase structure rules   Lexicon

                       ↘   ↙

               Deep Structure   ⟶   Semantic component

                       ↓

               Transformations

                       ↓

               Surface Structure   ⟶   Phonological component
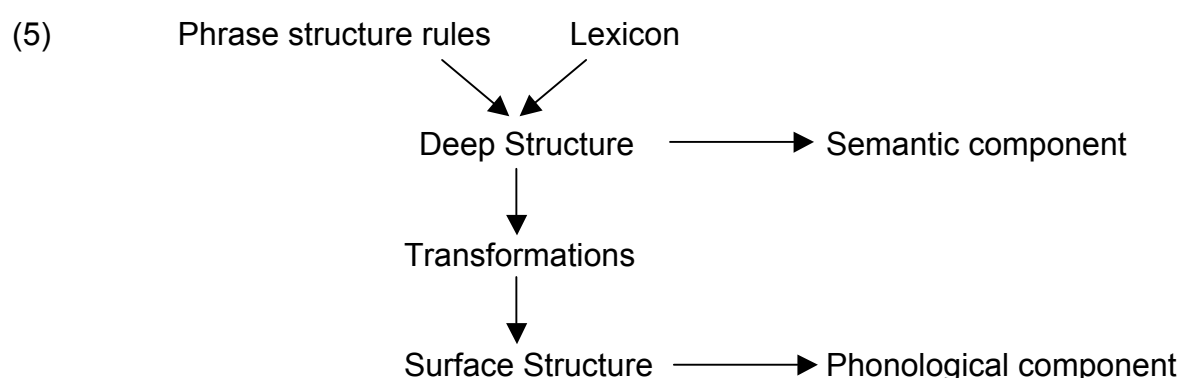
Crucially, since Deep Structure is the level before any transformations apply, this model implies the claim that transformations preserve meaning and, accordingly, that transformations do not affect the meaning of sentences. However, look at the following example, which shows that the active sentence (6a) is not synonymous with its

passive given in (6b) and which thus demonstrates "that not even the weakest se-
mantic relation (factual equivalence) holds in general between active and passive"
(Chomsky 1957: 101):

(6) a. *Everyone in the room knows at least two languages.*

     b. *At least two languages are known by everyone in the room.*

According to Chomsky (1957: 100-101), (6b) strongly suggests that, in contrast to
(6a), it is the very same two languages that everyone knows. Accordingly, examples
like (6) are regarded as cases in which transformational operations seem to affect
the meaning of utterances. Specifically, passive transformations involving quantifiers
like *a, some,* and *every* have truth conditional consequences. That is, (6a) may be
true, while (6b) is false, if one person knows, for instance, only German and English
and another only Chinese and Spanish.

   In light of such cases, Chomsky (1965) expresses first doubts regarding the
model given in (5) and rejects rather strong claims that even in cases like (6) "both
actives and passives […] are full paraphrases of each other" (Katz & Postal 1964:
72). So, in light of examples such as (6), he proposes that there are two 'senses' of
semantic interpretation. In particular, regarding cases like (7), he argues for both
"identity of the deep structures […] and […] an overriding consideration involving or-
der of quantifiers in surface structures" (Chomsky 1965: 224, n. 9). These considera-
tions ultimately led to an architecture according to which both Deep Structure and
Surface Structure contribute to semantic interpretation, as (7) shows graphically (cf.
Jackendoff 2002: 109):

(7) Phrase structure rules   Lexicon

                    Deep Structure

          Transformations   Semantic component

          Surface Structure   Phonological component
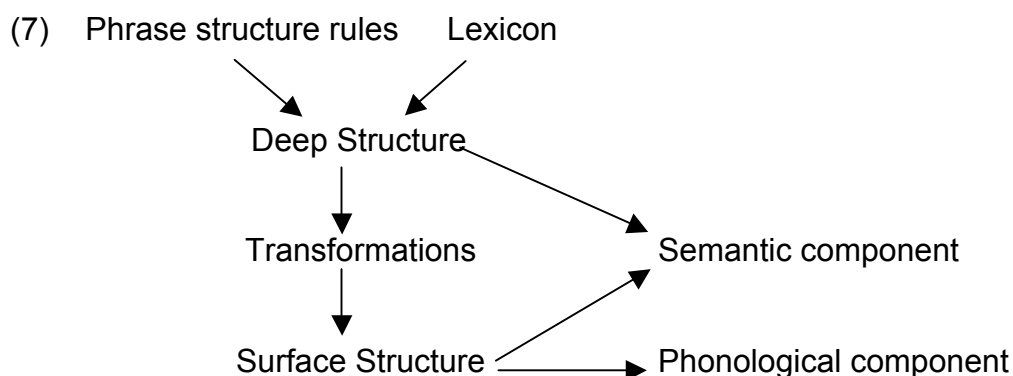
According to this model, the level of Deep Structure determines semantic interpretation with respect to grammatical properties like 'subject of,' and the level of Surface Structure determines discourse properties like scope relations that involve quantifiers, as exemplified by (6). Hence, this model articulates explicitly "that properties of surface structure play a distinctive role in semantic interpretation" (Chomsky 1970a: 89).

Since the role of Deep Structure is thus constrained to determining grammatical relations and since the role of Surface Structure, in the light of cases like (6), is viewed as indispensable for semantic interpretation, Chomsky (1973), in an attempt to arrive at a more uniform theory, argues for an account that provides an option to "interpret sentences directly in terms of the surface structure position" (Chomsky 1973: 283). This is accomplished by postulating that syntactic elements leave 'traces' after having been displaced in the structure. To make this concrete, examine the following example:

(8)   *Hans seems to drink beer.*

In constructions like (8), the verbs *seem* and *drink* differ regarding the grammatical relations they imply. On the one hand, *seem* takes the proposition 'Hans drinks beer' as a complement, but it does not involve a subject that is associated with a semantic role, that is, its subject position is 'semantically empty.' This can be illustrated by (9), where the subject position is occupied by the pronominal expletive *it*:

(9)   *It seems that Hans drinks beer.*

On the other hand, the verb *drink* assigns two semantic roles – 'theta-roles,' in Chomsky's (1981) terms – that is, there must be a 'drinker' (*Hans*) and there must be a 'drink' (*beer*). Given these specific grammatical relations, the Deep Structure of (8), conceived of as determining these relations, must transparently represent these semantic role configurations. This can be schematized as in (10), where 'Δ' represents the semantically empty position that, as we saw above, can be optionally filled with pronominal heads like *it*:

(10)   [∆ *seems* [*Hans*drinker *to drink beer*drink]]

Now, having explored cases like (10), Chomsky (1973: 266) argues that, when *Hans* moves to the subject position of the matrix clause, "it leaves behind a 'trace' which it controls." This 'control relation,' established by a specific "transformation that […] takes the subject of the complement sentence […] and substitutes it for the pronominal head" (Rosenbaum 1967: 64), is indicated by coindexing the trace ('t') with *Hans*. Accordingly, the resulting Surface Structure is the following:

(11)   [*Hans*$_i$ *seems* [t$_i$ *to drink beer*]]

Given this technical innovation of traces, the movement operations involved in cases like (8) do not destruct the configurational properties that are needed for determining grammatical relations because the coindexed trace indicates the relevant control relation. In other words, this interpretive relation is still accessible at the level of Surface Structure because the structure given in (11) still represents that *Hans* is the subject of the embedded sentence.

Recall now that there is an independent motivation to suppose that Surface Structure is involved in determining semantic interpretation, as the discussion of scope properties of sentences above has shown. So, given this motivation and assuming that movement leaves behind traces, it was soon argued that "it seems reasonable to postulate that *only* surface structures undergo semantic interpretation" (Chomsky 1975: 96, emphasis in the original). This resulted in an architecture that is depicted in the following schema (cf. Jackendoff 2002: 109), in which the phrase structure rules are replaced with 'X-bar theory,' which was by then established in generative linguistics, as we sketched in section 2.1:

(12)
```
         X'-Theory    Lexicon
               ↘    ↙
           Deep Structure
                 │
                 ↓
          Transformations
                 │
                 ↓
          Surface Structure
              ↙      ↘
Phonological component    Semantic component
```
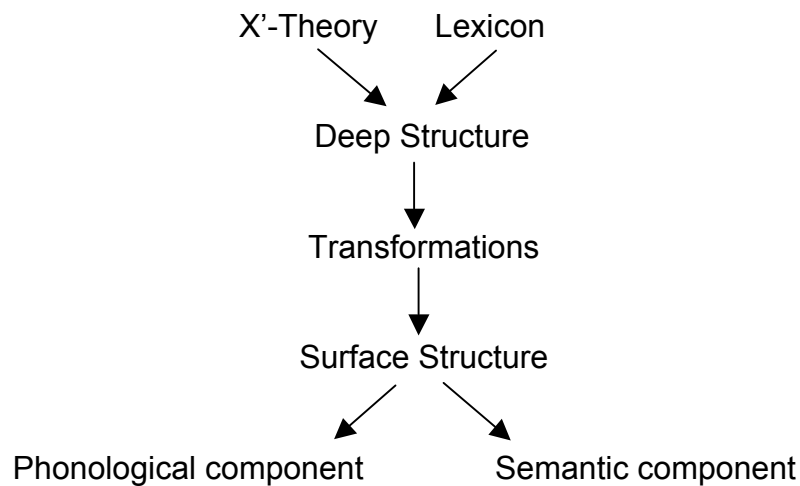
However, it was soon argued that Surface Structure is not an adequate level to account for scope relations. Look at the following sentence (cf. May 1977: 13):

(13)   *Every man loves some woman.*

While this sentence has only one Surface Structure, it has two different scope readings. That is, it is ambiguous between a reading according to which all men love one and the same woman – reading (i) – and a reading according to which it is asserted that each man loves a different woman – reading (ii). In order to analyze this sentence, May (1977: 13) argues for "two distinct […] logical forms which may be generated from its surface structure." According to May, this generation is accomplished by moving quantifiers like *every* and *some* from their surface positions to positions that serve to disambiguate their scope. So, our sentence (13), then, has two distinct logical forms, which can be represented as follows (cf. May 1977: 13), where (14a) corresponds to reading (i), and (14b) correlates with reading (ii):

(14)   a.   *some woman*$_j$ [*every man*$_i$ [t$_i$ *loves* t$_j$]]

      b.   *every man*$_i$ [*some woman*$_j$ [t$_i$ *loves* t$_j$]]

Given this distinct level of representation – termed 'Logical Form' (LF) – together with its counterpart 'Phonetic Form' (PF), we arrive at the model of grammar depicted in the following schema (cf. Jackendoff 2002: 109), in which Deep Structure and Surface Structure are replaced by D-Structure and S-Structure, respectively, to

account for their different format in contrast to earlier conceptions of these two levels:

(15)

X'-Theory   Lexicon

D-Structure

Transformations

S-Structure

Transformations

PF   LF
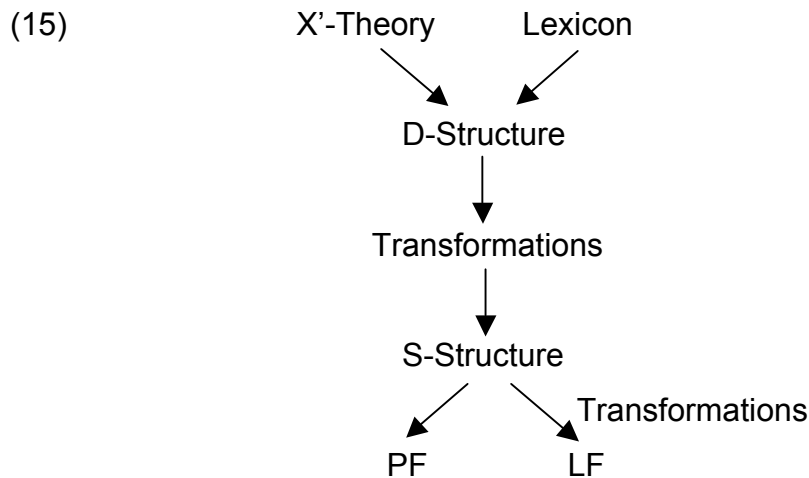
The model provided in (15) postulates four levels of representation. In addition to D-Structure and S-Structure, there are two extra levels of representation, PF and LF. Note that D-Structure is limited to the role of being the starting point of the derivation and that S-Structure is constrained to serve as the so-called 'branching' point, on one branch leading to LF and on the other to PF. The theoretical implications of this architecture were elaborated in great detail in the framework of G(overnment-and-)B(inding) theory (cf. Chomsky 1981). Within this theoretical framework, syntactic operations associated with the different levels of representation are subject to a number of principles that filter out unwanted derivations. These principles and the levels of representation were conceived of as accounting for many empirical phenomena, some of which we illustrated above. Therefore, this conception is regarded as achieving 'descriptive adequacy.' Furthermore, when it is assumed that this rich technology of constraining principles contains certain parameters to be set in the process of language acquisition (for an overview of this approach, see Chomsky & Lasnik 1993), this model of grammar together with the 'Principles-and-Parameters approach' was considered to provide an explanation for language acquisition, and thus, it was viewed as gaining 'explanatory adequacy.' However, this representational view of the language faculty, implying an enriched conception of syntax with several levels of representation, has been abandoned stepwise within recent mainstream generative linguistics – a development seriously marginalized in Jackendoff's discussion of syntactocentrism, to which we now turn.

3.2 Derivational Syntactocentrism

In recent mainstream generative grammar, in addition to descriptive and explanatory adequacy, other criteria of theoretical adequacy (re)emerged. In particular, a premium is placed now on "a level of explanation deeper than explanatory adequacy, asking not only *what* the properties of language are but also *why* they are that way" (Chomsky 2004: 105, emphasis in the original). In asking this question, linguists working within this framework tend to assume that language is the way it is because that is the simplest way it could be. This assumption can be traced back to the early expressed condition "that one of the considerations involved in choosing among alternative analyses is the simplicity of the resulting grammar" (Chomsky 1975 [1955]: 113). In this section, I will point out the serious consequences of this (re)emerged condition for the enriched model sketched at the end of the preceding section. In doing so, I will demonstrate why statements like "[t]he Minimalist Program […] assumes that the structures and derivations of Principles and Parameters Theory are essentially correct" (Culicover & Jackendoff 2005: 88) disregard the changes that took place in mainstream generative grammar. Moreover, outlining this recent version of syntactocentrism will serve as a starting point for comparing recent syntactocentrism with its theoretical alternatives in chapters 5 and 6.

3.2.1 Rethinking D-Structure and S-Structure

Given the condition of simplicity mentioned in the preceding paragraph, the syntactic component should only contain operations truly necessary for pairing sound and meaning of sentence structures. While we look at the model given in (15), the interface levels PF and LF can be regarded as necessary for relating the syntactic structure to sound (i.e. to an articulatory-perceptual system) and to meaning (i.e. to a conceptual-intentional system), respectively. Now, given both the limited role of D-Structure, providing only the starting point of the derivation, and the constrained function of S-Structure, serving merely as the 'branching' point, Chomsky began to take into account the conceptual option "that there are no levels of linguistic structure apart from the two interface levels PF and LF; specifically, no levels of D-structure or S-Structure" (Chomsky 1995: 219). Note that all considerations concerning the role of D-Structure and S-Structure mentioned in section 3.1 are crucially associated with

the analysis of empirical phenomena that generative linguists wanted to cover in their theory. Accordingly, if one adopts the hypothesis that syntax only requires the interface levels PF and LF, it remains to be shown how such empirical phenomena, which justify the postulation of D-Structure and S-Structure, can still be accounted for. Let us first consider one alternative account that regards the role of S-Structure.

In GB theory, a major argument in favor of S-Structure concerns cross-linguistic variation – arguably an important empirical fact. Within GB theory, S-Structure provides a descriptive tool for characterizing cross-linguistic variation, since it helps to describe variation in terms of movement that takes place before or after S-Structure. To give an illustrative example, examine the following constructions from English and French, varying in grammatical order (for analogous examples, see Pollock 1989: 367):

(16) a. *Hans often drinks beer.*

  b. **Hans drinks often beer.*

(17) a. *Hans boit souvent de la bière.*

  b. **Hans souvent boit de la bière.*

In English, the verb *drink* can only be placed after adverbs like *often*, as (16a) illustrates; placing the verb before such adverbs results in an ungrammatical sentence, as (16b) shows. In French, the verb *boit* can only stand in front of the adverb *souvent*, as shown in (17a); if it follows the adverb, the sentence becomes ungrammatical, as illustrated in (17b). Now, in GB theory, these properties are regarded as variations of S-Structure. Since in both languages the sentences mean roughly the same, it is hypothesized that French and English share the same D-Structure, with a nominal subject that occupies the specifier of an 'Inflectional Phrase' (IP) and both the optional adverb and the verb located in the VP (cf. Pollock 1989: 366):

(18) [IP NP I [VP (Adv) V …]]

According to GB theory, the finite verb must ultimately move to the head of IP ('I') due to the agreement relation with the subject of the clause. According to Pollock

(1989), it has long been argued that, in French-type languages, main verbs move to I overtly (i.e. before S-Structure), while, in English-type languages, these verbs remain in situ. However, since the meaning of the sentences is considered to be the same in both languages, the main verb in English must move covertly (i.e. after S-Structure) to the head of IP in order to ensure that the structures in French and in English are identical at LF. Since differences in the two language types – in this case, main verbs either following or preceding VP adverbs – are thus explained by postulating movement operations that apply prior or after S-Structure, this level is essential for describing language variation. However, within the recent minimalist framework, this variation can be captured without postulating any extra level of representation that distinguishes overt from covert movement. Let us see how.

Chomsky (1993) elaborates on the view that lexical items are composed of certain properties, technically termed 'features.' In other words, a word can be thought of as a collection of phonetic, semantic, and syntactic features. Within this concept, finite verbs are considered to have inflectional features in the lexicon as an intrinsic property. Now, in our example, finite verbs move to the abstract inflectional element I, for reasons, as already noted, of agreement with the subject of the clause. So, in more technical terms, a lexical element α is adjoined to I to form the abstract complex [α I]. However, within minimalism, the abstract element I, in contrast to the lexical item α, is considered to be an illegitimate object at the PF-interface. Accordingly, Chomsky supposes that "[i]f the features of α and I match, I disappears and α enters the PF component […]; if they conflict, I remains and the derivation crashes at PF" (Chomsky 1993: 27-28). Assuming this process of 'feature checking,' Chomsky accounts for cross-linguistic variation like (16) and (17) by claiming that features are either 'weak' or 'strong.' If "'strong' features are visible at PF and 'weak' features invisible at PF" (Chomsky 1993: 30-31), 'strong' features must be checked before the grammar splits, otherwise the derivation 'crashes' at PF. On the other hand, 'weak' features are phonologically invisible and only need to be checked at the level of LF. Based on the analysis proposed by Pollock (1989), Chomsky (1991: 422) hypothesizes "that the AGR [= agreement, A.T.] element is 'stronger' in French than in English." So, in other words, inflectional features of English and French differ in terms of strength. In particular, while the head of IP has a 'strong' feature of agreement in French, this position is associated with a 'weak' feature of agreement in English. Since 'strong' features are not legitimate objects at PF, the main verb must raise

overtly (i.e. before S-Structure) in French in order to ensure convergence at PF. By contrast, this overt raising is not forced for convergence in English. Since syntactic operations like movement are considered to be subject to economical conditions (in less technical terms, they are viewed as being 'lazy'), the grammar does not check features unless it has to. Consequently, movement of the verb 'procrastinates' and thus takes place covertly (i.e. after S-Structure). When we apply the notation of Hornstein et al. (2005: 39) to our example, it can be shown that S-Structure and Logical Form are, with respect to the verb position, identical in French, whereas in English, D-Structure and S-Structure are identical:

(19) a. D-Structure:

$[_{IP}$ *Hans* Infl$_{strong-V}$ $[_{VP}$ *souvent boit de la bière*$]]$

b. S-Structure/LF:

$[_{IP}$ *Hans boit*$_{i}$ + Infl$_{\text{strong-V}}$ $[_{VP}$ *souvent* t$_{i}$ *de la bière*$]]$

(20) a. D-Structure/S-Structure:

$[_{IP}$ *Hans* Infl$_{weak-V}$ $[_{VP}$ *often drinks beer*$]]$

b. LF:

$[_{IP}$ *Hans drinks*$_{i}$ + Infl$_{\text{weak-V}}$ $[_{VP}$ *often* t$_{i}$ *beer*$]]$

As demonstrated so far, the distinction of overt and covert movement that is necessary for describing cross-linguistic variation within mainstream generative grammar does not require an extra level of representation like S-Structure. Instead of postulating movement rules that apply prior or after S-Structure, a technology based on feature strength coupled with the economy principle 'Procrastinate' can also account for variation. Note that in the context of challenging the necessity of a level like S-Structure, the question is not whether a feature-based account is better or worse than earlier accounts that imply a level of S-Structure. As Hornstein et al. (2005: 42) point out, the above illustration merely aims at demonstrating "that we can deploy technology that is no less adequate and no less principled, but that does not need SS [= S-Structure, A.T.] at all." Since empirical phenomena can thus be described by alternative tools for description, the level of S-Structure may be just a technical artifact of GB theory.

After having illustrated the view that S-Structure can, in principle, be dispensed with, let us now briefly look at one argument in favor of abandoning the level of D-structure. One prominent empirical motivation for postulating D-structure is that it enables to account for the semantic differences between sentences like (21) and (8), repeated here for convenience as (22):

(21)   *Hans tries to drink beer.*

(22)   *Hans seems to drink beer.*

In (21), *Hans* is understood to be connected semantically both to the verb *try* and to the verb *drink*. To put it succinctly, *Hans* is understood as both a 'tryer' and a 'drinker.' By contrast, in (22), as already sketched in section 3.1, *Hans* is interpreted as playing only a semantic role that is associated with the embedded verb *drink*. In less technical terms, *Hans* is understood as a 'drinker' but not as a 'seemer.' As we saw in section 3.1, D-Structure serves as the starting point of the derivation, providing a pure representation of the grammatical relations (the 'thematic properties') of a sentence. Accordingly, the semantic differences mentioned above are postulated to be transparently represented at the level of D-Structure, which can be notated as follows (cf. Hornstein et al. 2005: 52):

(23)   D-Structure:
$$[Hans_{tryer} \ tries \ [PRO_{drinker} \ to \ drink \ beer_{drink}]_{proposition}]$$

(24)   D-Structure:
$$[\Delta \ seems \ [Hans_{drinker} \ to \ drink \ beer_{drink}]_{proposition}]$$

As (23) illustrates, the verb *try* requires two semantic roles, namely someone who is trying – *Hans* – and a proposition for a complement, that is, the action tried for. Likewise, the verb *drink* requires two roles, namely someone who is drinking and something to drink. These thematic properties are satisfied by *beer* – the 'drink' – and, according to GB theory, by the phonetically empty category 'PRO,' which will be coindexed with *Hans* in the course of the derivation. By contrast, in (24), the verb

seem takes a proposition for a complement, but it does not require a subject associated with any semantic role, as already mentioned in section 3.1. The verb *drink*, on the other hand, requires two semantic roles, a drinker (*Hans*) and a drink (*beer*), transparently indicated at the level of D-Structure, where *Hans* occupies the subject position of the embedded clause before this nominal element moves to the subject position of the matrix clause, thus leaving a trace.

Despite the virtue of D-Structure to depict crucial semantic differences between sentences like (23) and (24), there is, as Hornstein et al. (2005: 53-56) argue, an obvious redundancy in models that postulate this level of representation. Note first that, at D-Structure, understood as the pure representation of thematic properties, the so-called Theta-Criterion applies, requiring that "[e]ach argument bears one and only one θ-role, and each θ-role is assigned to one and only one argument" (Chomsky 1981: 36). Thematic properties of a structure must be, of course, preserved in the course of the derivation, since notions like agent, patient, etc. must be somehow accessible at LF, too, which is viewed as the input to semantic interpretation. So, if there is some principle at LF corresponding to the Theta-Criterion at D-Structure, then there is an unwanted redundancy in the system. If so, and, in Chomsky's words, "[i]f the empirical consequences can be explained in some other way and D-structure eliminated, then […] the θ-Criterion can be dispensed with" (Chomsky 1993: 20). Let us now consider such an option of how the semantic differences between (21) and (22) can be explained in a different way.

Hornstein et al. (2005) claim, roughly speaking, that the system can tell adequate from inadequate representations by merely focusing on the different empty categories in each structure. Look at the following representations (for an analogous example, see Hornstein et al. 2005: 54):

(25)   a.   LF:
            [*Hans*$_i$ *tries* [PRO$_i$ *to drink beer*]]

       b.   *LF:
            [*Hans*$_i$ *tries* [t$_i$ *to drink beer*]]

In (25a), the role 'drinker' was assigned to the empty category PRO, when it was merged with the embedded clause, and the thematic role 'tryer' was assigned to

Hans, when it was merged to the matrix clause. Thus, the subject positions of both verbs – *drink* and *try* – are filled. In (25b), by contrast, *Hans* receives the role of the 'drinker,' when it merges with the embedded clause, but *Hans* cannot receive the role of the 'tryer' in the matrix clause, since this lexical item is connected to the matrix clause by a movement operation and not by Merge. Given this account of singling out (25a) as the adequate representation at LF, Hornstein et al. (2005: 54) formulate a general condition on grammatical operations that replaces the Theta-Criterion, which operates at D-structure, with a principle that applies at the interface level LF:

(26) Theta-Role Assignment Principle
      θ-roles can only be assigned under a Merge operation.

This principle allows for distinguishing between structures like (23) and (24) by simply dictating that θ-roles can be assigned under Merge but not under a movement operation. By elaborating on this option, Hornstein et al. (2005) demonstrate that no 'extra' level like D-Structure has to be postulated to account for structures like (23) and (24). With this option in mind, and recalling that also S-Structure can, in principle, be dispensed with, let us now turn to the general picture we arrived at thus far.

At the outset of this section, I pointed out that, assuming a condition of simplicity, the syntactic component, considered from a conceptual point of view, should only contain the interface levels LF and PF, which are necessary for relating the syntactic structure to meaning and sound, respectively. Based on the illustrations in section 3.1, I demonstrated that the raison d'être of additional levels like D-Structure and S-Structure lies in the desideratum to account for certain empirical phenomena. In the preceding paragraphs, however, it has been shown how prominent empirical phenomena, originally motivating the postulation of D-Structure and S-Structure, are described by assuming alternative tools of description that do not, from a conceptual point of view, require these 'extra' levels of representation. Consequently, according to the minimalist view, the only necessary levels of representation are the interface levels LF and PF. Put together, these considerations yield the minimalist model of the grammar that was assumed in the early 1990s, which can be illustrated as follows (cf. Jackendoff 2002: 110):

(27)   Lexicon
          |
          v
    Merge + Movement
          |
          v
      'Spell-Out'
        /   \
       v   v
     PF   LF
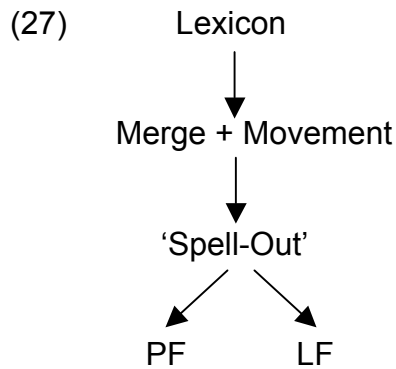
According to this model, lexical items are selected from the lexicon,[9] and Merge together with movement operations build the syntactic structures. At some point in the derivation, termed 'Spell-Out,' the syntactic computation must bifurcate to LF, ultimately semantics, and PF, ultimately phonetics, since sentences are in effect finite in length. Note that this point of Spell-Out is regarded as being different from S-Structure because, as linguists working within minimalism argue, "Spell-Out is not a level of representation. No conditions, principles, etc. can be forced to apply *at that point*" (Boeckx 2006: 80, emphasis in the original).

To conclude, the minimalist conception of derivations, relegating most (if not all) principles of GB theory to the interface levels LF and PF, differs crucially from preceding models. Specifically, different descriptive devices are postulated, and these new tools allow for abandoning the levels of D-Structure and S-Structure. So, even the illustration so far qualifies the claim, cited at the outset of the section, that "[t]he Minimalist Program […] assumes that the structures and derivations of Principles and Parameters Theory are essentially correct" (Culicover & Jackendoff 2005: 88). Crucially now, the changes sketched in this section have been already in place in the

---

[9] Strictly speaking, Jackendoff's (2002) graphic illustration is inaccurate in this regard. Within minimalism, a syntactic derivation does not start with the lexicon itself but with a set of lexical items selected from the lexicon, with a 'numeration,' understood as "a set of pairs (LI, *i*), where LI is an item of the lexicon and *i* is its index, understood to be the number of times that LI is selected" (Chomsky 1995: 225). This 'numeration' replaces D-Structure as the starting point of a derivation. However, it is not a level of representation, since it is not subject to certain linguistic conditions, as D-Structure is (e.g. to the 'Theta-Criterion' mentioned above and to X-bar theory mentioned in section 2.1). Rather, the numeration merely functions as the formal object that provides the relevant lexical items that feed the computational system. Such an object is necessary within a minimalist framework because, as we will see in the next section, derivations are compared for economy purposes. Since, for obvious reasons, it only makes sense to compare derivations for economy purposes, if they start from the same numeration, this descriptive device can be regarded as a conceptual necessity.

early 1990s, at the time when the Minimalist Program began to take shape. Tellingly, as already mentioned, Jackendoff merely refers to this early developmental stage, even in his recent publications (cf. Jackendoff 2007: 37). However, the model introduced so far has undergone further significant modifications. Let us take a closer look at these recent changes in the following section.

3.2.2 Rethinking LF and PF

As mentioned in the last section, the early minimalist model cited in Jackendoff's publications is merely the starting point, the first conceptual consequence of applying conditions of simplicity to mentalist models of the human language faculty within mainstream generative grammar. By questioning the conceptual necessity of levels like D-Structure and S-Structure and by relegating the principles that operate at these levels to the interfaces, generative linguists let their focus shift more and more to the derivational process itself. As a consequence, one prominent issue in minimalism is to compare derivations that start from the same set of lexical items, the same 'numeration,' for economy purposes. These economy purposes concern issues like the number of derivational steps and the length of movement operations. Within minimalism, aspects of derivational economy and simplicity are crucial for the convergence of a derivation, that is, for the two options that a derivation has: it can either be interpreted at the interface levels LF and PF (thus converges), or it cannot be interpreted (thus 'crashes'). Pushing this economical view further, Chomsky (1995: 220) argues that "[l]ess economical computations are blocked even if they converge." Hence, according to this claim, even if two derivations that are compared with each other can both be interpreted at LF and PF, it might be that one of them manages to converge more economically than the other. To illustrate this, let me show you the following examples (cf. Chomsky 1995: 344):

(28) a. *There seems to be someone in the room.*

  b. **There seems someone to be in the room.*

Both derivations of the sentences given in (28) are regarded as starting with the same numeration, as (29) illustrates in set-theoretic format, in which a numeral index indicates the number of instances of a particular item available for the computation. For the sake of illustration, I will adopt the simplified notation of Lasnik et al. (2005: 164), which ignores purely grammatical categories such as Tense:

(29) $\{there_1, seems_1, to_1, be_1, someone_1, in_1, the_1, room_1\}$

The derivations of both (28a) and (28b) are considered to be identical up to the derivational step that is represented in the following schema, in which the selection of a lexical item from the numeration is indicated by reducing its index by 1 (for a similar illustration of what follows, see Lasnik et al. 2005: 165-166):

(30) $\{there_1, seems_1, to_0, be_0, someone_0, in_0, the_0, room_0\}$
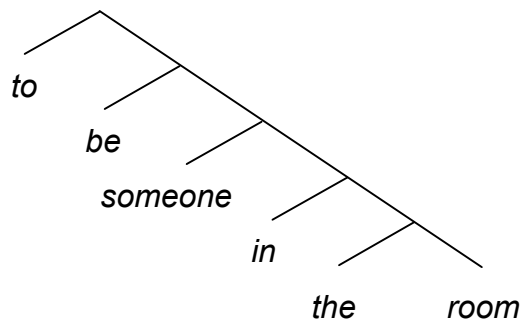
However, the derivations of (28a) and (28b) differ from the next derivational step on, as shown in (31), where (31a) represents the next derivational step that is implied by (28a) and (31b) depicts the step that yields the ungrammatical structure (28b):

(31) a. $\{there_0, seems_1, to_0, be_0, someone_0, in_0, the_0, room_0\}$

there
to
be
someone
in
the   room

b. $\{there_1, seems_1, to_0, be_0, someone_0, in_0, the_0, room_0\}$
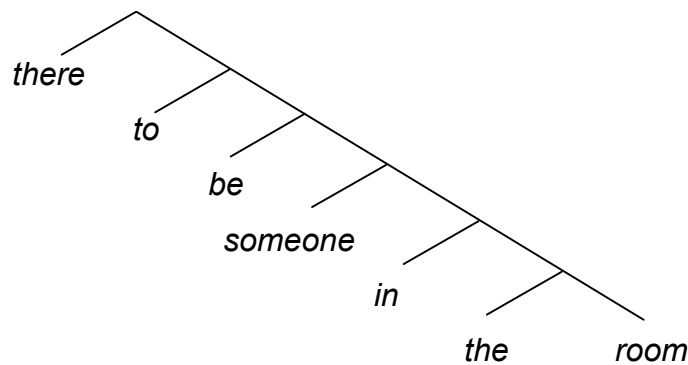
someone
to
be
t
in
the   room
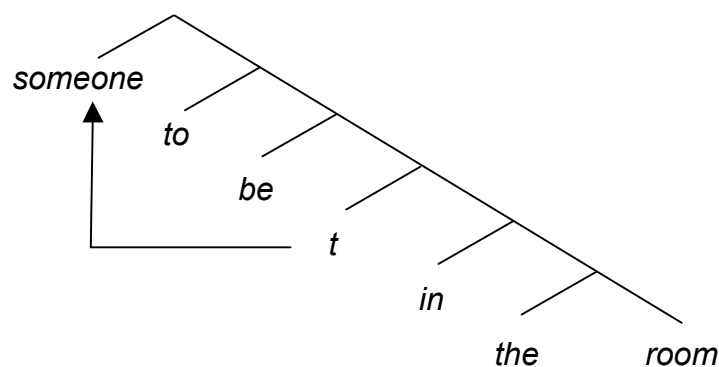
Let us study both derivations in detail. In (31a), the expletive *there* is merged in order to occupy the infinitival subject position, while *someone* remains in situ. By contrast, in (31b) *someone* is moved to the infinitival subject position, whereas *there* remains in the numeration. Once a specific option to occupy the subject position is taken, the next steps of the derivation, according to minimalist syntax, necessarily follow. So, after having merged the finite verb *seems in* (31a), the subject position of the matrix clause must be filled with *there*, since *someone* is not allowed to cross the intervening expletive *there*, as (32) illustrates:

(32) *[*someone*$_i$ *seems* [*there to be* t$_i$ *in the room*]]

Accordingly, *there* has to move and thus leaves a trace, which can be represented as follows:

(33)   a. $\{there_0, seems_0, to_0, be_0, someone_0, in_0, the_0, room_0\}$

there

seems

t

to

be

someone

in

the   room
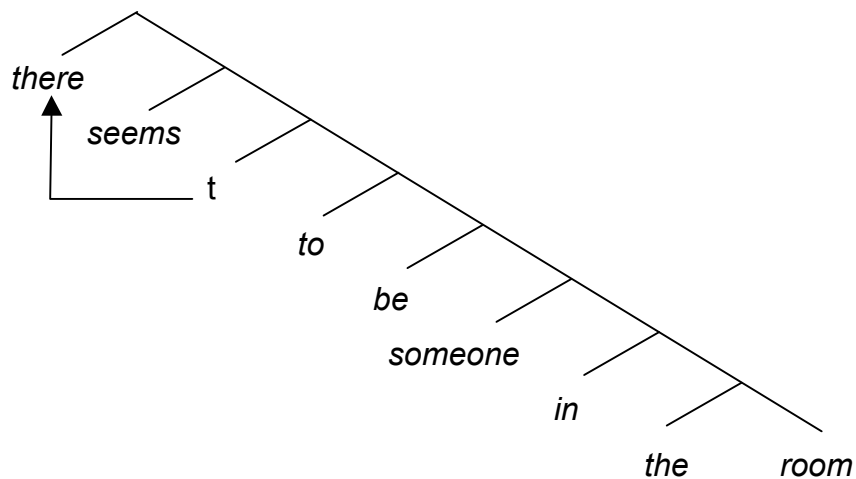
By contrast, examine the steps following (31b). After having merged the verb *seems*, the expletive *there* can be merged to occupy the subject position of the matrix clause, only if *there* is available in the numeration. This can be depicted as in (34):

(34) $\{there_0, seems_0, to_0, be_0, someone_0, in_0, the_0, room_0\}$

there

seems

someone

to

be

t

in

the   room
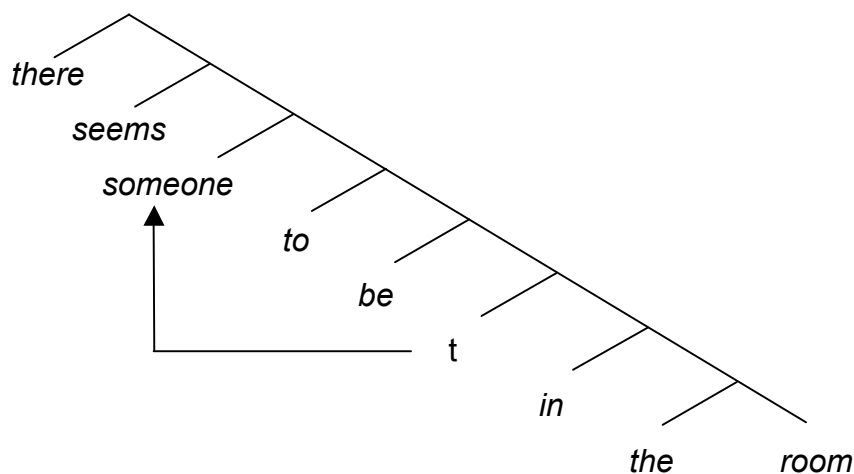
Since both derivations converge with respect to filling both the infinitival and the matrix subject position, the question is why (28a) is grammatical and (28b) is odd. Chomsky's suggestion is that the derivation that yields (28a) is more economical than the derivation underlying (28b). Note that looking at the whole derivation does not help us in this case, since both derivations, taken as a whole, contain the same number of steps. Accordingly, general criteria like length of derivation fail to capture the difference concerning grammaticality. Even if one adopts Chomsky's (1995: 373)

view that, in contrast to movement operations, "Merge is costless, so it can apply freely," we do not arrive at an adequate explanation, since both derivations contain exactly one movement operation. In order to explain, on economical grounds, why (28a) is fine and (28b) is ungrammatical, Chomsky adopts a more local, that is, a stepwise view of derivations. As the derivations differ from the steps (31a) and (31b) on, let us once again focus on these crucial points, where the two derivations split.

After the steps that were taken up to (30), there are, as noted above, only two options to fill the infinitival subject position: merging the expletive *there* or moving *someone*. Since the movement of *someone*, as argued within minimalism, is more costly, the most economical way to fill the infinitival subject position at this specific point of the derivation is merging *there*. So, as at this point the option of merging *there* is regarded as more economical than moving *someone*, the difference between (28a) and (28b) concerning grammaticality can be accounted for by referring to a criterion of derivational economy. That is, to put it simply, (28a) is fine because *there* is merged at this point, whereas (28b) is odd, since *someone* is moved at this specific stage of the derivation, although the more economical option of merging *there* would have been available. Note that it does not matter that, at the end of the derivation of (28a), *there* is moved from the embedded to the matrix sentence because, according to this strict stepwise view, derivational decisions are made locally. So, while the derivational step given in (31b) is an economy violation, the movement of *there* to the matrix subject position in (33a) works because no alternative to this option exists at this stage of the derivation.

With this preference for 'Merge over Move' and its local application in mind, let us now turn to the following case (cf. Castillo et al. 1999: 6), whose analysis will help to capture the need of reconsidering the early minimalist model of grammar with its two levels of representation, LF and PF:

(35)   *There was a rumor that someone was in the room.*

In this case, *someone* has moved to fill the subject position of the embedded clause, although the expletive *there* was present in the numeration and would have been available for insertion, as schematically shown in the following representation (for an analogous example, see Lasnik et al. 2005: 237-241):

(36) $\{there_1, was_1, a_1, rumor_1, that_1, someone_0, was_0, in_0, the_0, room_0\}$

someone

was

t

in

the   *room*
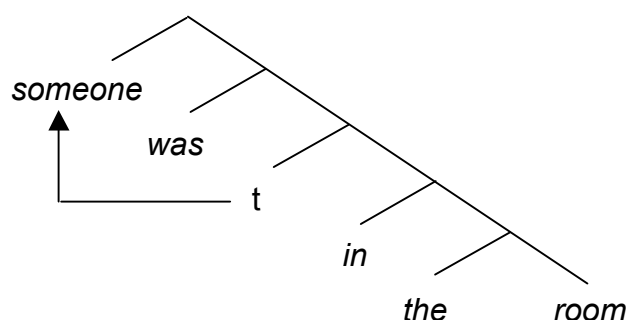
According to the economy considerations discussed above, the derivational step in (36) should result in an ungrammatical sentence, since it seems to violate the economy-based preference 'Merge over Move.' However, (35), resulting from the derivational step schematized in (36), is perfectly fine. In order to explain constructions like (35), Chomsky (2000: 106) formulates an additional assumption concerning the nature of numerations, or 'lexical arrays' (LAs), namely "that at each stage of the derivation a subset LA_i is extracted, placed in active memory." In his discussion of such 'subnumerations,' he introduces the term 'phase' to refer to syntactic objects (SO) derived from such subnumerations. So, in his words, he takes "a *phase* of a derivation to be an SO derived […] by choice of LA_i" (Chomsky 2000: 106, emphasis in the original). Therefore, derivations proceed by phases, and each phase has its own subnumeration. When we apply this proposal to (35), the embedded clause is considered to have its own subnumeration, indicated by the following notation:

(37) $\{there_1, was_1, a_1, rumor_1, \{that_1, someone_0, was_0, in_0, the_0, room_0\}\}$

According to (37), the expletive *there* is not present in the subnumeration that corresponds to the embedded clause, and, accordingly, *there* is not available for lexical insertion at this point of the derivation. Consequently, the economy-based preference 'Merge over Move' is not violated. After this illustration of rather 'strong' derivational thinking within recent mainstream generative linguistics, let us now broaden our presentation again and return to the issue of linguistic levels.

When we assume that the computation of convergence proceeds phase by phase, the model of the grammar no longer contains fully-fledged structures at the interface levels LF and PF, as postulated in the early minimalist model given in (27).

In other words, since computed phases are conceived of as already interpretable at the interfaces, there is not only one point of Spell-Out that hands the whole syntactic structure to LF and PF, but there are, depending on the number of phases, multiple points of Spell-Out handing over partial bits and pieces of structure to the interface components. So, while syntactic structures were hitherto considered to represent the whole sentence at some particular level of representation, it is now argued that only parts of these structures are represented at the interfaces. To put it differently, "while there are still what might be called PF and LF components, there are no *levels* of PF and LF" (Lasnik 2005: 82, emphasis in the original). Consequently, the grammar model resulting from these considerations can in effect be regarded as 'level-free,' as indicated in (38) by 'lf' and 'pf' in lower case (cf. Boeckx 2006: 77):

(38)
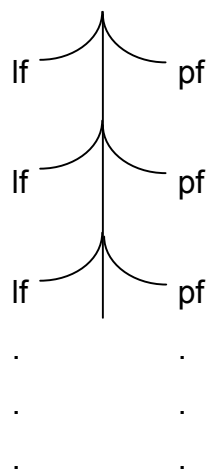
As we will see in chapter 5 of this thesis, this strong derivational perspective on the language faculty has changed the relation of syntactic computation to aspects of phonology and meaning substantially, thus challenging, as I will argue, the somehow antiquated notion of syntactocentrism provided by Jackendoff. Moreover and more importantly, this level-free architecture together with the minimalist conception of impoverished syntactic structures, sketched in section 2.1, entails significant points of convergence with theoretical alternatives to syntactocentrism – alternatives to which we now turn.

4. Alternatives to Syntactocentrism

The preceding chapter provided an illustration of the significant shift from representational to derivational syntactocentrism within recent mainstream generative grammar. As was argued in light of the conceptual changes sketched above, Jackendoff's notion of syntactocentrism can be regarded as antiquated, inasmuch as his discussions are based on a grammar model of the early 1990s. Since an up-to-date comparison of syntactocentrism with its theoretical alternatives must, in my view, incorporate the recent theoretical shift illustrated so far, one key to approach such a comparison has now been established. The other key, of course, is to study theoretical alternatives to syntactocentrism. In this chapter, therefore, I will present two prominent alternative approaches. In particular, I will first turn to approaches within the general movement of Cognitive Linguistics, whose scholars, over the years, compared to syntactocentrism, "went to the other extreme and denied syntax *any* independent role" (Jackendoff 2007: 43, emphasis in the original). Having illustrated this diametrically opposed view, I will then turn to Jackendoff's own approach, the 'Parallel Architecture,' which can be regarded as an intermediate position between the two extremes Cognitive Linguistics and syntactocentrism or, in Jackendoff's (1996: 98) words, as an "appropriate middle ground." In the final section of this chapter, given these alternative approaches to modeling the mental architecture of the human language faculty, I will provide a short outlook to perspectives of convergence between syntactocentrism and its theoretical alternatives.

4.1 Cognitive Linguistics

Of course, Cognitive Linguistics is not a monolithic approach. Rather, it can be characterized as a general movement consisting of many different approaches. And yet, since it can be regarded as "a cluster of broadly compatible approaches" (Geeraerts & Cuyckens 2007: 3), there are some crucial assumptions that are widely shared among scholars that participate in that movement. In this section, for the sake of contrast to the syntactocentric view, I will focus on how syntactic representation is conceived of within Cognitive Linguistics and therefore, draw attention to so-called

'construction grammar models,'[10] commonly assumed within Cognitive Linguistics and particularly concerned with the nature of syntactic representation.

Historically, construction grammar models arose out of several problems in analyzing idioms within a syntactocentric framework. Since the meaning of idioms, such as *kick the bucket* (meaning: 'die'), cannot, in the eyes of many, be derived by compositional rules of semantic interpretation like the ones sketched in section 1.3, Chafe (1968: 127) has already pointed out "the clear inability of generative syntax to account for a phenomenon as pervasive in language as idiomaticity." Fillmore et al. (1988) argue that the pervasiveness of idiomaticity emphasized by Chafe strongly suggests that "[t]hose linguistic processes that are thought of as irregular cannot be accounted for by constructing lists of exceptions" (Fillmore et al. 1988: 534), as many linguists committed to syntactocentrism commonly do in order to explain these idiosyncratic phenomena that contradict compositionality. For Fillmore et al., the absurdity of excluding idiomatic expressions from the 'core' of the grammar by relegating them to the 'periphery' becomes clear in the case of idioms that are 'formal' to a greater or lesser degree instead of being 'substantive.' According to Fillmore et al. (1988: 505), substantive idioms are expressions whose "lexical make-up is (more or less) fully specified," whereas formal idioms are 'lexically open' idioms, meaning that they allow for a wide range of possible items to instantiate the idiom and thus show a great deal of productivity. Concerning this point, look at some of their examples (cf. Fillmore et al. 1988: 505-512):[11]

(1) *It takes one to know one.*

(2) a. *He trips the light fantastic.*

    b. *They trip the light fantastic.*

---

[10] Following Croft (2007: 463), I will set *construction grammar* in lower case, thereby indicating that the illustration in this chapter is not constrained to the specific theory originating in Fillmore et al. (1988) – 'Construction Grammar' – but rather aims at covering several theories of grammar within Cognitive Linguistics that assume the basic form of syntactic structures to be constructions.

[11] In presenting these examples, Fillmore et al. (1988: 512) use capital letters to mark main stress because they need this descriptive dimension for their in-depth analysis of '*let alone* constructions' (cf. Fillmore et al. 1988: 510-533). Since I will not delve into this aspect of their detailed analysis, this representational device is omitted here.

(3) a. *I blow/blew my nose.*

b. *You blow/blew your nose.*

(4) a. *I barely got up in time to eat lunch, let alone cook breakfast.*

b. *I doubt you could get Fred to eat shrimp, let alone Louise squid.*

These examples show that differentiating between substantive and formal idioms "is actually a gradient or cline rather than a simple two-way distinction" (Fillmore et al. 1988: 505, n. 3). In particular, example (1) illustrates a case of idiomaticity in which the lexical make-up is fully specified, that is, every lexical element is fixed. In (2), the idiomatic expression *trip the light fantastic* can also be regarded as highly fixed, except for inflectional categories. So, both the singular (*trips*) and the plural inflection (*trip*) work well. By contrast, idioms like (3) do not only possess this inflectional flexibility but also show flexibility regarding one argument of this expression because the 'nose possessor' can vary, as (3a) and (3b) demonstrate. Concerning (4), Fillmore et al. (1988: 511) argue "that *let alone* sentences possess a collection of properties that is unique to this particular family of expressions," and, accordingly, they regard these sentences as idiomatic expressions and thus as instances of one and the same construction. As (4a) and (4b) indicate, *let alone* sentences have open slots for all categories, except for the connective *let alone* itself. So, regarding structures like (4a) and (4b) as formal idioms, Fillmore et al. (1988: 534) hypothesize that their approach could be "powerful enough to be generalized to more familiar structures, in particular those represented by individual phrase structure rules." Adopting this suggestion, Goldberg (1995), for instance, claims that there are formal idioms whose elements are all lexically open. One instance discussed by her is the 'ditransitive construction,' as exemplified by the following sentence (for an in-depth analysis of this construction in English, see Goldberg 1995: 141-151):

(5) *Hans gave Peter a beer.*

Instead of presupposing autonomous syntactic operations like the rewrite rules in (6), which generate a structure to be interpreted by semantics and phonology, Goldberg analyzes these rules as the most formal, that is, the most schematic constructions of

a language, as illustrated in (7), where the syntactic structure of this construction is analyzed as "a pairing between a semantic level and a syntactic level of grammatical functions" (Goldberg: 1995: 51):

(6) S → NP + VP
     NP → (Det +) N
     VP → V + NP + NP

(7) [Subj V Obj$_1$ Obj$_2$]
     X CAUSES Y to RECEIVE Z

Representations like (7) demonstrate the substitution of autonomous syntactic rules by means of schematic constructions. Hence, approaches like Goldberg's exemplify that construction grammar models aim at a uniform representation of all grammatical knowledge in terms of generalized constructions and thus abandon the concept of postulating distinct components like the phonological, the semantic, and the syntactic component. Note that, by linking phonological, semantic, and syntactic information without supposing an autonomous rule component, constructions resemble lexical items within mainstream generative linguistics. However, they differ from lexical items in being not always atomic. So, as construction grammar models aim at a uniform representation format of both atomic and structurally more complex items, "lexical items […] may be viewed […] as constructions themselves" (Fillmore et al. 1988: 501), and a single word is thus regarded as the smallest case of a construction. As a result, the syntactocentric distinction between syntax and lexicon – the syntax deals with multiword expressions generated by syntactic rules, whereas the lexicon provides fixed atomic elements – is replaced by a conception according to which "lexicon, morphology, and syntax form a continuum of symbolic units" (Langacker 1987: 35). The notion of 'symbolic unit' is important here. Within construction grammar models, grammatical units such as morphemes, words, or phrases are viewed as fundamentally symbolic, that is, as (at least partially arbitrary) pairings of form and meaning. The following schema may help to bring out this symbolic conception of a construction (cf. Croft 2007: 472):

(8)

```
┌─────────────────────────────────────┐  ◄───── CONSTRUCTION
│  ┌───────────────────────────────┐   │
│  │      Syntactic properties     │   │
│  │    Morphological properties   │◄──┼──── FORM
│  │     Phonological properties   │   │
│  └───────────────────────────────┘   │
│              ┃◄──────────────────────┼──── Symbolic correspondence (link)
│  ┌───────────────────────────────┐   │
│  │      Semantic properties      │   │
│  │      Pragmatic properties     │◄──┼──── MEANING
│  │  Discourse-functional properties │
│  └───────────────────────────────┘   │
└─────────────────────────────────────┘
```
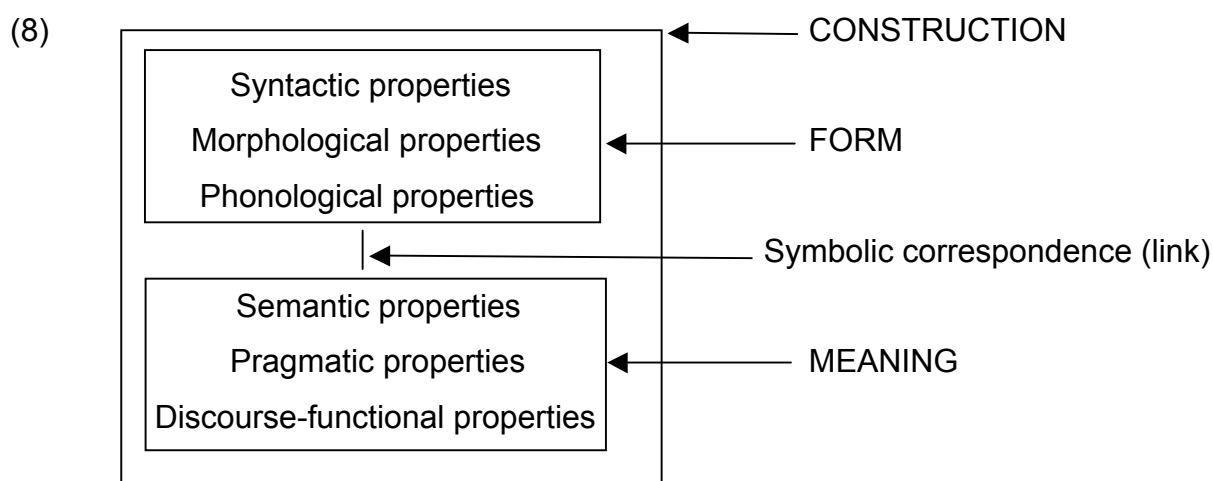
This symbolic account of linking form and meaning is emphasized the most in 'Cognitive Grammar,' one branch of Cognitive Linguistics. So, while mainstream generative grammar hypothesizes that linking form and meaning is accomplished by syntactic operations, Langacker (1987: 76), who focuses on this linking, argues that the correspondence between form and meaning is due to "conventional symbolization of semantic structure." In Cognitive Grammar, all aspects of grammatical structure are ultimately derived from a 'semantic space,' that is, in Langacker's (1987: 76) words, from the part of "the multifaceted field of conceptual potential within which thought and conceptualization unfold." So, for instance, consider phonological properties. Langacker argues that sounds are actually concepts, since "the cognitive representation of linguistic expressions derives most directly from auditory impressions, and only indirectly from the sound waves that give rise to these impressions" (Langacker 1987: 78). More specifically, Langacker regards these conceptualized auditory impressions as auditory 'images' that are used to categorize acoustic input. Note that these images are viewed as primary not only in the context of speech perception but also in the context of speech production, insofar as these cognitive representations determine the specific motor sequence associated with producing a specific sound. So, regarding sound as being derived from conceptual representations, Langacker (1987: 79) concludes that "phonological space should […] be regarded as a subregion of semantic space." Given this view that all structural properties of language can ultimately be derived from semantics, this version of Cognitive Linguistics can be viewed as diametrically opposed to syntactocentrism. Compared to this 'semanti-

cocentrism,'[12] Jackendoff's own proposal can be regarded as a less extreme alternative to syntactocentrism, as we will see in the next section.

4.2 The Parallel Architecture

As we saw in the previous section, the construction grammar model within Cognitive Linguistics rejects the concept of postulating a syntactic component that determines both phonology and semantics. More importantly, this approach abandons the componential model of grammatical knowledge as a whole and argues in favor of a uniform representation of this knowledge in terms of constructions. As I illustrated, some approaches like Cognitive Grammar, subscribing to this construction-based view, claim that every aspect of language can be derived from semantics and thus can be regarded as a diametrically-opposed alternative to syntactocentrism. Let us now turn to Jackendoff's own approach. In contrast to theories committed to the framework of Cognitive Linguistics, Jackendoff does not reject the componential model of grammar because he assumes an autonomous syntactic component. Yet, unlike mainstream generative linguistics, he claims that both the phonological and the semantic component must be regarded "as generative completely on a par with syntax" (Jackendoff 1997: 39). To capture his account, let us consider his arguments for an independent role of both the phonological and the semantic component. Since, over the years, Jackendoff has provided various arguments in favor of this view, most of them repeatedly, I will limit myself to prominent examples presented in his most concise illustration (cf. Jackendoff 2003: 655-658). Let us begin with the phonological component.

According to Jackendoff, the assumption that the phonological component is a generative system in its own right, in many aspects independent of syntax, is mainly due to insights from phonological theory in the mid-1970s, when phonology "came to be thought of having its own autonomous structure, in fact multiple structures" (Jackendoff 2003: 655). When he speaks of 'multiple structures,' Jackendoff refers to such entities as segments, syllables, and intonational phrases, all investigated intensively in the period he is concerned with. Jackendoff argues that these structures do not

---

[12] To my knowledge, this term has so far only been used by van der Hulst (2006: 666), but in a slightly different context. However, to sharpen the contrast between Langacker's approach and mainstream generative grammar, it may be, in my view, an appropriate term.

correspond one by one to the units proposed for the syntactic component. For a better understanding, notice the mismatches between intonational phrases and syntactic structures in the following example (cf. Jackendoff 2003: 656):

(9)   a.   Syntactic Structure:

[NP *this*] [VP *is* [NP *the cat* [CP *that* [VP *chased* [NP *the rat* [CP *that* [VP *ate* [NP *the cheese*]]]]]]]]]

       b.   Intonational Phrases:

[IntP *this is the cat*] [IntP *that chased the rat*] [IntP *that ate the cheese*]

(9a) illustrates with the help of bracketing notation that, on the syntactic side, the sentence is a complex right-embedded structure. In particular, it contains two relative clauses – *that caught the rat* and *that stole the cheese* – each notated as 'Complementizer Phrase' (CP), since they are introduced by the complementizer *that*. These two CPs – with their own internal structure – are both hierarchically right-embedded constituents of the NP *the cat*, which is the complement of the verb *is*. In contrast to this complex embedding structure, consider now the flat prosodic structure of this sentence given in (9b). While the sentence is pronounced with a standard intonation, the prosody of the sentence is balanced into three intonational phrases (IntP), indicating a pause between the three units. At first sight, the intonational structure seems to be determined by syntactic structure, since its boundaries fall at the beginning of major syntactic constituents. Yet, the ends of the IntPs, as the comparison of (9a) and (9b) makes clear, do not always correspond to the ends of the correlating syntactic constituents. Based on this observation, Jackendoff argues that the phonological component must contain principles for prosodic structuring that are independent of syntactic structure – for example, a tendency to generate structures according to rhythmicity or, to be more specific about our example, a preference for parallelism. All in all, based on this case and further examples, Jackendoff (2003: 657) concludes "that phonological structure is not just a passive hand-me-down derived from low-level syntax – it has its own role in shaping the tonality of linguistics structure" and thus can be regarded as an autonomous generative system.

Let us now turn to the semantic component. According to Jackendoff, this component must be conceived of as an independent generative system, too. Analogous

to multiple structures postulated for the phonological component, he claims that semantics can be factored into independent structures as well. To illustrate his reasoning, let us focus on the first argument he gives in favor of this view. Consider the following sentences, where capital letters indicate main stress (cf. Jackendoff 2003: 657):

(10)   a.   *The BEAR chased the lion.*

       b.   *The bear chased the LION.*

       c.   *The bear CHASED the lion.*

As (10) shows, in addition to the level of propositional structure, which, roughly speaking, encodes the information of 'who did what to whom,' there is also the level of information structure, encoding the fact that, as Halliday (1967: 242) defines it,

> the speaker maps on to the clause, as defined in sentence structure, a structure of a different kind in terms of information units, by which he organizes the discourse into message blocks and specifies the status of the components of the message as new information or otherwise.

While we suppose the partitioning of the message, which is conveyed by sentences, into 'new' versus 'old' information as a level of the semantic component, the sentences given in (10) demonstrate, as Jackendoff argues, that changes concerning this level of the semantic component do not show up in syntax at all. In particular, while the propositional structures in (10a)-(10c) are identical, the element interpreted as providing new information (the 'focus') is different: in (10a), the element interpreted in this way is *the bear*, in (10b) it is *the lion*, and in (10c) it is the verb *chased*. Based on this simplified illustration of information structural properties of a sentence, Jackendoff (2003: 658) concludes "that there are aspects of semantics that have no impact on syntax."

On the basis of other cases, provided to underline the independence of the semantic component, and by referring to his argumentation concerning the independence of the phonological component, Jackendoff, for over ten years now, argues for a conception of grammar according to which "[t]he grammatical structure of a sentence can be regarded as a triple" (Jackendoff 1997: 38). According to this view,

three independent components must be postulated in the overall picture of the grammar. This can be graphically depicted as the so-called 'Parallel Architecture' (cf. Jackendoff 2003: 659):

(11)   Phonological   Syntactic   Semantic

formation   formation   formation

rules   rules   rules

Phonological   Syntactic   Semantic

structures   structures   structures

Interface   Interface

Interface
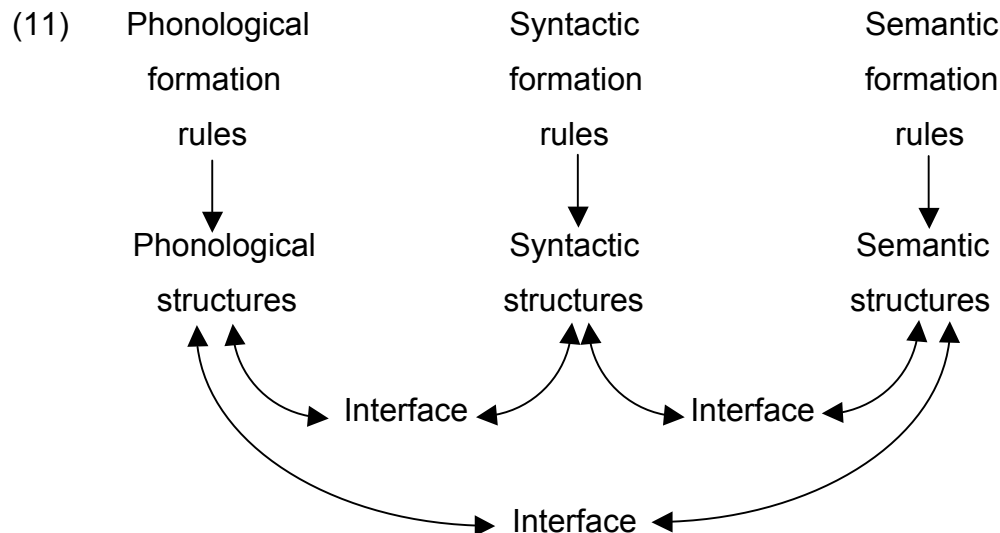
The model given in (11) contains three independent generative components of formation rules – the phonological, the syntactic, and the semantic formation rules – each determining their own type of structure. These structures are linked by interface components, which can be viewed as rules that determine the specific linking at each interface. One example for a rule linking semantic and syntactic structures is the 'Theta-Criterion,' already mentioned in section 3.2.1, which dictates that each syntactic argument must bear one and only one θ-role, and each θ-role is assigned to one and only one argument. Since Jackendoff, as I indicated by having presented two of his major examples, hypothesizes that such an isomorphism does not exist in every case, he claims that several aspects of each of the three structural components are 'invisible' for the interfaces, thereby supporting his view that each of the three components plays a rather independent role in the grammar.

In addition to the independence that is hypothesized with respect to the phonological and the semantic component, one significant aspect, distinguishing this model from syntactocentric models, is that there is no lexicon component that serves as the input for syntactic derivations. As already proposed by Jackendoff (1997), a lexical element is, according to his approach, regarded as a triple of phonological,

syntactic, and semantic features, which can be formalized as follows (cf. Jackendoff 1997: 89):

(12)   Wd$_i$ N$_i$ [$_{Thing}$ TYPE: CAT]$_i$

         /|\ |
        k æ t   ┌──────┐
                         │ count │
                         │ sing │
                         └──────┘
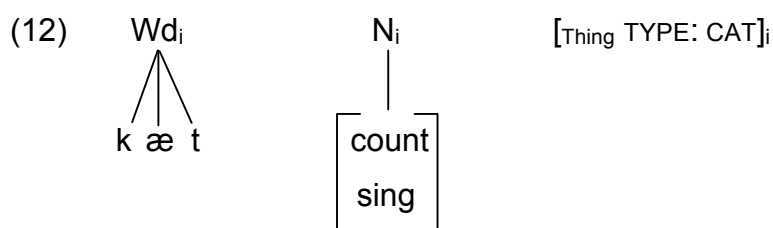
According to (12), lexical items consist of a feature set that can be represented formally as a triple. Specifically, every structure of that triple is associated with an index ('i') that indicates the correspondence relations between the structures. So, when we return to the overall picture of the Parallel Architecture given above, it is essential that "*a lexical item is to be regarded as a correspondence rule, and the lexicon as a whole is to be regarded as part of the interface modules*" (Jackendoff 1997: 89, emphasis in the original). In other words, in contrast to the different syntactocentric models sketched in chapter 3, lexical items are viewed as not belonging to a separate component (the 'lexicon') that is added to the syntactic rules and principles of the grammar. Rather, lexical items are regarded as rules themselves. Crucially now, elements like idioms, mostly considered to be problematic cases within a syntactocentric framework, can be easily accounted for by regarding them as lexical entries that are larger than words and coindexed in a special way. In more precise words, look at, for instance, the idiom *kick the bucket* once again. Since its meaning, according to Jackendoff, cannot be derived compositionally, an account of inserting individual words independently, complete with their meanings, is not an adequate solution to account for such items because in cases like *kick the bucket*, the individual words can be regarded as having no meaning in isolation. Now, within the framework of the Parallel Architecture, *kick the bucket* can be represented as follows (cf. Jackendoff 1997: 169):

(13) Wd_i Cl_j Wd_k VP_x $[\text{DIE}([\ ]_A)]_x$

       k ɪ k   ð ə   bʌkət V_i NP

                                   Det_j N_k
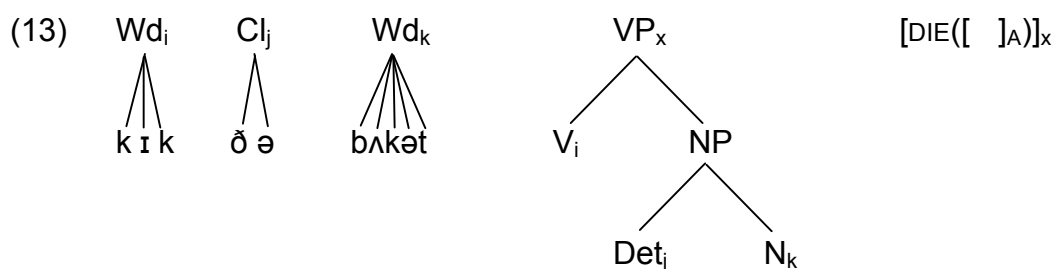
According to this representation, the whole VP *kick the bucket* is treated as a lexically listed unit that is coindexed with the phonological structure – containing two phonological words (Wds) and the clitic (Cl) *the* – in a 'normal' way, but that is coindexed as a whole (by the subscript 'x') with the semantic structure, which contains the concept DIE and an empty argument slot, subscripted 'A.' So, the interface rules, in this case, lack indices that connect individual words to individual parts of the semantic structure. Note that, by postulating that lexical elements account for matching sound and meaning, Jackendoff advances a view according to which "[a] strict separation of lexicon and grammar, like a strict separation of word lexicon and idiom lists, may prove to be but a methodological prejudice" (Jackendoff 1995: 155-156). Given this aspect of Jackendoff's theory, proponents of construction grammar models have already pointed out that by developing and spelling out his Parallel Architecture, Jackendoff "is apparently moving in the direction of construction-based grammar, which makes the 'interface' the heart of the entire grammar" (Goldberg 1996: 14). Nonetheless, since Jackendoff proposes an independent syntactic component, his account does not truly fit to construction grammar models and can therefore better be characterized as an intermediate position between models of grammatical knowledge within Cognitive Linguistics, and conceptions within mainstream generative grammar. As, to me, this mediating enterprise seems to be worth pursuing, the remainder of the chapter provides a short reflection on perspectives of convergence between syntactocentrism and its theoretical alternatives.

4.3 Syntactocentrism and its Alternatives: Perspectives of Convergence

One of the aspects characterizing Jackendoff's work over the years is his impetus "of restoring some degree of much-needed unity to the field of linguistics" (Jackendoff 2002, xii). He places a premium on this unity, since, throughout his work, he is concerned with the integration of linguistics with other areas of cognitive science. Since most of his publications are explicitly addressed at an interdisciplinary audience, his attitude of taking into account insights from a variety of linguistic and non-linguistic disciplines is widely accepted as "a unifying starting point" (Ritter 2005: 121) to discuss and further strengthen the role of linguistics within cognitive science. Accordingly, on general grounds, his open-mindedness to both Cognitive Linguistics and mainstream generative grammar seems to be worth supporting. Adopting this open-mindedness, one can indeed argue that syntactocentrism and its theoretical alternatives share not only their history (cf., e.g., Harris 1993) but also, if only at a highly abstract level, their general conceptual underpinnings. So, as already indicated at the beginning of this thesis, Goldberg (2006: 4), referring to constructionist approaches and mainstream generative grammar, points out that "[b]oth approaches agree that it is essential to consider language as a cognitive (mental) system," and Fillmore et al. (1988: 501) explicitly situate their approach "in the broad generative tradition." Promoting a dialogue between these approaches, as Jackendoff steadily does, is, therefore, a quite realistic enterprise worth endorsing.

However, as the illustration of recent syntactocentrism in chapters 2 and 3 has already demonstrated, Jackendoff is not always aware of current developments within mainstream generative grammar, thus causing linguists working within that framework to note that "Jackendoff often operates with antiquated notions of grammar" (Boeckx & Piattelli-Palmarini 2007: 406). On the other hand, proponents of approaches that are situated within Cognitive Linguistics point out, in a similar manner, that they are "struck by the lack of understanding Jackendoff brings to topics in which we ourselves are far more deeply immersed" (McClelland & Bybee 2007). To put it succinctly, Jackendoff seems to be caught in the middle.

Of course, to keep up with the rapidly growing literature in all areas of the different approaches we touched on so far is beyond anyone's ability. In this thesis, as chapters 2 and 3 make most clear, my aim is to amend some of Jackendoff's assertions regarding recent syntactocentrism. However, instead of claiming that the alter-

native descriptions he provides are essentially misguided, I argue that, once the consequences of the major changes within syntactocentrism are taken seriously, some (not all!) of Jackendoff's objections to this view of grammar disappear. In order to explore this hypothesis in detail in the next chapter, I will focus on one specific issue presented by Jackendoff as a crucial argument in favor of the Parallel Architecture: the place of information structure within the grammar. After the reevaluation of Jackendoff's argument in light of recent syntactocentric analysis provided for phenomena associated with this issue, I will then turn to some final considerations concerning points of convergence between recent syntactocentrism and Cognitive Linguistics – certainly the more complex case to explore.

5. Derivational Syntactocentrism and the Parallel Architecture: Approaches to the Pragmatics of LP-Movement in German

We ended the last chapter with the general perspective of bridging the gulf between syntactocentrism and its theoretical alternatives. Specifically, I argued that it seems worth evaluating the relation between these different theoretical views in light of the amended notion of syntactocentrism developed in chapters 2 and 3. Since Jackendoff's Parallel Architecture, as we argued, can be regarded as an intermediate position between syntactocentrism and other alternatives like Cognitive Linguistics, his approach seems to be a good starting point to explore the relation between recent syntactocentrism and its theoretical alternatives. As we saw in section 4.2, one argument that Jackendoff gives in favor of the Parallel Architecture concerns the information structure of a sentence. In this chapter now, I argue that by adopting the recent derivational view of grammar within syntactocentrism, some approaches to information structural properties of a sentence within mainstream generative grammar show significant points of convergence with the conceptual underpinnings of accounting for this phenomenon within the framework of the Parallel Architecture.

To explore this hypothesis, I will bring down the comparison of recent syntactocentrism and the Parallel Architecture to tractable size by focusing on the analysis of one specific phenomenon, namely the pragmatics of left-periphery-movement (LP-movement) in German, a topic that is described in discourse-related, information structural terms in the literature. Since, as we already saw in chapter 3, recent derivational approaches within syntactocentrism define themselves via modifying aspects of preceding models, I will first illustrate a prominent representational account – the 'cartographic approach' – of LP-movement in German. After I will have outlined the general representational framework and after I will have sketched a specific analysis of LP-movement in German that is committed to this framework, I will point out essential problems this approach faces. With this in mind, in section 5.2, I will outline a recent derivational analysis of LP-movement in German. Based on this illustration, I will finally show that recent derivational approaches within syntactocentrism share crucial assumptions regarding the place of information structure in the grammar with frameworks like the 'Parallel Architecture.' Before we begin to argue towards this conclusion, let me first briefly recall the phenomenon that Jackendoff

cites in favor of the Parallel Architecture, as we saw in section 4.2, and that I will be concerned with in this chapter.

Since the extensive literature on information structure has established several categories such as focus, presupposition, topic, comment, theme, rheme, etc., thus demonstrating that "[t]here is no consensus on what and how many categories of information structure should be distinguished, or how these can be identified" (Büring 2007: 445), I will limit myself, as a first approximation, to Jackendoff's perspective on this issue. Although we have already presented examples in section 4.2, let us consider a more illuminating illustration of focus phenomena, in which Jackendoff uses question-answer pairs (cf. Jackendoff 2002: 408-409). In the following examples from English, prosody plays a central role in the expression of information structure (as it does in German, as we will see in the following sections). Again, like in the examples in section 4.2, words bearing a focal stress are marked with capitals:

(1)   J: *Who went to the party?*

      K: a. *PAT went to the party.*

           b. **Pat went to the PARTY.*

(2)   J: *Where did Pat go?*

      K: a. *Pat went to the PARTY.*

           b. **PAT went to the party.*

In these examples, one part of K's answer corresponds to the wh-element in J's question. Crucially now, only this corresponding piece can be stressed in the answers given in (1) and (2). In particular, the question in (1) can be answered with (1a) and not with (1b). Likewise, (2a) works fine, given the question *Where did Pat go?*, whereas (2b) is odd in this context. Referring to cases like (1) and (2), Jackendoff argues, as I have already mentioned, that "there are aspects of semantics that have no impact on syntax but *do* have an effect on phonology" (Jackendoff 2003: 658, emphasis in the original). Accordingly, following Jackendoff's argument, these cases seem to require a direct phonology-semantics interface, postulated in the Parallel Architecture but denied within syntactocentric approaches, according to

which phonology and semantics are mediated by the syntactic component. Concerning information structural properties of sentences, this mediation is traditionally accounted for by encoding properties of information structure in the syntactic representation – an approach that we now turn to.

5.1 The Representational View: The Cartographic Approach

Within generative grammar, Jackendoff was one of the first scholars who was concerned with issues raised by examples like the ones cited above. In particular, Jackendoff (1972: 230) uses the term 'focus of a sentence' to denote newly-supplied information, that is, "the information in the sentence that is assumed by the speaker not to be shared by him and the hearer." In contrast to that, the remaining part of the sentence is referred to as the 'presupposition of the sentence,' thus as "information in the sentence that is assumed by the speaker to be shared by him and the hearer" (Jackendoff 1972: 230). Jackendoff was also the first to suggest an encoding of notions like focus in the syntactic representation. Specifically, he introduced "a syntactic marker F which can be associated with any node in the surface structure" (Jackendoff 1972: 240). So, in case of (1a), repeated here for convenience as (3), the syntactic surface representation can be notated as follows (cf. Jackendoff 1972: 273):

(3)   *PAT went to the party.*

(4)   [$_S$ [$_{NP \atop F}$ *PAT* [$_{VP}$ *went* [$_{PP}$ *to* [$_{NP}$ *the party*]]]]]

As shown in (4), the syntactic representation contains a formal device to trigger both the prosodic aspects of focus and the focus interpretation of the relevant constituent in the semantics of the sentence. However, as cases like (3), expressing focus through prosody alone, demonstrate best, the syntactic feature [F], assumed within several syntactocentric models, merely provides a device to pass information from semantics to phonology. Accordingly, Jackendoff (2002: 409) concludes "that a syntactic feature [+F] is simply an artifact of syntactocentrism, the assumption that everything in meaning has to be derived from something generated in syntax."

However, while Jackendoff has replaced such 'mediating' devices with the postulation of a direct phonology-semantics interface within his Parallel Architecture, many approaches within mainstream generative linguistics are still committed to the general representational view that notions of information structure have to be encoded somehow in the syntactic representation. Moreover, many theoretical accounts do not merely encode information structural notions in terms of diacritics, which are assigned to the output, that is, to the surface representation at the end of the syntactic derivation (for a recent approach following this general idea of Jackendoff's early proposal, see, e.g., Zubizarreta 1998). Rather, these accounts consider information structural properties to be encoded in the syntax as formal categories present and actively determining the syntactic derivation. This branch of mainstream generative linguistics aims at representing every conceivable interpretive aspect in terms of syntactic structures, or, to put it more metaphorically, it attempts "to draw maps as precise and detailed as possible of syntactic configurations" (Cinque & Rizzi 2010: 51) – thus the name 'cartographic approach.' To capture the logic of this approach, let us study its foundations.

One important step towards the cartographic approach was the extension of X-bar theory to so-called 'non-lexical elements' of the clause, an idea that was first suggested by Chomsky (1986a). In the mid-1980s he focuses on the distinction between lexical categories, like noun, verb, adjective, and preposition, and non-lexical (also termed 'functional') categories, like complementizer (C) and inflection (I). Remember from section 2.1 that, within X-bar theory, the symbols NP, VP, AP, and PP are used to denote maximal projections of lexical categories. Now, Chomsky argues that the clausal categories S, denoting matrix clauses, and S', denoting subordinate clauses that are introduced by a complementizer, can be replaced by IP and CP, which can be represented as follows (cf. Chomsky 1986a: 3):

(5)   a.   S = IP = [Spec [$_{I'}$ [$_{VP}$ V …]]]

      b.   S' = CP = [Spec [$_{C'}$ C IP]]

Accordingly, when we adopt the suggestion "to extend conventional notation for the lexical categories to the nonlexical ones" (Chomsky 1986a: 3) and presuppose the

general X-bar schema for lexical categories sketched in section 2.1, the clausal structure can be roughly represented as follows:

(6)

```
                CP
              /    \
          Spec      C'
                   /   \
                  C     IP
                      /    \
                  Spec      I'
                          /    \
                         I      VP
```
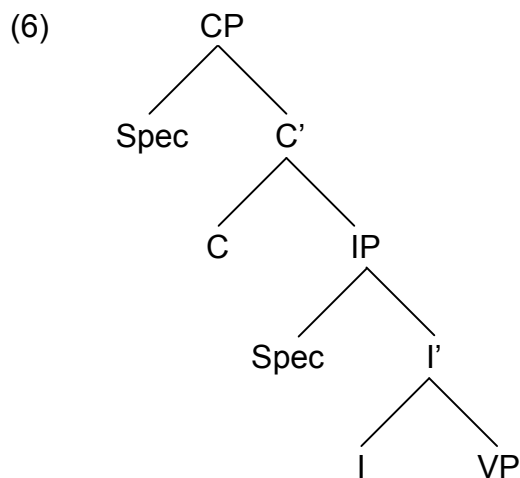
Crucially now, once the 'IP-zone' and the 'CP-zone' were isolated, the cartographic idea of drawing more precise 'maps' of IP and CP arose, and it was thus argued to split the functional projections IP and CP into more elementary elements. The first proposal to enrich the representational devices of the X-bar schema in this way concerned the splitting of IP, as Pollock (1989) suggested in his discussion of verb movement in French and English.

Pollock (1989) claims that a single head within IP cannot account for the different positions that infinitival and finite verbs can occupy in French. To capture Pollock's proposal, recall from section 3.2.1 that he regards the structure in (7), in this case supplemented with the optional negative adverb *pas*, as the D-Structure, hence as the starting point of the derivation (cf. Pollock 1989: 366):

(7) $[_{IP}$ NP I (*pas*) $[_{VP}$ (Adv) V …]]

With this assumption in mind, let us now examine the following examples (cf. Pollock 1989: 377-379), containing both a negative and a VP-initial adverb:

(8)   a.   *Il ne perd pas complètement la tête pour les belles étudiantes, c'est bien.*

      b.   **Il ne pas complètement perd la tête pour les belles étudiantes, c'est bien.*

According to (8), the movement of the finite verb (*perd*) to I, as in (8a), is obligatory in French. If the finite verb remains in situ, as in (8b), the clause is ungrammatical. Now, with this rule in mind, look at the following examples:

(9) a. *Ne pas complètement perdre la tête pour les belles étudiantes, c'est bien.*

   b. **Ne perdre pas complètement la tête pour les belles étudiantes, c'est bien.*

   c. *Ne pas perdre complètement la tête pour les belles étudiantes, c'est bien.*

In the case of infinitives, unlike in cases involving finite verbs, the verb (*perdre*) may remain in situ, as shown in (9a). If the nonfinite verb *perdre* moves to I across the negative *pas*, the clause becomes ungrammatical, as (9b) illustrates. However, the infinitive can be moved to a higher position across VP-initial adverbs like *complètement*, as (9c) demonstrates. Since this optional movement is possible, the movement operation involved in (9c), as Pollock argues, must target some intermediate position between the negative adverb (*pas*) and the VP-initial adverb (*complètement*). So, based on observations like this, Pollock claims that the category I must be splitted in order to have both a position in front of *pas* and a position between *pas* and the VP-initial adverb. In particular, he hypothesized that the 'short' movement of the infinitive verb illustrated in (9c) targets the head 'Agreement' (Agr) and that the 'long' movement of finite verbs, as illustrated in (8a), targets the head 'Tense' (T). When we adopt this splitting of the functional projection IP, the resulting structure of the clause can be represented with the following schema, in which a Negation Phrase (NegP) is added to account for the negative adverb *pas* (cf. Pollock 1989: 397):

(10)

```
                    CP
                   /  \
               Spec    C'
                      /  \
                     C    TP
                         /  \
                     Spec    T'
                            /  \
                           T    NegP
                               /   \
                           Spec     Neg'
                                   /    \
                                 Neg     AgrP
                                        /    \
                                    Spec      Agr'
                                             /    \
                                           Agr     VP
```
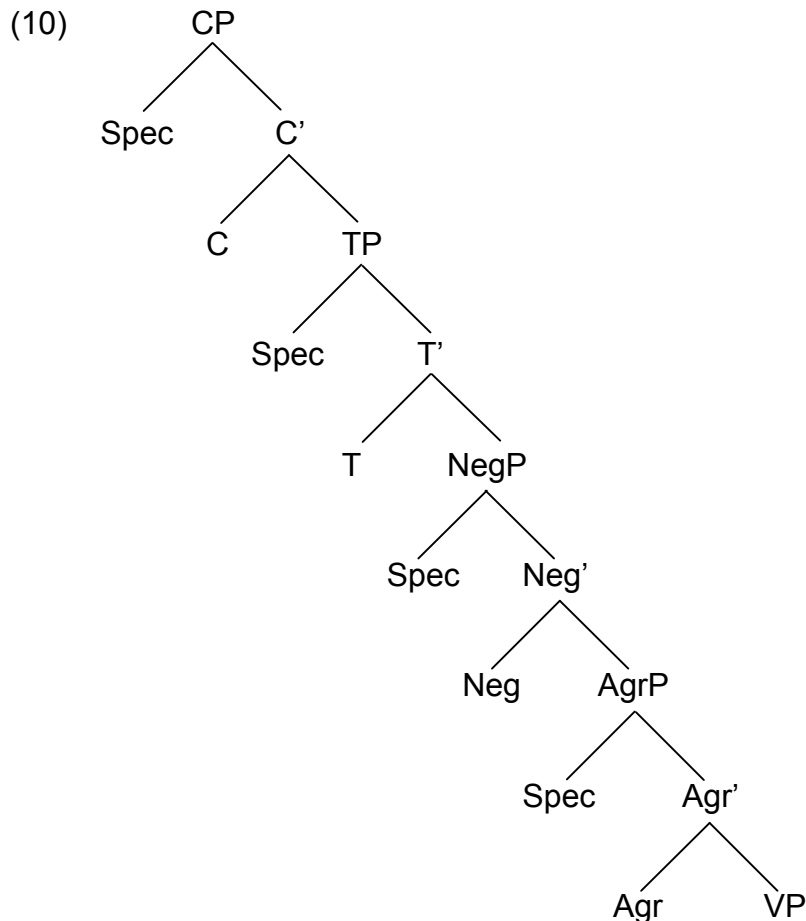
Knowing this approach of dissolving IP into more elementary functional projections, let us now turn to first attempts to split the CP-domain, which introduces the general idea of postulating dedicated categories for information structural notions in the syntactic representation.

In his extensive discussion of the CP-domain, Rizzi (1997) suggests that a single head within this domain (viz. C) is not enough to account for the left periphery of the clause, given the different interpretive imports hosted in this pre-IP-zone. The first interpretive import Rizzi discusses is the widely-assumed aspect that "C is basically an indicator of mood or force […]: declarative, interrogative, and so on" (Chomsky 1995: 240) and thus determines clause types. To make this 'specification of force' more concrete, consider the following examples, in which the choice of the complementizer *whether* results in an interrogative and the choice of *that* in a declarative:

(11)   a.   (*I wonder*) *whether Hans drank beer.*

      b.   (*I know*) *that Hans drank beer.*

In addition to this 'specification of force,' the second kind of informational contribution of the C-domain concerns the fact that a complementizer is connected to certain properties of the verbal system. For example, since "COMP may be realized as *that* before tensed sentences and as *for* before infinitives" (Chomsky & Lasnik 1977: 434), the choice of the complementizer co-occurs (in English) with the choice of a tensed or an infinitival verb. Yet, based on Italian data, Rizzi (1997: 283) argues that "the 'temporal' properties encoded by C are very rudimentary," for a complementizer can co-occur with many tenses in Italian. Therefore, he claims that "C expresses a distinction related to tense but more rudimentary than tense and other inflectional specifications on the verbal system: finiteness" (Rizzi 1997: 284). When we adopt this assumption that the C-system merely expresses an abstract specification of 'finiteness' (Fin) and put this together with the property of specifying the 'force' of a sentence, the structural representation of CP, as a first step, can be dissolved into a force-finiteness system, as illustrated in the following schema:

(12)   ForceP
          /   \
       Spec   Force'
                /   \
            Force   FinP
                       /   \
                    Spec   Fin'
                            /   \
                          Fin   IP
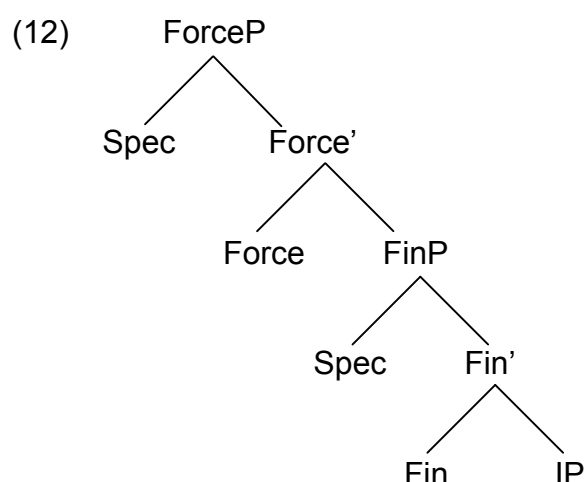
Crucially now, as Rizzi (1997) argues, also information structural notions are encoded in the C-domain. Look at the following examples (cf. Rizzi 1997: 286):

(13)   [*Il   tuo   libro*]$_i$ *lo ho letto*   t$_i$
         the   your   book   it   I   read
        'As for your book, I read it.'[13]

---

[13] Throughout the thesis, I will provide glosses and translations of examples taken from languages other than English and French, as is common in most of the literature I rely on. In

(14) [*IL TUO LIBRO*]_i *ho letto*   t_i (*non il   suo*)

        the your book   I   read   (not the his)

        'I read your book.'[14]

In both cases, the phrase *il tuo libro* is preposed, that is, has moved to the left pe-riphery of the clause, thereby receiving some discourse-relevant interpretation. In particular, in (13), *il tuo libro* expresses the topic, that is, roughly speaking, informa-tion available and salient in previous discourse. In this case, the rest of the sentence expresses new information predicated of the topic (the 'comment'). In (14), unlike in the case of (13), *il tuo libro* bears focal stress and expresses new (contrastive) in-formation, indicated by the possible Italian expansion *non il suo*. So, since (14) im-plies that the hearer believes that the speaker has read something different from the hearer's book, the speaker corrects the hearer's believe by uttering (14). While *il tuo libro* is the (contrastive) focus, the rest of the sentence expresses given information, presupposed by the speaker to be shared with the hearer.[15]

So, in sum, the topic-comment and the focus-presupposition articulation can be regarded as two different interpretive options with respect to the left periphery of the Italian clause. Accordingly, Rizzi (1997: 286-287) extends the C-domain by postulat-ing two more functional projections, as shown in (15) and (16):

(15)

TopP
├── Spec
└── Top'
    ├── Top
    └── Complement

---
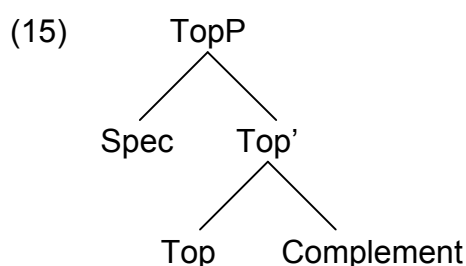
providing the glosses, I will abstract away from grammatical markers like Case or Tense, as they are not relevant in the contexts wherein the given examples are provided.

[14] For the sake of consistency, I will not translate the specific information structural proper-ties of examples that involve a contrastive or focal interpretation into English by means of cleft constructions or the like. Since such translations are impossible in case of some ex-amples provided in later parts of this thesis and since these interpretive aspects are also omitted in the translations provided by the literature I will be concerned with, I will limit my-self to bring out these special interpretive properties by discussing them in the text.

[15] Since it makes no difference for the point to be made in this section, I abstracted away from the fact that the two structures discussed are not only formally distinguished concern-ing their prosodic properties but also with respect to the clitic *lo* (for elaboration on this as-pect, see Rizzi 1997: 285-286 and references therein).

(16)   FocP
             /   \
         Spec   Foc'
                  /   \
               Foc   Complement
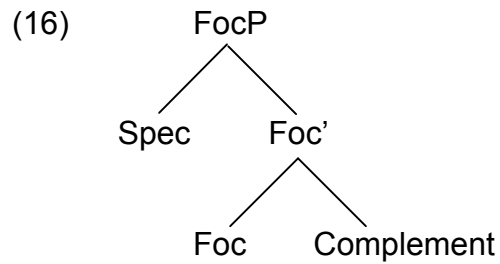
According to Rizzi (1997), the head 'Top' takes the topic as its specifier and the comment as its complement. Likewise, the specifier of the schema given in (16) hosts the focal element, while the complement of the head 'Foc' is the presupposition. So, in addition to the force-finiteness system sketched above, Rizzi argues for a topic-focus field in the left periphery of the clause. Consequently, the articulated structure of the complementizer system proposed by Rizzi can be represented as in the following schema, in which '*' indicates that the node is recursive (cf. Rizzi 1997: 297):

(17)

```
                    ForceP
                   /      \
                Spec      Force'
                         /     \
                      Force    TopP*
                              /     \
                           Spec     Top'
                                   /    \
                                 Top    FocP
                                       /    \
                                    Spec    Foc'
                                           /    \
                                         Foc    TopP*
                                               /     \
                                            Spec     Top'
                                                    /    \
                                                  Top    FinP
                                                        /    \
                                                     Spec    Fin'
                                                            /    \
                                                          Fin    IP
```
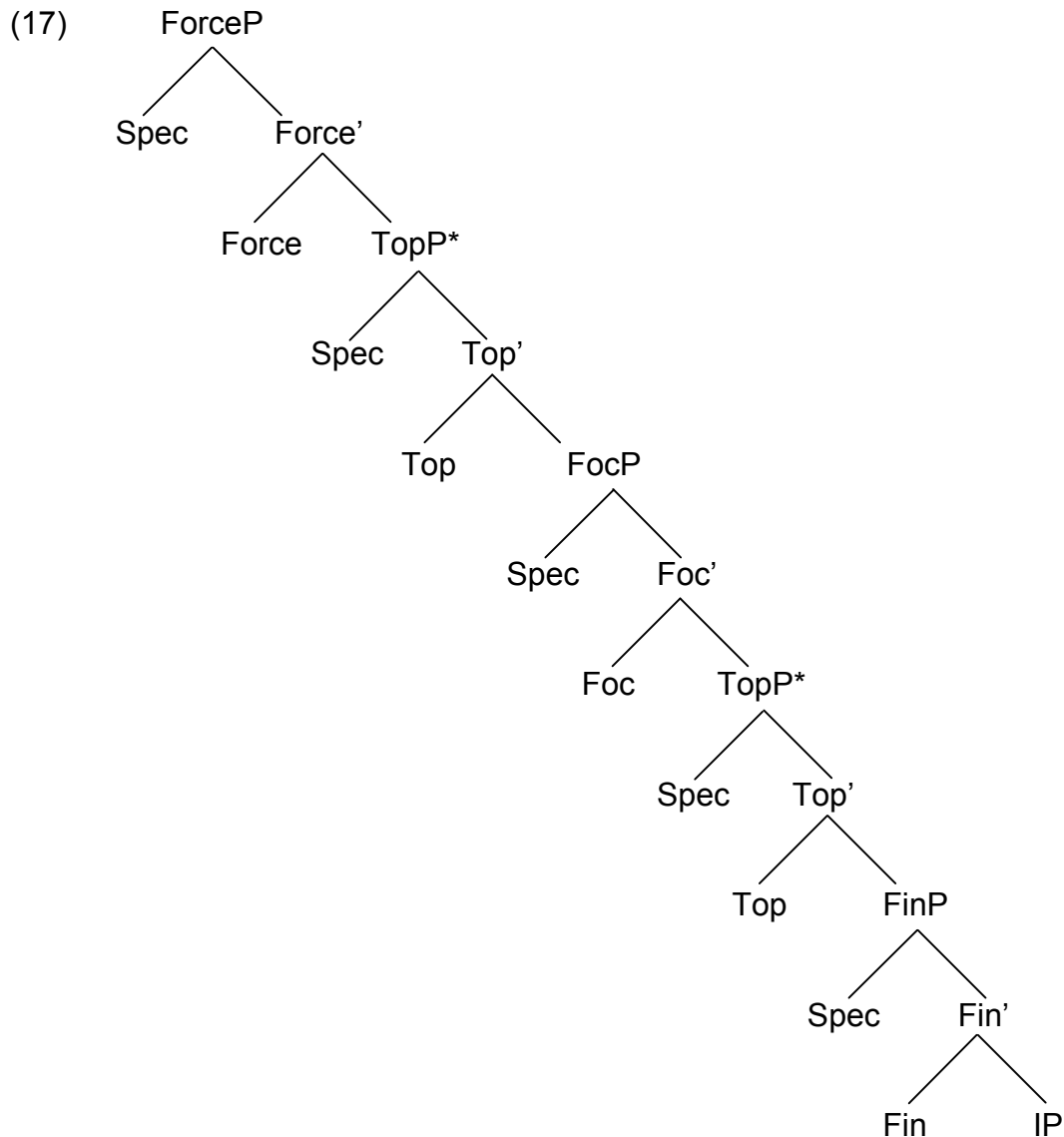
Rizzi (1997) provides many arguments for why a clause can contain several topics – as marked by '*' – and only one focal element, for why the focalized element can be both preceded and followed by topics, and he also provides evidence for situating the topic-focus field between the force-finiteness system. However, a detailed exposition of his arguments, based on a large amount of data from Italian, would take me too far afield. My aim was merely to illustrate the general idea of drawing 'structural maps' by splitting the functional projections of the clause, as initiated by Pollock's (1989) proposal concerning IP and Rizzi's (1997) approach to the CP-domain. With this general conception in mind, let us now focus on a cartographic approach to information structural properties of the German left periphery.

5.1.1 A Cartographic Analysis of LP-Movement in German

As we saw in the previous section, the cartographic approach to syntactic structures attempts to draw detailed 'maps' of syntactic configurations. In doing so, proponents of this approach have shown how functional projections can be dissolved into more precise representations. The initial impulse for splitting functional projections came from closer investigations of IP. However, it was argued soon that a splitting of CP into more elementary elements should be put forward, given the different interpretive imports that are hosted in the left periphery. Having introduced the general conception of this approach and how it is applied to the left periphery of the clause, let us now examine a cartographic analysis of pragmatic properties of LP-movement in German.

As for German, it is a long-established view within generative linguistics that in light of early investigated "fronting processes as […] topicalization" (Thiersch 1978: 85), the left periphery of the clause structure hosts pragmatically relevant elements that are associated with information structural notions. In the following, I will focus on one particular approach that adopts the cartographic view to analyze the pragmatics of the German left periphery. In particular, Frey (2004a et seq.) argues for an additional functional projection in the CP-domain that accounts for the contrastive interpretation some elements are argued to receive, when they are moved to the left periphery. To capture his argument, we must first study the different options of movement to the German left periphery that he proposes.

Frey (2004a) claims that there are basically two types of movement for filling the left periphery in German: 'Formal Movement' and a movement operation resulting in the contrastive interpretation of the moved element. Let us first look at the operation of Formal Movement. The following example, taken from Frey (2004a: 3), may help to bring out the crucial aspects of this movement type:

(18) *In Europa spielen Jungen gerne Fußball.*
      in Europe play   boys   Adv.   football
      'In Europe, boys like to play football.'

In (18), the constituent *in Europa* is regarded as unmarked, that is, it is neither interpreted contrastively nor receives prominent stress. Now, Frey (2004a) argues that

fronting the adverbial in a sentence such as (18) is obtained by Formal Movement. The movement *in Europa* has undergone is viewed as purely formal because it only took place to meet the formal condition of filling a specific position in the left periphery that has to be occupied in a German verb-second clause. Note that in most cases of German verb-second clauses this requirement is fulfilled by moving the subject of the clause to this position, which is the specifier of CP, as indicated in (19), where the example in (18) is modified by leaving out the adverbial *in Europa* and (19a) represents the starting point of the derivation:
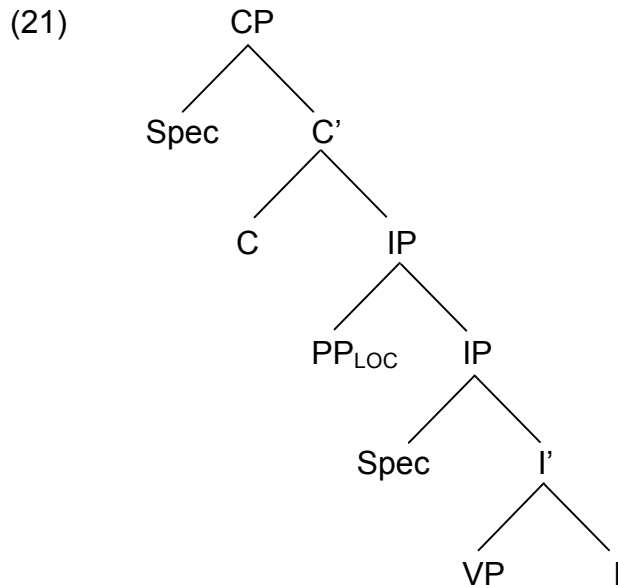
(19) a. [$_{CP}$ [$_C$ (*dass* ('that')) [$_{IP}$ *die Jungen* [$_{VP}$ *gerne Fußball spielen*]]]]

      b. [$_{CP}$ *die Jungen*$_i$ [$_C$ *spielen*$_j$ [$_{IP}$ t$_i$ [$_{VP}$ *gerne Fußball* t$_j$]]]]

It was Bhatt (1999) who first pointed out that in verb-second clauses not only subjects but also other (non-subject) elements such as "adverbs can appear clause-initially without any focal stress associated with them" (Bhatt 1999: 89). According to Bhatt, this is due to the so-called 'Minimal Link Condition.' That is, in order to meet the formal condition of filling the specifier of CP, "the derivation will choose the […] candidate closest to the target" (Bhatt 1999: 89). In other words, the highest constituent of the IP-zone is chosen to move to the specifier of CP. Adopting this proposal for German, let us again concentrate on our example (18). As Maienborn (1996: 111-118) has observed, in contrast to local adverbials referring to the inner proposition of the clause, 'frame adverbials' like *in Europa* are setting the 'frame' for the whole proposition that the clause expresses. For the sake of contrast, consider the following example, again implicitly assuming the configurations of verb-final clauses as the starting point:

(20) [$_{CP}$ *Hans*$_i$ [$_C$ *will*$_j$ [$_{IP}$ t$_i$ [$_{VP}$ *in der Bar Bier trinken* t$_j$]]]]
          Hans   wants.to   at the bar beer drink
      'Hans wants to drink beer at the bar.'

Applying several tests with respect to scope-related properties of sentences, Maienborn argues that in contrast to local adverbials like *in der Bar* in (20), which are hosted within the VP of the clause, frame adverbials like *in Europa* in (18) are 'base-

generated' above the subject and thus external to the VP. Consequently, Maienborn (1996: 117) suggests the following schema to capture the base position for these adverbials, which are labeled as PP_{LOC}:

(21)

```
                CP
               /  \
           Spec    C'
                  /  \
                 C    IP
                     /  \
                PP_LOC   IP
                        /  \
                    Spec    I'
                           /  \
                         VP    I
```

According to the schema in (21), frame adverbials like *in Europa* are situated in the highest position of the IP-domain (in this case extended by a further IP projection). Therefore, when we adopt the 'Minimal Link Condition' mentioned above, the adverbial can move to the preverbal position because it is the closest element to the specifier of CP. So, following the proposal of Maienborn (1996), Frey (2004a) argues that the movement of *in Europa* to the left periphery can be regarded as Formal Movement because in this case, the adverbial is the highest constituent within the IP-domain and can thus be moved to meet the condition of filling the specifier of CP. This can be formally represented as follows (cf. Frey 2004a: 9):

(22) $[_{CP}$ *in Europa*$_i$ $[_C$ *spielen*$_j$ $[_{IP}$ t$_i$ *Jungen* $[_{VP}$ *gerne Fußball* t$_j$]]]]

From the discussion so far, it should become clear that, according to this proposal, constituents like subjects or frame adverbials can show up in the left periphery of the German clause without receiving any special pragmatic interpretation that could be accounted for in information structural terms. However, as Frey (2004a: 12-15) points out, the left periphery can also host topical material, as shown in (23):

(23)   I tell you something about Max.

   Den   Max   sollte   unsere   Gruppe   unterstützen.

   the   Max   should   our   group   support

   'As for Max, our group should support him.'

As Frey argues, topical material, before moved to the left periphery, has already undergone a movement operation. Specifically, in light of subordinate sentences like (24a), in which *den Max* receives a topical interpretation, Frey (2004a) supposes a dedicated position for topics between the complementizer and the whole IP, as (24b) spells out in detail (for an elaboration on this topic position, see Frey 2004b):

(24)   a.   … *dass* ('that') *den Max unsere Gruppe unterstützen sollte.*

   b.   [CP [C (*dass*) [TopicP *den Max*i [IP *unsere Gruppe* [VP ti *unterstützen sollte*]]]]]

According to Frey, the constituent *den Max* has already received its topical interpretation by occupying the dedicated position indicated in (24b), and this position, again, is the closest one to the specifier of CP, *den Max*, and, under this hypothesis, is also moved to the left periphery by Formal Movement, as represented in (25):

(25)   [CP *den Max*i [C *sollte*j [TopicP ti [IP *unsere Gruppe* [VP ti *unterstützen* tj]]]]]

To summarize, when we hypothesize that the movement of topical material to the left periphery "does not seem to be related to any semantic or pragmatic property but seems to be a purely formal one" (Frey 2004a: 8), fronting *den Max* is viewed as just another instantiation of the movement type that is applied in order to prepose subjects and frame adverbials.

   With all these options to fill the German left periphery in mind, let us now study a case of LP-movement that, according to Frey, is associated with a special pragmatic interpretation that can be captured in information structural terms. Look at the following example:

(26) *GRÜN wird Maria die Tür streichen (und nicht rot)*.

      green   will   Maria the door paint   (and not   red)

      'Maria will paint the door green.'

According to Frey, in (26), the constituent *grün* receives focal stress and is interpreted contrastively or, more specifically, as a contrastive focus. A natural German expansion, therefore, would be *und nicht rot*, as indicated in parentheses. In contrast to the elements preposed in (18) and (23), *grün*, as (27a) shows, is neither base-generated in the highest position of the IP-zone – like frame adverbials – nor can it be moved to the position that hosts topical material, as (27b) demonstrates (cf. Frey 2004a: 19):

(27) a. *[$_{CP}$ [$_{C}$ (*dass*) ('that') [$_{TopicP}$ [$_{IP}$ *grün Maria* [$_{VP}$ *die Tür streichen wird*]]]]]

     b. *[$_{CP}$ [$_{C}$ (*dass*) ('that') [$_{TopicP}$ *grün*$_i$ [$_{IP}$ *Maria* [$_{VP}$ *die Tür* t$_i$ *streichen wird*]]]]]
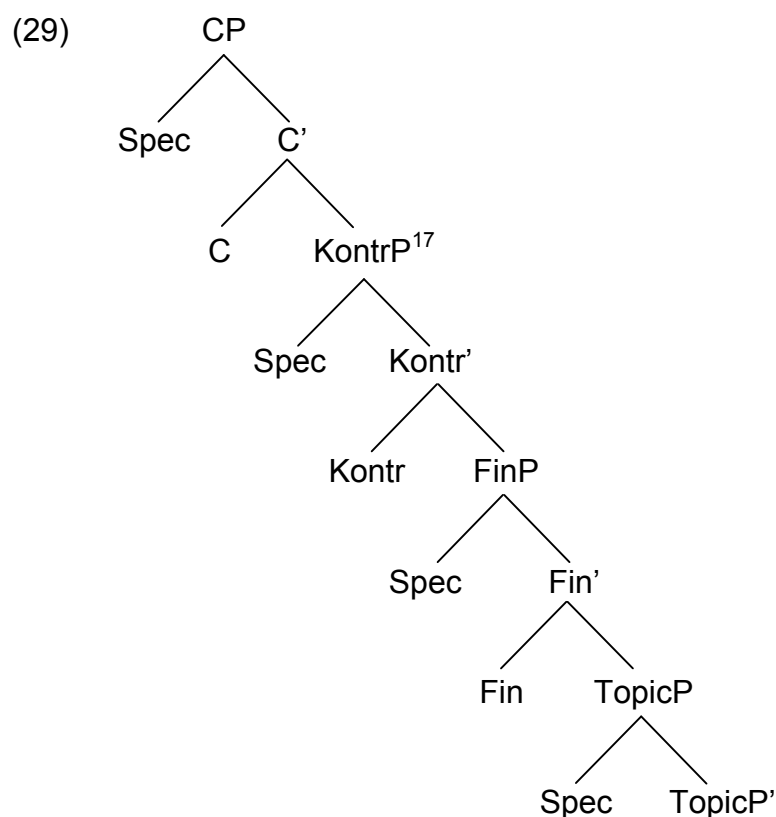
Since (26), as (27) shows, cannot result from Formal Movement, *grün* has reached its position, according to Frey, via another form of LP-movement, which can be spelled out as follows:

(28) [$_{CP}$ *grün*$_i$ [$_{C}$ *wird*$_j$ [$_{IP}$ *Maria* [$_{VP}$ *die Tür* t$_i$ *streichen* t$_j$]]]]

So, as (28) illustrates, this type of movement differs syntactically from the other ways to fill the German left periphery that were discussed above. Moreover, as Frey claims, it is the only type of LP-movement that is associated with focal stress on the moved item and with a pragmatic interpretation in terms of contrast.[16] Based on these observations, and adopting the cartographic enterprise to transparently represent all interpretive properties in the syntactic representation, Frey (2004a) argues

---

[16] In more recent work, Frey goes on to refine the interpretive import of the element preposed to the left periphery in cases like (26). In particular, he replaces the coarse notion of contrast by "a notion of emphasis which reflects a ranking with a highest element imposed by the speaker among the alternatives evoked by the stressed element" (Frey 2010: 1434). However, since the aim of this section is merely to illustrate the general concept of accounting for pragmatic properties of LP-movement in German in terms of the cartographic approach, I will not go into this recent proposal.

that the CP-zone of the German clause has to be splitted, and he proposes the fol-lowing structure for the left periphery (cf. Frey 2006: 254):

(29)

```
                    CP
                   /  \
                Spec   C'
                      /  \
                     C   KontrP^17
                        /    \
                     Spec   Kontr'
                           /    \
                        Kontr   FinP
                               /   \
                             Spec   Fin'
                                   /   \
                                 Fin   TopicP
                                      /    \
                                   Spec   TopicP'
```

Applying this structure of the German left periphery to the examples given above, we can substitute the representation in (28) with the more elaborated structure given in (30), which indicates that *grün* moves to the specifier-position of a functional projec-tion that hosts contrastive elements (KontrP):

(30)   [$_{KontrP}$ *grün*$_i$ [$_{Kontr'}$ *wird*$_j$ [$_{FinP}$ [$_{TopicP}$ [$_{IP}$ *Maria* [$_{VP}$ *die Tür* t$_i$ *streichen* t$_j$]]]]]]

When we adopt the proposal of Rizzi (1997) that was sketched in the preceding sec-tion, the projection not associated with any pragmatic effect is called 'FinP.' Accord-ingly, both *in Europa* in (18) and *den Max* in (23), since they are not associated with any specific pragmatic interpretation, are moved, according to this approach, to the

---

[17] In this case, by using 'kontrast' in lieu of 'contrast,' Frey adopts a terminological conven-tion according to which "the idiosyncratic spelling indicates that the term is not to be un-derstood as covering all instances of what has been dubbed contrast in semantics, syntax, and phonology" (Vallduví & Vilkuna 1998: 81).

specifier of FinP by Formal Movement. However, whereas frame adverbials like *in Europa* are adjoined to IP, the topic *den Max*, as we illustrated above, first moves to TopicP before it is fronted to the specifier of FinP. This difference can be formally represented as follows:

(31) [$_{KontrP}$ [$_{FinP}$ *in Europa*$_i$ [$_{Fin'}$ *spielen*$_j$ [$_{TopicP}$ [$_{IP}$ t$_i$ *Jungen* [$_{VP}$ *gerne Fußball* t$_j$]]]]]]

(32) [$_{KontrP}$ [$_{FinP}$ *den Max*$_i$ [$_{Fin'}$ *sollte*$_j$ [$_{TopicP}$ t$_i$ [$_{IP}$ *unsere Gruppe* [$_{VP}$ t$_i$ *unterstützen* t$_j$]]]]]]

Let us summarize. In this section, I sketched a cartographic analysis of the left periphery of German. Needless to say, a section is not enough to touch on all of the detailed technical aspects of the issues involved. However, I extracted the basic syntactic arguments, perhaps sometimes oversimplifying, to illustrate how a specific phenomenon associated with the pragmatics of information structure is accounted for according to the representational view of the cartographic approach. With this strong version of representational syntactocentrism in mind, which aims at transparently encoding all interpretive properties of a sentence in the syntactic representation, let us now concentrate on some problems that arise with respect to the specific analysis I illustrated.

5.1.2 Conceptual and Empirical Problems

As we saw, according to the strong representational view of the cartographic approach, syntactic structures have to provide a transparent representation of interpretive properties, including, as was shown in the context of LP-movement in German, pragmatic properties associated with information structural notions like focus or contrast. In this section, I will briefly point out some conceptual and empirical problems this approach faces. Let us first focus on conceptual issues.

According to the proponents of the cartographic approach, there is no tension between the enriched representations illustrated in the preceding sections and the minimalist idea of reducing the syntactic computation to a minimum (see section 2.1) and of dispensing with levels of representation considered to be not conceptually necessary (see section 3.2). In particular, as Rizzi (2004) argues, a crucial point of

connection between the cartographic approach and the Minimalist Program is the core idea of simplicity. The cartographic approach, according to Rizzi, contributes to this notion by decomposing functional projections into simple structural units. According to this account, one unit is associated with exactly one interpretive property, thus transparently indicating interpretation at the interfaces. Thus, in the eyes of linguists working within the cartographic framework, "[l]ocal simplicity is preserved by natural languages at the price of accepting a higher global complexity, through the proliferation of structural units" (Rizzi 2004: 8). And yet, from the view of the Minimalist Program, which is committed to a strong derivational concept, there is an essential problem with an approach that supposes such enriched syntactic representations.

Recall from section 3.2.1 that movement operations, adopting the economy conditions of the Minimalist Program, must always be motivated somehow. In particular, in the context of 'strong' and 'weak' Agreement-features, we discussed the conception that the process of obligatory 'feature checking' justifies movement operations within the Minimalist Program. According to this proposal, proponents of the cartographic approach are forced to postulate that a functional head in the left periphery possesses a [foc] or, in our case, a [kontr] feature, with which the preposed phrase matches. According to Chomsky (1995: 228), however, "any structure formed by the computation […] is constituted of elements already present in the lexical items selected […]; no new objects are added in the course of computation apart from rearrangements of lexical properties." In other words, this 'Inclusiveness Condition' implies that syntactic operations can only refer to lexical features. Of course, lexical items cannot be viewed as inherently focused or contrastive. Consequently, such features, as Neeleman & Szendrői (2004: 155) note, "must be inserted after an element has been taken from the lexicon," and thus the postulation of such features violates the minimalist 'Inclusiveness Condition.' In addition to this general conceptual problem of presupposing discourse-related features in the syntactic representation, let us now look at an empirical phenomenon of German that challenges the specific cartographic analysis I presented above.

In his discussion of the two options of LP-movement in German – Formal Movement and movement that yields some pragmatically relevant interpretation – Fanselow (2003) demonstrates some cases that do not fit the cartographic approach. The first observation he builds on is the fact that in German, given certain pragmatic

conditions, verb prefixes can be moved to the left periphery, as the following examples illustrate (cf. Fanselow 2003: 32):

(33) a. auf-machen ('open,' lit. 'open-make')

[*AUF*]$_i$ *hat er die Tür* t$_i$ *gemacht* (*und nicht zu*).

open has he the door made (and not shut)

'He has opened the door.'

b. vor-haben ('intend,' lit. 'before-make')

[*VOR*]$_i$ *haben wir das schon* t$_i$ *gehabt.*

before have we that well had

'We had intended that.'

In both cases, the verb prefix – *auf* in (33a) and *vor* in (33b) – is moved to the left periphery. However, as Fanselow argues, there is a crucial difference between these two examples. Examine first the case of moving *auf*. Since the prefix-verb *aufmachen* has, according to Fanselow, a compositional interpretation, there is, in principle, no problem with a cartographic approach to the movement of *auf*. In particular, demonstrating that the prefix *auf* is not preposed to the left periphery by Formal Movement, Fanselow points out that *auf* receives focal stress and is contrastively interpreted, as the possible expansion *und nicht zu* indicates. Examine now (33b). Here, unlike in the case of (33a), the interpretation of the prefix-verb *vormachen* is regarded as a non-compositional one. In other words, the prefix *vor*, as Fanselow claims, is itself meaningless and makes no identifiable semantic contribution to the particle-verb combination. So, while the case of *auf* can be accounted for by a cartographic approach like Frey's, data such as (33b) do not fit this account because, according to Fanselow (2003: 33), "[i]t is difficult to accept the idea that a meaningless element can be interpreted as a focus or a topic […] phrase." Nevertheless, the fronting of *vor* has a clear interpretive import because it results in the contrastive interpretation of the whole predicate. Accordingly, as for pragmatic interpretation, preposing *vor* is equivalent to fronting the whole prefix-verb, as in (34):

(34) [*VORgehabt*]ᵢ *haben wir das schon* tᵢ*.*
      before.had   have   we   that well
      'We had intended that.'

Given this interpretive equivalence between moving a part and fronting the whole constituent to the left periphery, Fanselow introduces the term *"pars-pro-toto-movement"* (Fanselow 2003: 32, emphasis in the original) to refer to syntactic constructions in which only a part of the phrase receiving a pragmatic interpretation is fronted. Based on these initial observations concerning prefix-verbs, Fanselow goes on to show that parts of idiomatic verb phrases can be moved to the left periphery, too. Look at the following example (cf. Fanselow 2004: 20):

(35) schöne Augen machen ('to make eyes,' lit. 'beautiful eyes make')
      [*Schöne AUgen*]ᵢ *hat er ihr* tᵢ *gemacht.*
      beautiful eyes   has he her   made
      'He made eyes at her.'

Like in the cases of (33a) and (33b), as Fanselow (2004) argues, *schöne Augen* cannot be moved to the left periphery by Formal Movement and thus receives focal stress and is associated with a contrastive interpretation. However, as in the case of *vor* in (33b), the preposed part *schöne Augen* of the idiom *schöne Augen machen* is regarded as meaningless in isolation. Accordingly, fronting this element poses a problem concerning the assumption of a dedicated syntactic position in the left periphery associated with focal or contrastive interpretation of the element that occupies this position. Again, as was the case with (33b), *schöne Augen* is moved to the left periphery 'pars-pro-toto,' since the pragmatic interpretation involved is equivalent to preposing the whole constituent, as shown in (36):

(36) [*Schöne AUgen gemacht*]ᵢ *hat er ihr* tᵢ*.*
      beautiful eyes   made   has he her
      'He made eyes at her.'

As a final point, let me add that Fanselow (2003: 34) argues that this type of movement also applies in non-idiomatic constructions such as (37):

(37)   Was he an anarchist?

[*HÄUser*]ᵢ *hat   er   jedenfalls   nie*   tᵢ   *angezündet.*

houses   has   he   in any event   never   set on fire

'He has never set houses on fire.'

In (37), the element contrastively interpreted is not the fronted noun *Häuser* but the syntactic elements referring to the action one would expect from someone who is an anarchist (e.g. setting houses on fire). Hence, the whole predicate is contrastively interpreted, as the variant in (38) may help to bring out:

(38)   Was he an anarchist?

[*HÄUser   angezündet*]ᵢ   *hat   er   jedenfalls   nie*   tᵢ.

houses   set on fire   has   he   in any event   never

'He has never set houses on fire'

Since the case of 'pars-pro-toto-movement,' as the preceding non-idiomatic example shows best, is thus not an 'exotic,' marginal phenomenon, this empirical observation concerning LP-movement in German poses serious problems for cartographic accounts that suppose a specific position in the left periphery – in Frey's approach, the specifier of 'KontrP' – that is associated with a contrastive interpretation of the element that occupies this position, and only of this element. Together with the conceptual objections mentioned at the beginning of this section, this calls, in the eyes of linguists committed to a more derivational view of syntax, for analyzing the data differently. Let us see how this can be done.

5.2 The Derivational View: Cyclic Linearization in Minimalism

As I illustrated in section 3.2, recent syntactocentrism is characterized by a strong derivational view of grammar. In particular, we arrived at a picture in which the computation of convergence is regarded as proceeding 'phase by phase' and thus does not postulate fully-fledged levels of representation like LF and PF. Having shown conceptual and empirical problems of analyzing German LP-movement within a representational framework like the cartographic approach in the last section, we will now focus on an analysis of LP-movement in German that adopts the phase-based approach within recent syntactocentrism, since it assumes the concept of 'cyclic linearization.' Based on the illustration of this approach, I will then reevaluate Jackendoff's point of abandoning syntactocentrism in light of information structural properties of sentences. Specifically, I will show that his arguments concerning this matter do not hold given analyses committed to a strong derivational view and, what is more, that there are in fact significant points of convergence between the conceptual underpinnings of his Parallel Architecture and recent derivational syntactocentrism. However, before I will begin to argue in this direction, let me first briefly clarify the concept of 'cyclic linearization.'

The explicit formulation of this concept goes back to Fox & Pesetsky (2005), who explore the mapping between syntax and phonology in light of recent phase-based approaches to grammar. In particular, they propose, in accordance with the derivational model illustrated in section 3.2.2, that the mapping from syntax to phonology does not take place at a single point in the course of the derivation but rather at various points that correspond to different cycles. Based on this cyclic view on Spell-Out, Fox & Pesetsky focus on one crucial property of the mapping between syntax and phonology, namely on the process concerned with the linear ordering of words, in short: linearization. Their basic claim concerning this process is that linearization implies the property of 'Order Preservation,' meaning that "*information about linearization, once established at the end of a given Spell-out domain, is never deleted in the course of the derivation*" (Fox & Pesetsky 2005: 6, emphasis in the original). To concretize, let me exemplify this abstract concept with some aspects of one phenomenon they discuss in their paper.

Among other data, they test their proposal with respect to so-called 'Object Shift' in Scandinavian languages. In Scandinavian languages such as Swedish, like in the

cases of German we sketched in the previous sections, the finite verb can appear in the second position of the clause by moving to the complementizer position 'C.' In embedded clauses, however, the main verb does not move to C. What is crucial in the present context is that, when the verb moves to C, an object can move to a position that precedes the VP, that is, to a position left of adverbs. To gain more concreteness, consider the following examples (cf. Fox & Pesetsky 2005: 17):

(39) a. $[_{CP}$ *jag*$_i$ $[_C$ *kysste*$_j$ $[_{IP}$ t$_i$ *henne*$_k$ *inte* $[_{VP}$ t$_j$ t$_k$]]]]

             I   kissed   her   not

    'I did not kiss her.'

    b. *$[_{CP}$ $[_C$ (*att*) $[_{IP}$ *jag henne*$_i$ *inte* $[_{VP}$ *kysste* t$_i$]]]]

             (that)   I   her   not   kissed

    '… that I did not kiss her.'

In case (39a), the finite verb (*kysste*) moves to C and thus the movement of the object *henne* is possible. By contrast, in the case of embedded clauses such as (39b), the main verb remains in situ, and, as a consequence, the object cannot move to the position left of adverbs. Fox & Pesetsky argue that these facts follow from the property of 'Order Preservation' mentioned above. In particular, adopting a cyclic view of Spell-Out, they claim that acceptability of 'Object Shift' depends on the condition that the ordering of elements in the VP Spell-Out domain does not contradict the ordering in the CP-domain. Let us see how this account works in cases like (39a) and (39b).

    Regarding (39a), Fox & Pesetsky (2005: 17) sketch the following abstract ordering statements, where '<' means, roughly speaking, 'precedes:'[18]

---

[18] In fact, they develop a more precise formulation of the relation this symbol indicates (cf. Fox & Pesetsky 2005: 8-10).

(40)   a.   VP: [$_{VP}$ V O]

             Ordering: V < O

       b.   CP: [$_{CP}$ S$_i$ V$_j$ [$_{IP}$ t$_i$ O$_k$ Adv [$_{VP}$ t$_j$ t$_k$]]]

             Ordering: S < V

                       V < O

                       O < Adv

                       Adv < VP

According to (40), the verb (V) precedes the object (O) within the Spell-Out domain –
the 'phase' – VP, as indicated in (40a). Since V precedes O, when CP is spelled out
as well, because V has moved to C, there is no contradiction between the ordering
statements given in (40a) and (40b). With this in mind, study the ordering statements
that hold true for cases like (39b):

(41)   a.   VP: [$_{VP}$ V O]

             Ordering: V < O

       b.   *CP: [$_{CP}$ C [$_{IP}$ S O$_i$ Adv [$_{VP}$ V t$_i$]]]

             Ordering: C < S

                       S < O

                       O < Adv

                       Adv < VP → Adv < V

(41b) schematically illustrates that the presence of a complementizer, in our case
(39b) *att*, blocks movement of V to C. Accordingly, the verb remains in situ. When
the object moves to the position left of adverbs, the result is an unacceptable sen-
tence like (39b). According to this approach, the ordering statements 'O < Adv' and
'Adv < V' given in (41b) imply that O precedes V (O < V) and thus contradict the
statement V < O that was established in the VP Spell-Out domain, as shown in
(41a). Hence, the unacceptability of (39b) is, according to Fox & Pesetsky, due to
this violation of Order Preservation.

Fox & Pesetsky (2005) go on to verify their cyclic, phase-based theory of lineari-
zation with regard to further facts about Scandinavian Object Shift and show that the

predictions of Order Preservation apply to various other cross-linguistic evidences as well. In light of all this evidence, they conclude "that a heterogeneous set of syntactic constructions appear to be constrained by a single factor – the avoidance of ordering contradiction" (Fox & Pesetsky 2005: 39). Having sketched fundamental aspects of this model, which accounts for the linear ordering of words by referring to different cycles, which means to different domains of Spell-Out, let us now return to the question of how this model of cyclic linearization can be applied to LP-movement in German.

5.2.1 A Minimalist Analysis of LP-Movement in German

Adopting crucial aspects of the derivational concept sketched above, Fanselow & Lenertová (2008) propose an analysis of LP-movement in German that abandons cartographic, representational devices and illustrates that "[t]here is no reason […] for coding information structure in the syntactic representation" (Fanselow & Lenertová 2008: 38). Specifically, and broadly speaking, instead of postulating representational categories of pragmatic interpretation, they argue that prosody is the primary means of expressing those interpretive properties, and, accordingly, they concentrate on prosodic rather than on purely syntactic properties, as the cartographic approach sketched in section 5.1 does. Consider how LP-movement is accounted for according to this perspective.

Fanselow & Lenertová (2008) claim that parts of, for instance, a focus can be moved to the left periphery because they bear the leftmost accent. Recall that, from a minimalist point of view, information concerning focality, contrastiveness, etc. – as we pointed out with respect to the 'Inclusiveness Condition' postulated in Minimalism – cannot be represented as syntactic information of lexical items. Rather, as Fanselow & Lenertová argue, in German, this information is linked to purely prosodic aspects. To gain concreteness, recall our example (35). We have already seen that there are basically two options of LP-movement with respect to the idiom *schöne Augen machen*. In particular, either *schöne Augen* is fronted, as in (35) – repeated below as (42) – or the whole idiom *schöne Augen gemacht* is preposed, as in (36) – represented here as (43):

(42)   [*Schöne AUgen*]ᵢ *hat er ihr* tᵢ *gemacht.*
      beautiful eyes   has he her   made
      'He made eyes at her.'

(43)   [*Schöne AUgen gemacht*]ᵢ *hat er ihr* tᵢ.
      beautiful eyes   made   has he her
      'He made eyes at her.'

As for prosodic structure, since in "German VPs the object precedes the verb and systematically receives the phrasal stress" (Truckenbrodt 2007: 444), the object *schöne Augen*, located between the indirect object *ihr* and the verb complex before moving to the left periphery, receives the phrasal stress, the structural accent of the clause. As Fanselow (2004) has already argued, note that the pars-pro-toto-movement in case of *schöne Augen machen* is only allowed for the part of the predicate that bears the structural accent (in our case: the object *schöne Augen*), whereas the verb of the idiom alone cannot move to the left periphery, as the following example shows (cf. Fanselow 2004: 22):

(44) *[*Gemacht*]ᵢ *hat er ihr schöne AUgen* tᵢ

By elaborating on this observation that syntactic movement is crucially constrained by prosodic properties like pitch accent, Fanselow & Lenertová (2008) demonstrate that in cases in which the sentence contains two structural accents, only the leftmost accented phrase can be moved to the left periphery. Look at the following example, where the primary accent of the clause is marked by capitals and the less strong accent is indicated by small caps (cf. Fanselow & Lenertová 2008: 9):

(45)   a. *Er hat den NAgel auf den KOPF getroffen.*
          he has the   nail   on the head hit
          'He clearly expressed the truth.'

     b. [*Den NAgel*]ᵢ *hat er* tᵢ *auf den KOPF getroffen.*

     c. *[*Auf den KOPF*]ᵢ *hat er den NAgel* tᵢ *getroffen.*

According to the general observation that "[o]n the level of the intonation phrase, the rightmost stress of the level of the p-[= phonological, A.T.]phrase is strengthened" (Truckenbrodt 2007: 446), the second phrasal part in constructions like (45a) – in our case, *den Kopf* – bears the primary accent of the clause, and the first part (*den Nagel*) receives a less strong accent. Crucially now, while moving the less accented (but leftmost accented) phrase is acceptable – as (45b) shows – the accented constituent *auf den Kopf* cannot be moved to the left periphery without losing its idiomatic reading, as (45c) illustrates. Accordingly, as Fanselow & Lenertova argue, it is not the prosodically most prominent accent that moves but rather the leftmost accent.

In light of examples like (45), Fanselow & Lenertová (2008: 15) claim that "[s]tructural accents are assigned in the context of immediate linearization." In doing so, they refer to the concept of 'cyclic linearization' sketched above, inasmuch as they assume that, when constituents are linearized, that is, are spelled out, their ordering cannot be changed at any later stage of the derivation. For clarification, let us apply this abstract conception to our example (45). Since the phrases *auf den Kopf* and *den Nagel* receive a structural accent, they have been linearized immediately when they were merged, and, consequently, the ordering statement '*den Nagel < auf den Kopf*' has been created. As this ordering cannot be contradicted, *den Nagel* can move to the left periphery, as (45b) shows, but *auf den Kopf* cannot move across *den Nagel*, as (45c) illustrates.

With these considerations, let us now turn to their analysis of LP-movement that is associated with a contrastive interpretation. According to their proposal, as mentioned, a phrase that receives structural accent is linearized immediately. So, as a consequence, a phrase that remains unaccented or receives a non-structural accent is regarded as being not linearized immediately. Crucially now, concerning accents that yield a focal or a contrastive interpretation, they claim that "[t]he accents corresponding to these functions do not depend on structural position, rather, they are assigned freely" (Fanselow & Lenertová 2008: 20-21). Thus, contrastive stress is not regarded as a structural accent and, according to their proposal, is not linearized immediately. To give an example of the consequences of this assumption within the concept of cyclic linearization, examine the following example, where the structural

accent is indicated by capitals and the non-structural accent is marked by small caps
(cf. Fanselow & Lenertová 2008: 24):

(46) The teacher is angry at Peter's friends.

    Aber [*beLEIdigt*]$_i$ *hat sie nur PEter* t$_i$

    but offended has her only Peter

    'But only Peter has offended her.'

In (46), the constituent *Peter* bears a structural accent and, according to the proposal
of Fanselow & Lenertova, is thus linearized immediately. The element *beleidigt*, on
the other hand, is contrastively accented. So, since it does not receive a structural
accent, it is not serialized immediately. Crucially now, if *beleidigt* has not been line-
arized immediately, according to this approach, its movement options are unre-
stricted. Note that this aspect modifies the proposal of Fox & Pesetsky (2005). Ac-
cording to Fox & Pesetsky's account, spelling out the structurally accented *Peter*
necessarily creates the ordering statement '*Peter < beleidigt*,' since, as Fox & Pe-
setsky hypothesize, domains of Spell-Out are linearized completely. In contrast,
Fanselow & Lenertová claim that the linearization of *Peter* within a syntactic structure
merely means that *Peter* is entered into a set of elements that are ordered in this
structure. Only when further material is entered into this set, must it be ordered rela-
tive to, for example, *Peter*. To sharpen this proposal, let us remember our example
(45) and compare it with (46) in light of this special concept of linearization. In case
of (45), according to the approach of Fanselow & Lenertová, both *den Nagel* and *auf
den Kopf* are entered into the set of elements that are ordered in the structure be-
cause both elements receive structural accent and, accordingly, are serialized im-
mediately. As a result, the ordering statement '*den Nagel < auf den Kopf*' has been
created and cannot be contradicted by later syntactic operations. In contrast to cases
such as (45), in (46), the set of ordered elements has only a single member (*Peter*),
so ordering statements are trivial because, only when further elements are entered
into the set of ordered elements, are relevant ordering statements created. Concern-
ing this 'relational' aspect of their theory, Fanselow & Lenertová adopt crucial proper-
ties of recent generative approaches according to which "ordering statements are
relativized, in the sense that only a subset of the ordering statements that could in

principle be generated [… is] in fact generated by syntactic operations" (Müller 2007: 83). So, in sum, since the object *Peter* is the only constituent that bears a structural accent, no ordering statement is created that prevents *beleidigt* from moving across *Peter*. Therefore, elements with a non-structural accent that is associated with pragmatic properties like contrast undergo 'late' – 'secondary,' as Fanselow & Lenertová (2008) call it – linearization.
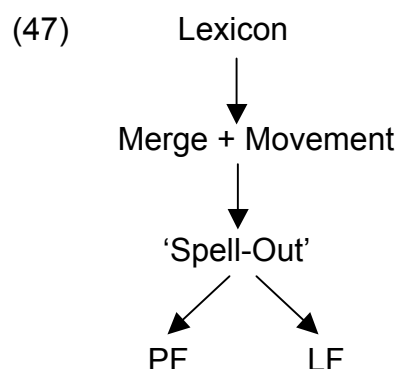
Let us take stock at this point. By adopting a phase-based view of linearization theory, postulating multiple Spell-Out domains, Fanselow & Lenertová account for the phenomenon of LP-movement in German without referring to information structural properties. In particular, according to their model, syntactic movement to the CP-domain is constrained by linearization processes that are directly linked to accentuation. In contrast to the cartographic approach discussed in section 5.1, their model claims that information structural properties do not play a role in syntactic operations. In other words, pragmatic interpretations like 'focus' or 'contrast' are only linked to freely-assigned accents and thus only interface with phonology. Since "there is no direct link between information structure and syntactic movement" (Fanselow & Lenertová 2008: 2), pragmatic properties of a sentence that are associated with its information structure do not have to be encoded in the syntactic representation at all. Turning now from this rather narrow, 'microscopic' perspective on the place of information structure in the grammar to the general 'bird's-eye-view' again, let us point out the conceptual implications of this approach and their consequences for reevaluating Jackendoff's critique concerning the view of information structure within syntactocentrism.

5.2.2 Derivational Syntactocentrism and the Parallel Architecture: Perspectives of Convergence

In the last section, I presented an analysis of the pragmatics of LP-movement in German that abandons the encoding of information structural notions in the syntactic representation by postulating that only prosodic properties, thus the phonological component of the grammar, is directly linked to information structure. More precisely, this analysis is based on a version of cyclic linearization in which only structural accents have an effect on locality constraints that are involved in syntactic movement;

non-structural accents like focal stress, by contrast, are assigned 'freely' and hence show no interaction with the syntactic computation at all. In this section, after having dealt with these specifics of German syntax, I will broaden the view again and demonstrate in what sense this cyclic, phase-based approach shows significant points of convergence with conceptual underpinnings of Jackendoff's Parallel Architecture, sketched in section 4.2. To do so, I will first turn to general issues concerning the architecture of grammar and then end with a reflection on the relation between syntax and pragmatics, and I will focus on its conception in both approaches.

Concerning architectural issues, it may be worth recalling that, according to Jackendoff (2003: 658), "[i]n a syntactocentric theory, one is forced to generate […] sentences with a dummy syntactic element such as [+Focus], which serves only to correlate phonology and meaning." As we saw in section 5.1 of this chapter, this conception of postulating 'dummy elements' in the syntactic representation finds its most extreme expression in cartographic approaches, which assume dedicated functional projections and corresponding features in order to transparently encode notions like focus or contrast in the syntactic representation. And indeed, this encoding seems to be necessary on conceptual grounds, when one looks at the most recent grammar model of mainstream generative grammar that Jackendoff discusses, which is depicted here again:

(47)   Lexicon
             |
             v
    Merge + Movement
             |
             v
        'Spell-Out'
           /   \
          v   v
        PF   LF

Regarding this architecture, an encoding of notions such as focus or contrast is conceptually required because, as has often been noted in the context of research on information structure, this model "permits no direct interaction between the PF and the LF parts of the derivation" (Horvath 2007: 112). Such an interaction, as we indicated in preceding sections, may be crucial for phenomena like prosodically expressed focus in, for instance, English and German. According to this architecture,

however, both the prosodic information concerning some particular stress and the information regarding, for instance, contrastive interpretation have to be transported through the syntax. Now, as Jackendoff argues, since information structural interpretation is primarily expressed by prosodic properties, at least in the languages we are concerned with here, this architecture and the required machinery to encode pragmatic notions like focus and contrast in the syntax seem to be a rather artificial result of the syntactocentric view. He clarifies this point best, as we already saw, by citing examples in which identical syntactic structures can have different meanings with respect to information structural properties of the sentence. To illustrate, consider the following case again:

(48)   J: Who went to the party?

     K: a.   *PAT went to the party.*

         b.   **Pat went to the PARTY.*

We saw in section 5 that the question in (48) can be answered with (48a) but not with (48b) because, although the syntax of both replies is identical, only in (48a) the prosody marks *Pat* as the focus, the new information required by the question given in (48). As we saw in the discussion in section 5.1.1, the examples cited by proponents of the cartographic approach also imply a strong correlation between, for instance, a contrastive interpretation and a prominent stress on some specific constituent. In particular, only the type of movement that is associated with placing prominent stress on the preposed constituent results in a contrastive interpretation of this particular fronted element. So, for example, while in (49) the left peripheral element is not stressed and does not receive a contrastive interpretation, the stressed case given in (50), which is, according to the cartographic approach, due to a different type of movement, can only be interpreted contrastively, similar to our paradigm case (51):
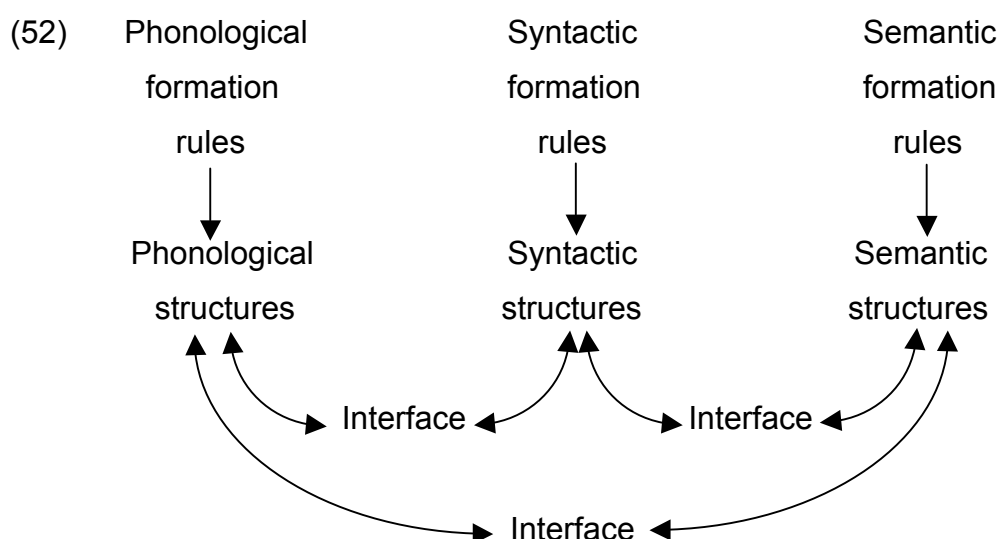
(49)   *In Europa spielen Jungen gerne Fußball.*
       in Europe play   boys   Adv.   football
       'In Europe, boys like to play football.'

(50) *In EuROpa spielen Jungen gerne Fußball* (*und nicht in Asien*).
     in Europe play   boys   Adv.   football   (and   not   in   Asia)
     'In Europe, boys like to play football.'

(51) *GRÜN wird Maria die Tür streichen* (*und nicht rot*).
     green   will   Maria   the   door   paint   (and   not   red)
     'Maria will paint the door green.'

Observations like these strongly suggest, in Jackendoff's view, that "a direct phonology-semantics interface […] is attractive for the correlation between prosody and information structure" (Jackendoff 2003: 658). Consequently, his solution, as we saw in section 4.2, is the Parallel Architecture, which permits a direct interaction between phonological and semantic structures without any mediation of syntax, as shown in (52):

(52)

| Phonological formation rules	Syntactic formation rules	Semantic formation rules
↓	↓	↓
Phonological structures	Syntactic structures	Semantic structures

Interface   Interface   Interface

With this proposal of a direct interaction between phonology and semantics in mind, let us now turn to the derivational analysis illustrated in the last section. Since this analysis, as we noted, also focuses on the prosodic means of expressing information structural properties, this approach, like Jackendoff's, also criticizes the syntactic encoding of notions like contrast, inasmuch as "[t]he pertinent features and machinery reflect little more than an attempt to uphold an architectural assumption (PF does not influence LF) in the light of counterevidence (prosody is crucial for interpretation)" (Fanselow 2008: 398). Let us see how the claim that PF does influence LF without

mediation of syntax fits the architecture of grammar in recent derivational syntacto-centrism.

As illustrated, Fanselow & Lenertová (2008) assume a version of cyclic linearization. This concept implies an architecture with multiple Spell-Outs of partial pieces (lf and pf) rather than one single mapping of fully-fledged structures (PF and LF) at the end of the derivation. I have already sketched the underlying architectural assumption in section 3.2.2, and I have depicted the resulting model as follows:

(53)

```
        lf          pf

        lf          pf

        lf          pf

        .           .
        .           .
        .           .
```
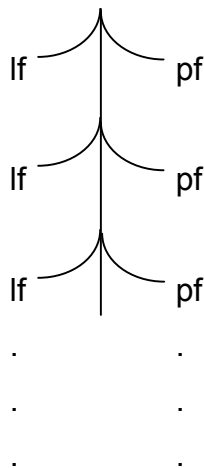
An illustrative example of how this assumption of multiple Spell-Out works within the approach of Fanselow & Lenertová (2008) has been the following:

(54) a. *Er hat den NAgel auf den KOPF getroffen.*
          he has the nail   on the head hit
          'He clearly expressed the truth.'

      b. [*Den NAgel*]$_i$ *hat er* t$_i$ *auf den KOPF getroffen.*

      c. *[*Auf den KOPF*]$_i$ *hat er den NAgel* t$_i$ *getroffen.*

As I mentioned, (54) clarifies that, in this case, the leftmost accented phrase *den Nagel* can be moved to the left periphery, as shown in (54b), while the accented constituent *auf den Kopf* cannot be preposed without losing its idiomatic reading, as (54c) indicates. Assuming their concept of linearization, Fanselow & Lenertová argue that both *auf den Kopf* and *den Nagel* have been linearized immediately, that is, spelled-out as pieces of the whole syntactic structure because they show a structural

accent. Accordingly, the ordering statement 'den Nagel < auf den Kopf' has been created, which cannot be contradicted in subsequent steps of the derivation, as (45c) demonstrates. According to this view, the assignment of structural accents is a product of Spell-Out, that is, structural accents are assigned at the interfaces. So, given this analysis, the products of two Spell-Out operations, that is, the structural accents on *auf den Kopf* and on *den Nagel*, determine subsequent aspects of the syntactic derivation, inasmuch as these products of Spell-Out either allow or ban the syntactic operation of movement to the left periphery. In light of general architectural concerns, this illustrates a significant difference to earlier models. In particular, in models like (47), accentuation properties only emerge after one single, one final Spell-Out of a fully-fledged structure at the end of a derivation. In contrast to this conception, the analysis above postulates that these aspects of accentuation can determine factors of the syntactic derivation before the derivation has been finished as a whole.[19] With this aspect of spelling out partial pieces of the derivation in mind, let us look at the contrastively-accented case again:

(55)   The teacher is angry at Peter's friends.

   Aber [beLEIdigt]$_i$ *hat   sie   nur   PEter* t$_i$

   but   offended   has   her   only   Peter

   'But only Peter has offended her.'

In this case, although *beleidigt* has been merged with *Peter*, as indicated by the trace, it does not have to be spelled out together with *Peter*, since, within a phase-based architecture like (53), *Peter* alone can be spelled out as one piece and already receive its structural accent, whereas *beleidigt* can receive its accent after the whole syntactic derivation has been finished. Accordingly, as the accent of *beleidigt* does

---

[19] Strictly speaking, in most recent approaches, the Spell-Out of small pieces and the syntactic computation of more complex parts are regarded as "operations proceeding in parallel" (Chomsky 2008: 147). That is, the subparts of the derivation are constructed in parallel, inasmuch as small syntactic objects generated in the course of the derivation are mapped to the phonological and the semantic component, but at the same time they function as an element of another more complex subset in the syntactic derivation. In our context, however, it is not necessary to elaborate on this point, since the concept that prosodic properties were only present after the whole derivation in earlier models, whereas these properties can now influence ongoing syntactic derivations, can be clarified without delving into complex issues of parallel computation.

not depend on its syntactic position, otherwise it has been spelled out like *Peter*, the accentuation of *beleidigt* after the whole derivation is 'invisible' for the syntax, as we indicated in the last section by stressing that its movement options are unrestricted. Crucially now, this analysis implies that, when *beleidigt* is interpreted contrastively, the correlation between PF – non-structural stress – and LF – contrastive interpretation – has not been established by syntax. Because of this conceptual implication of such strong derivational analysis, some authors speak of an interaction between PF and LF with respect to recent models of mainstream generative grammar and point out that "PF has access to both, the syntactic derivation of the phase […] and the semantic interpretation" (Winkler 2005: 24). In other words, recent derivational syntactocentrism allows, in some cases, for an interaction between sound and meaning without the mediation of syntax. This point of convergence between syntactocentrism and the Parallel Architecture has, to my knowledge, so far only been mentioned by Winkler (2005), who notes that the strong derivational model within minimalism "turns out to be conceptually closer to Jackendoff's […] tripartite parallel model of grammar than might be recognized at first sight" (Winkler 2005: 231, n. 8). However, she does not elaborate on this point and, unfortunately, leaves it at a short comment in a footnote.

Of course, the question that arises in this context is what accounts for the mediation of sound and meaning. According to the analysis sketched above, the stress on *beleidigt* is a "pragmatically determined accent" (Fanselow & Lenertová 2008: 14) and thus, as the authors claim, is only determined by the pragmatic context the utterance is used in. However, given this proposal, one still has to explain the fact that the German cases with contrastive interpretation I discussed in previous sections involve movement to the left periphery. Here again, Fanselow (2008: 406) argues that these structural configurations are due to broad pragmatic concepts like "increasing the listener's attention by the choice of an unexpected syntactic construction." This strict separation of syntax from pragmatics has further implications for some objections to syntactocentrism that are raised by Jackendoff.

Another argument Jackendoff gives in favor of viewing the semantic component as largely independent of syntax concerns the aspect of illocutionary force, that is, "whether certain words […] *had the force of* a question, or *ought to have been taken as* an estimate and so on" (Austin 1962: 99, emphasis in the original). As in the case of information structure, he claims that this interpretive property shows no system-

atic, isomorphic relation to syntactic structure. Look at the following examples (cf. Jackendoff 2003: 657):

(56) a. (*Now, Hans:*) *Is Kehl in France?*

 b. *Is the toilet over there?*
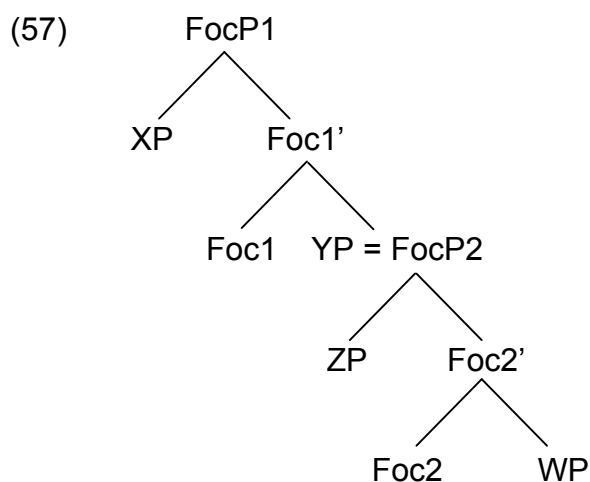
 c. *Is the Pope Catholic?*

The examples in (56) show that the syntactic form of a question can be used to test someone's knowledge (56a), elicit information (56b), or sarcastically express an affirmative answer to a prior question (56c). In light of such cases, Jackendoff (2003: 657) concludes that "choices of illocutionary force are not mapped into syntactic structure." When we adopt the view that pragmatics is a separate component, however, these observations create no problem for mainstream generative grammar. Like in Jackendoff's architecture, these properties are relegated to an independent component, to a "system of rules and principles constituting pragmatic competence" (Chomsky 1980: 224). The difference between both approaches, in this case, lies only in the fact that Jackendoff tries to make this pragmatic competence explicit in terms of Conceptual Semantics, "in which semantic structure is subsumed under conceptual structure" (Jackendoff 1983: 19), whereas mainstream generative grammar limits itself more and more to focus 'on pure syntax' – thus the title of Fanselow (2006).

In sum, some of Jackendoff's objections to syntactocentrism disappear in light of recent derivational analyses like the cyclic linearization account of LP-movement in German sketched in this chapter. In particular, we saw that recent syntactocentrism allows, in some cases, for direct interaction between phonology and semantics, much like Jackendoff's Parallel Architecture. Furthermore, I illustrated that the recently-sharpened separation of syntax from pragmatics qualifies further objections to syntactocentrism. Certainly, both approaches still differ in numerous respects. However, regarding architectural issues – and that is what the notion of syntactocentrism originally refers to – the gulf between derivational syntactocentrism and the Parallel Architecture is not as huge as Jackendoff never tires to suggest.

Concerning this slight convergence, a question that arises in light of the considerations above is if there is any criterion within recent syntactocentrism for relegating

aspects of language to syntax or to pragmatics. In other words, what does belong to 'pure syntax' and what can be excluded? This question pertains to the general conception of the syntactic component in recent syntactocentrism, and I will now address this issue briefly – admittedly, quite speculatively – in the remainder of this chapter.

Remember from section 2.1, that the syntactic component in recent syntactocentrism "has one operation that comes 'free,' in that it is required in some form for any recursive system: the operation Merge" (Chomsky 2004: 108). As this quotation emphasizes, the main feature of this operation is recursion because Merge takes previous applications of Merge as its input. Accordingly, recursion as a tool for creating hierarchically embedded structures might be a good criterion for including or excluding issues in syntactic research that is committed to minimalist guidelines. According to this criterion, it seems quite reasonable to relegate information structural notions to pragmatics and exclude them entirely from syntax. Interestingly, a hint in this direction can be found in Rizzi's (1997) discussion of focus. Rizzi stresses that the following structure, in which a FocP is embedded into another FocP, poses an interpretive problem (cf. Rizzi 1997: 297):

(57)

```
          FocP1
         /     \
       XP       Foc1'
               /     \
            Foc1   YP = FocP2
                   /       \
                 ZP        Foc2'
                          /     \
                       Foc2      WP
```

According to the cartographic approach, the specifier of FocP2 (ZP) must host a focal constituent, that is, an element that is interpreted as new information. However, FocP2 is located in the position YP, which is dedicated to the presupposition of the higher focal head Foc1. Since FocP2 can thus only specify given information, Rizzi (1997: 297) concludes that "recursion of FocP is banned by the interpretive clash that would arise." In non-cartographic terms, since notions like focus only imply a

bipartite distinction like new vs. old, there cannot be any embedding of one part within a larger part, of one focus within another focus. And yet, there are some proposals that argue for the existence of hierarchical embedding in information structural configurations. In particular, the phenomenon of 'second occurrence focus' (cf. Büring 2008) and the claim of recursive topic-marking in Japanese (cf. Kuroda 2005) are often cited in order to argue for recursion in information structure. However, given even these observations, to most scholars – and I concur – it seems safe to conclude "that structure in IS [= information structure, A.T.] is rather rudimentary, and […] we have to look hard for any linguistic relevance that such structural characteristics may bring about" (Tomioka 2007: 97). Accordingly, since notions like focus, contrast, or topic do not clearly show the feature of recursive hierarchical embedding, they can be excluded from syntactic research within recent mainstream generative grammar.

Turning to the main track once again, I pointed out some conceptual points of convergence between recent syntactocentrism and the Parallel Architecture in light of a recent derivational analysis of LP-movement in German. This analysis, with its exclusion of interpretive properties from syntax, is a good example of the more general minimalist framework, which dictates "to examine every device […] to determine to what extent it can be eliminated in favor of a principled account […], going beyond explanatory adequacy" (Chomsky 2004: 106). It is this 'going beyond explanatory adequacy' that has recently initiated research into the evolutionary origins of language – a field in which even points of convergence between derivational syntactocentrism and Cognitive Linguistics can be demonstrated.

6. Derivational Syntactocentrism and Cognitive Linguistics: Approaches to Language Evolution

As we saw in the last chapter, the strong derivational view of syntactocentrism implies significant points of convergence with conceptual aspects of the Parallel Architecture. In particular, both approaches share the assumption that a direct interaction between phonology and semantic interpretation is required in order to deal with phenomena like prosodically-expressed focus or contrast. This interaction is, according to these approaches, established by pragmatic factors rather than by anything that could be accounted for in syntactic terms. In the componential view of both approaches, pragmatics can be regarded as an independent component, which is described in the Parallel Architecture and omitted in the descriptive apparatus of mainstream generative linguistics but not denied as a crucial component for human language.

As already mentioned in section 4.2, Cognitive Linguistics does not share this componential view. Consequently, there is less potential for convergence with respect to architectural issues. However, if we elaborate on the aspect, exemplified in the context of LP-movement in German, that much of the representational richness of syntax is abandoned in order to focus on combinatorial essentials, some points of convergence even between derivational syntactocentrism and Cognitive Linguistics can be demonstrated, too. In this chapter, I will argue in this direction by focusing on how both approaches conceive of language evolution. Needless to say, the issue I touch on here is among the richest and most controversial fields of linguistics and related disciplines like psychology and biology. Since I cannot hope to exhaustively cover even those approaches that I selected, I will favor intellectual coherence over exhaustive coverage and only mention aspects that I need for my argument to go through. Therefore, in section 6.1, I will briefly sketch an approach to language evolution that is associated with the strong derivational view on syntactic computations. Then, in section 6.2, I will illustrate an approach that focuses on language as a communicative system and crucially rests on the conceptual underpinnings of Cognitive Linguistics. Finally, in section 6.3, I will point out significant points of convergence between both approaches.

6.1 The Computational View: Recursive Syntax & FLN

In a widely-cited paper, Hauser et al. (2002) aim at setting an agenda for productive interchange between biologists and linguists in the context of research on the evolutionary origins of language. They argue that this interchange has often been blocked due to the lack of a unified concept of language. To avoid several confusions that arise from missing a conceptual unity, they emphasize that "it is important to distinguish between questions concerning language as a communicative system and questions concerning the computations underlying this system" (Hauser et al. 2002: 1569). Having stressed this distinction, they focus on language as a computational system and adopt the mentalist perspective that postulates the entity of grammatical knowledge as its object of inquiry. However, since theories that adopt this mentalist view abound, as we indicated in section 4.2, Hauser et al. claim that the theoretical diversity within this paradigm calls for further constraining of the mentalist conception of language. In this context, they draw the already famous distinction between 'Faculty of language – broad sense (FLB)' and 'Faculty of language – narrow sense (FLN).'

According to this distinction, FLB is an inclusive term of broad range that includes both FLN and the mechanisms that can be relegated to more general cognitive capacities, which, as Hauser et al. (2002: 1573) claim "are shared with other animals […] with difference of quantity rather than kind." The aspects of the human language faculty that are hypothesized to be shared with other animals concern, as they call it, the 'sensory-motor system' and the 'conceptual-intentional system.' Broadly speaking, these two components contain both the capacity to physically perceive and produce sounds and to master both semantic and pragmatic aspects of meaning. In contrast to FLB, as Hauser et al. (2002: 1573) hypothesize, "FLN comprises only the core computational mechanisms of recursion as they appear in narrow syntax and the mappings to the interfaces". Based on a large amount of data, they argue that the capacity for syntactic recursion, described in minimalism as the operation Merge, cannot be derived from other more general cognitive capacities, and thus, they suggest the hypothesis that "the computational mechanism of recursion […] is recently evolved and unique to our species" (Hauser et al. 2002: 1572). Note that singling out a particular aspect that gave rise to the evolution of human language opens the door for theoretical alternatives to the prominent view that "human language, like other

specialized biological systems, evolved by natural selection" (Pinker & Bloom 1990: 726), and thus, this system is due to stepwise, fine-grained processes of adaptation. The additional scenario that can be hypothesized by positing only one feature is that of a 'great leap forward,' caused by the fact that "the brain was rewired, perhaps by some slight mutation" (Chomsky 2005: 11-12). However, since these different evolutionary scenarios are part of a more substantial debate and involve complex concepts like exaptation and adaptation, I cannot go into this here (for a more extensive illustration of these two views and their conceptual foundations, see Trotzke 2008: 30-51). Instead, let us now turn to the empirical line of research that this constrained hypothesis regarding language evolution has initiated both in comparative psychology and in neuroscience. Let us first concentrate on comparative psychology.

Fitch & Hauser (2004: 378) focus on the issue that humans, due to their capacity for syntactic recursion, "can embed strings within other strings, thus creating complex hierarchical structures [...] and long-distance dependencies." To make this concrete, remember our discussion of syntactic rule systems in section 1.1. There, we saw that a finite-state grammar (FSG) cannot account for long-distance dependencies, while a phrase structure grammar (PSG) like the rewrite-rule system can capture these embedded structures. To recall, consider the following cases:

(1)   [*The man*]$_A$ [*the dog*]$_A$ [*bit*]$_B$ [*comes*]$_B$.

(2)   [*The man*]$_A$ [*comes*]$_B$; [*the dog*]$_A$ [*bit the man*]$_B$.

In chapter 1, I pointed out that a finite-state grammar cannot account for center-embedded structures like (1) because it computes a sequence in a strictly local manner and thus does not 'know' what states it has been in and how many times it has been in some particular state. Accordingly, with respect to (1), it cannot ensure that it will generate an equal number of As and Bs and thus cannot capture the non-local relation between *the man* and *comes*. Due to this computational limitation, a finite-state grammar can only generate paratactic constructions like (2), in which one A is followed by one B in a strict local manner. Given this distinction between a FSG and a PSG, Fitch & Hauser (2004) tested the parsing abilities of cotton-top tamarins (*Saguinus oedipus*) regarding both grammar types. Using two different classes of

syllables that were modeled to correspond to A and B, respectively, they create stimuli that correspond to the artificial grammars A^nB^n, generating structures like (1), and $(AB)^n$, yielding structures like (2). Again, I will cut sharply here, since I have already illustrated this particular study and an analogous experiment with songbirds at length elsewhere (cf. Trotzke 2008: 53-64). The result of their study was "that tamarins suffer from a specific and fundamental computational limitation on their ability to spontaneously recognize or remember hierarchically organized acoustic structures" (Fitch & Hauser 2004: 380). That is, while tamarins were able to process structures generated by the $(AB)^n$ grammar, they were not capable of mastering structures according to the A^nB^n formula. Accordingly, this experimental study supports the hypothesis that "the acquisition of hierarchical processing ability may have represented a critical juncture in the evolution of the human language faculty" (Fitch & Hauser 2004: 380) and thus may be of direct relevance to the evolutionary scenario suggested by Hauser et al. (2002). Beside further experiments with nonhuman species, this comparative study also inspired neuropsychological studies that ask to what extent the core computational faculty of processing hierarchical embedded structure can be segregated from other brain functions. Let us briefly look at this field of research.

Friederici et al. (2006) build on the findings of Fitch & Hauser (2004) and hence assume that humans differ from non-human primates in their capacity to master sequences that are generated by the A^nB^n grammar. In their study, they ask, broadly speaking, whether the differences of processing the two grammars used by Fitch & Hauser (2004) are reflected in the human brain. To explore this question, they test human subjects by visually presenting sequences of consonant-vowel syllables that were modeled to represent the different grammar types. After having used these stimuli and after having applied several sophisticated testing procedures, they indeed conclude that there are differences in processing in the brain. In particular, Friederici et al. claim that processing of local transitions within a finite-state grammar is subserved by the left frontal operculum, whereas a specific section of Broca's area holds responsible for the computation of hierarchical dependencies involved in syntactic recursion within a phrase structure grammar. Crucially, Friederici et al. (2006: 2461) point out "that the grammar type processed by human and non-human primates is subserved by a brain area, which is phylogenetically older than the brain area subserving the processing of the grammar type only learnable by humans."

Thus, this neuropsychological study can also be regarded as empirical support for the hypothesis that syntactic recursion has only recently evolved as a unique feature of human language. However, some qualification is in order here because stimuli that are generated by $A^n B^n$ and $(AB)^n$, as used both in this neuropsychological experiment and in the study by Fitch & Hauser (2004), face some serious problems.

It was soon pointed out that natural human language requires the ability to process sequences in which a consistent coupling of AB-pairs is involved. That is, structures like AABB, since they indicate no specific relation between particular As and Bs, are inadequate to model human languages. To visualize, the more exact representation of our sentence (1), according to this objection, must be (3), where the pairing of particular As an Bs is marked by numbers:

(3) $[$*The man*$]_{A1}$ $[$*the dog*$]_{A2}$ $[$*bit*$]_{B2}$ $[$*comes*$]_{B1}$.

To ground this critique in experimental evidence, Perruchet & Rey (2005) demonstrate, by applying several probe procedures, that human subjects, when tested to process AABB sequences that do not indicate this pairing, "did not process the material as a center-embedded structure" (Perruchet & Rey 2005: 310). Rather, as they argue, humans master such structures by alternative strategies such as simple counting. In particular, subjects can merely count the As and match the resulting number with the amount of Bs in order to master such tests as conducted by Fitch & Hauser (2004).

Given this fundamental problem, it is now a central concern in this field of inquiry "that the relation between artificial language studies and natural language must be clarified" (Hauser et al. 2007: 127). As a consequence, both empirical and conceptual contributions to investigate the nature and format of recursive embedding have become more differentiated. As for empirical studies, turning to brain studies again, Bahlmann et al. (2008) ask whether the activation in Broca's area observed by Friederici et al. (2006) is independent of any other more simpler processing strategy. So, adopting the proposal of pairing the particular As and Bs, they created stimuli material that, in their view, "ensures that in the hierarchical dependency condition sequences are processed by participants in an embedded manner" (Bahlmann et al. 2008: 526). By using these more refined stimuli, they were able to show „that the

activation in Broca's area is indeed due to the processing of hierarchical dependencies, and not to a more simple processing strategy" (Bahlmann et al. 2008: 526). On the conceptual side, it is now explicitly pointed out that the "notion that A^nB^n requires recursion is incorrect, and appears to reflect an inadequate grasp of computer science and formal language theory" (Fitch 2010: 87).

In sum, whether one believes in syntactic recursion as *the* unique feature of human language or not, recent and ongoing research that has been initiated by Hauser et al. (2002) has led to deepen both the theoretical notion of recursion in language and the methodological discussion concerned with the creation of adequate experimental designs to test the processing of hierarchically-embedded structures.

6.2 The Communicative View: Shared Intentionality

In addition to the computational approach to language evolution, there is another line of research that "places the human ability for complex symbolic communication at the centre of the evolution of language" (Christiansen & Kirby 2003: 6). One prominent approach within this 'communicative view' is Tomasello's (2008) recent theory, which fits well to several conceptual assumptions of Cognitive Linguistics and is, in fact, the only full-blown evolutionary theory associated with this framework that does not only deal with processes of cultural transmission, of 'cultural' evolution, but also with aspects of 'biological' evolution. In this section, I will sketch the core ideas of this theory and also, as I did regarding the computational approach, mention some objections. This illustration, together with the short overview of the view on language evolution associated with recent syntactocentrism in the last section, will serve as the basis to point out some points of convergence between approaches to language evolution committed to derivational syntactocentrism and accounts in line with Cognitive Linguistics.

Tomasello (2008) claims that all of the grammatical complexity has evolved to serve the special functional demands of human communication. Since Tomasello, like Hauser et al. (2002), also argues that human language, compared to other species, is a unique capacity, he claims that there is some unique feature about human communication. According to his communicative view, the crucial difference to other species lies in the human capacity for 'shared intentionality.' This capacity has al-

ready been analyzed extensively in Analytical Philosophy as 'we intentionality' or 'collective intentionality' and refers, in this theoretical context, to a particular mental state that accounts for the "[b]ackground sense of the other as a candidate for cooperative agency" (Searle 1990: 414). To approach this capacity more psychologically, Tomasello divides it into two components: "(i) the cognitive skills for creating joint intentions […] with others; and (ii) the social motivations for helping and sharing with others" (Tomasello 2008: 73). Let us consider both components in turn.

First, the cognitive capacity for generating joint intentions with others contains the ability "to determine both what the communicator is directing attention to (his referential intention) and why he is doing it (his social intention)" (Tomasello 2008: 75). To make this capacity more vivid, imagine that Hans sits in his favorite bar, and he points to his empty glass to request another beer from the bartender. Concerning the 'referential intention,' the bartender must know, in this case, that Hans is pointing to the emptiness of the glass and not to its color, its shape, etc. Regarding the 'social intention,' the bartender understands the pointing because both Hans and the bartender know that customers like Hans, under normal circumstances, are at the bar to drink, and an empty glass does not enable drinking. Note that reading the social intention in the right way is not trivial, since it could also be the case that the bartender knows that Hans is an alcoholic that wants to quit, in some context of exposure therapy, by sitting at the bar without drinking. In this, admittedly 'exotic,' case, the pointing indicates to the bartender that Hans has still managed to resist drinking. However, the fact that the communication between Hans and the bartender works well in most cases is due to complex processes involved in determining referential and social intention that have been investigated in the literature under the roof of 'common ground' (for an overview of this notion, see Clark 1996: 92-121).

Let us now turn to the second component of shared intentionality mentioned above. As for the social-motivational infrastructure of shared intentionality, Tomasello postulates

> three general types of evolved communicative motives […]: **Requesting:** I want you to do something *to help me* […]; **Informing:** I want you to know something because *I think it will help or interest you* […]; **Sharing:** I want you to feel something so that *we can share attitudes/feelings together.* (Tomasello 2008: 87, emphasis in the original)

Notice that both the communicative motive of requesting and the one of informing involve helping: Concerning the motive of requesting, one complies with a request by helping the requesting person; regarding the motive of informing, one helps others by informing them of useful things. Accordingly, Tomasello summarizes these two motives in some contexts, like in the definition given above. The assumption of these three basic human communicative motives is based on a large amount of experimental studies that demonstrate that nonhuman primates do not show these pro-social motives in various tested contexts (for a concise review, see Tomasello et al. 2005: 684-686), while human infants show both the motive of sharing an attitude with an adult and the motive of providing the adult with useful information in the context of pointing (cf. Tomasello et al. 2007). Putting these basic pro-social motives together with the cognitive capacity to generate joint intentions creates, according to Tomasello, the mutually-known assumption that human communication is cooperative and, accordingly, forms the 'cooperative model of human communication' (for a summarizing schema of this model, see Tomasello 2008: 98). Since Tomasello assumes that language, in the sense of grammatical structures, has developed to fulfill the needs of this cooperative model of human communication, let us now focus on the question of how these structures have emerged according to this theory.

Regarding human language as a device to serve cooperative communication, Tomasello (2008: 244) claims "that the purpose for which one communicates determines […] what kind of grammatical structure is needed." Consequently, supposing the three basic motives to communicate mentioned above, Tomasello (2008: 243-295) describes the emergence of complex grammatical structure as follows: Requestive motives are already present in our ancestors and involve only 'me and you in the here and now,' thus creating no functional pressure for any serious syntactic marking. With the informing motive, understood as offering help, grammatical devices evolved for such complex functions as indicating referents displaced in time and space. Finally, the emergence of the sharing motive demands a syntax that provides devices for the narration of complex series of events displaced in time and space. Note that, while the pro-social motives underlying the grammatical structure have their roots in the biological evolution of a specific 'psychological infrastructure,' as Tomasello calls it, "the actual grammatical conventions are […] not created by evolutionary processes at all: they are created by […] the conventionalization of grammatical constructions" (Tomasello 2008: 317). As this quote indicates, in outlining

this cultural-historical process of conventionalization, Tomasello heavily relies on basic concepts of Cognitive Linguistics, namely on a construction-based view of grammar and on models that highlight the necessity to "look to the diachronic dimension to learn how the conventions of grammar arise if we are to know why they take the particular form that they do" (Bybee 1988: 351). With this approach to language evolution in mind, which adopts core ideas of Cognitive Linguistics and thus exemplifies that, "[m]ore recently, cognitive psychologists have been influenced by cognitive linguistic research" (Croft & Cruse 2004: 329), let us now turn to some problems with this account.

As mentioned, the assumption of the three basic human communicative motives is grounded in numerous experimental studies, mainly on infant pointing. In these studies, Tomasello et al. have observed that, in contrast to nonhuman primates, there are instances of infant pointing that cannot be classified as imperative actions but are driven by the cooperative motives of helping and sharing. However, based on various – and to my mind, compelling – experimental findings, Southgate et al. (2007: 735) argue "that each of the instances of infant pointing may in fact have a more selfish motive." In particular, they claim that each case of pointing explored by Tomasello et al. can be regarded as a case in which pointing actions function to specify a referent that infants want to obtain information about. So, infant pointing in general can be classified as an interrogative act that serves the selfish need to, roughly speaking, 'learn about the world.'[20] This different interpretation of the data reveals a more fundamental problem with the experimental grounding of pro-social motives.

To point out this problem, it may help to contrast the recent hypothesis concerning pro-social motivation with former accounts of human uniqueness, proposed by Tomasello et al. in the 1990s. Tomasello & Call (1997: 353) argue "that no nonhuman primates understand the behavior of conspecifics as intentional or mental," and thus, they claim that the cognitive capacity to represent mental states of others is a uniquely human feature. However, their subsequent research shows more and more that nonhuman primates "also have some natural social cognitive skills to understand and predict many aspects of the behaviour of others" (Hare et al. 2000: 784)

---

[20] Interestingly, Tomasello (2008) himself suggests this more selfish interpretation in some contexts by pointing out, for instance, that "[h]elping motives […] can flourish in mutualistic collaboration in which helping you helps me" (Tomasello 2008: 198).

and, accordingly, can represent (at least some) mental states of others. Crucially, compared to the claim of the 1990s, the recent hypothesis is no longer a representational one that denies that nonhuman primates have some cognitive capacity concerning mental representation. Instead, Tomasello et al. now argue that, in some contexts, for instance of sharing, nonhuman primates "are not *motivated* in the same way as humans to share emotions, experiences, and activities with others" (Tomasello et al. 2005: 686, emphasis added). This shift from a representational hypothesis that makes firm predictions – either nonhuman primates have this capacity or they do not possess it – to an account that relies on notions of motivation poses serious problems for experimental studies, since "the inherent subjectivity of motivation as an explanatory construct makes Tomasello et al.'s […] hypothesis dangerously close to non-falsifiable" (Lyons et al. 2005: 708).

To summarize this section, Tomasello (2008) argues that the uniqueness of human language is due to the distinguished character of human communication. Human communication differs from other forms of communication because it implies shared intentionality. This aspect of human communication concerns both the generation of joint intentions and the existence of pro-social motives. As we indicated, while the generation of joint intentions is well-grounded in both the literature of pragmatics and in earlier studies of Tomasello et al., the existence of pro-social motives, when we judge from the objections mentioned above, can be seriously questioned, inasmuch as there seem to be several problems to define motivation in an experimentally quantifiable, and that is, falsifiable way. Given these problems, in the final section of this chapter, I will abstract away from these motivational aspects and focus on the cognitive component of Tomasello's theory in order to compare this aspect of his account with the computational view on language evolution sketched in section 6.1.

6.3 Derivational Syntactocentrism and Cognitive Linguistics: Perspectives of Convergence

As we saw in the preceding sections, both the account associated with recent syntactocentrism and the account committed to the general outlook of Cognitive Linguistics assume that human language – be it from the computational or from the communicative point of view – is unique and, accordingly, that its evolution is due to some capacity that is not shared with other species. In section 6.1, I pointed out that the theory committed to recent syntactocentrism claims that the difference lies in the capacity for syntactic recursion, as conceived of in minimalist syntax. In section 6.2, I demonstrated that the unique feature hypothesized in Tomasello's approach is the capacity for shared intentionality, which ultimately gave rise to the emergence of language with a fully-fledged grammar. As pointed out in the final paragraphs of the last section, one aspect of shared intentionality, the pro-social motives, is questionable on methodological grounds. Looking now at the cognitive side of Tomasello's approach, although significant representational cognitive capacities, as we mentioned, have also been found in nonhuman primates, it has been argued soon that "this does not necessarily mean that chimpanzees have human-like social-cognitive skills" (Hare et al. 2001: 149) – and this view is still widely accepted (for a recent overview of cognitive capacities of chimpanzees shared and not shared with humans, see Call & Tomasello 2008). Accordingly, even if we abstract away from the motivational aspects Tomasello mentions, the cognitive basis of shared intentionality, according to this approach, may still be a human innovation, possibly sui generis. Importantly, this capacity is regarded as essential for the emergence of grammar because "the conventionalization of grammatical constructions – grammaticalization and similar processes – can occur only in species who have cognitive skills for constructing common ground" (Tomasello 2008: 307-308). Since my interest lies in comparing the conception of language evolution associated with recent syntactocentrism, which is mainly concerned with grammatical structures, with the approach of Tomasello, it may be worth focusing on this aspect involved in the development of grammar in more detail.

At some points, Tomasello (2008: 321) mentions that "[t]he basic cognitive skill of shared intentionality is recursive mindreading." That is, the cognitive capacity for creating common ground implies some kind of hierarchical embedding, since establishing a common ground between two persons requires each of them to know things

that she knows the other knows as well – and knows that the other knows this about her as well. To clarify, recall our example of Hans sitting in a bar, and let us take his perspective. Regarding the social intention of his pointing to the glass, Hans must know that the bartender knows that Hans knows that this pointing is, in this context, a sign to indicate that Hans wants another beer. Tomasello emphasizes that such embeddings cannot be replaced by a psychological primitive like 'we both know that p' because "the underlying recursive levels […] become clear when there is some kind of breakdown" (Tomasello 2008: 96). In other words, when Hans assumes he shares something with the bartender, and it turns out that Hans does not, this 'breakdown' can happen at different levels of embedding. So, for instance, if Hans says "How delicious!," it will not work, (i) if the bartender thinks Hans is attending to something Hans is not or (ii) if Hans thinks the bartender thinks Hans is attending to something Hans is not, and so on potentially ad infinitum. Of course, there can only be computations up to a certain point – but that, like in the case of recursive embedded syntax, can be regarded as a factor of memory limitations (of 'performance,' in generative terms) instead of being constrained by the underlying mechanism itself. So, the fact of breakdowns at different levels indeed suggests that we are dealing with a hierarchical embedded structure here, analogous to the kinds of structures we find in syntax. For the sake of pointing out this similarity, we could represent this embedding as follows:

(4)   [A knows [that B knows [that A knows [that p]]]]

Interestingly, similar to the concern in generative linguistics that some finite system to generate this potential infinity is needed, there are elaborated proposals in pragmatics according to which "[m]utual knowledge can […] be treated as a single mental entity instead of an infinitely long list of ever more complex mental entities" (Clark & Marshall 1981: 34). Let us take a one-sided definition of mutual knowledge, which captures the infinite embeddings by a single mental entity, a single rule (cf. Clark & Marshall 1981: 59, n. 3):

(5)   (r) A knows that p and that: B knows that p and that r.

Clearly, the loop-like device established by reintroducing 'r' is strongly reminiscent of the kind of rules postulated in mainstream generative grammar, for instance the rewrite-rule system illustrated in section 1.1. Having looked at this obvious analog, let us recall to what entity such recursive rule systems refer in the context of research on language evolution within mainstream generative linguistics. As Hauser et al. (2002: 1570) point out, "'internal language' or 'I-language' […] is the primary object of interest for the study of the evolution […] of the language faculty." When we assume an entity like 'I-language,' that is, a "system of knowledge of language attained and internally represented in the mind/brain" (Chomsky 1986b: 24), the rewrite-rule system or, in recent syntactocentrism, the operation Merge is viewed as the recursive computational device of a mental representational system that generates hierarchical structures. Importantly, Tomasello also argues for some representational format, since, in the context of generating joint intentions, he presupposes "perspectival cognitive representations […], which transformed human cognition from a mainly individual enterprise into a mainly collective cultural enterprise involving shared beliefs and practices" (Moll & Tomasello 2007: 646). As we indicated above, these cognitive representations crucially imply some mechanism of recursive mindreading, which, for reasons of memory limitations, must be modeled as finite entities (maybe in the form of rules), analogous to the finite computational system assumed for syntactic recursion. Here, a well-known and long-standing problem appears. On the one hand, we could suppose two different representational systems and posit an independent 'language of thought' and thus "commit ourselves to cognitive processes mediated by representational systems other than natural language" (Fodor 1975: 57). On the other hand, we can deny the intelligibility of any kind of thought in the absence of structures postulated for language and thus assume a single representational format accounting for both thought and language.

Returning to evolutionary issues, as for Chomsky, it is clear "that postulation of an independent or prior 'language of thought' LOT raises all the problems of evolution of language, but with the extra difficulty that we have almost no idea what LOT would be, independently of linguistic evidence" (Chomsky 2010: 226, n. 24). In other words, Chomsky argues for one representational format and presupposes an ontological identity regarding recursive processes that are involved in language and in thought. Crucially now, one can argue that Tomasello suggests, if only indirectly, this strong correlation between representations of language and thought, too. His functional

perspective implies that language, understood as the conventionalized grammatical structures, is nothing more than a device to meet the demands of human communication. These demands, as we pointed out above, are shaped by some special cognitive capacities like recursive mindreading. As a consequence, it does not contradict his theory to postulate that recursive mindreading, as it shapes human communication, is somehow reflected in the grammatical structure of language. Admittedly, this convergence of both approaches, if at all existing, is located at a very abstract level. However, even if we abstract away from these considerations, the fact that both approaches to language evolution regard recursive operations of the mind as a sine qua non for the emergence of grammar suffices to suggest some points at which these two accounts can cross-fertilize each other. With this intention, let us first turn to Tomasello's approach again.

Since Tomasello et al. hypothesize that "[a] child raised on a desert island would have all of the biological preparations for participation in interactions involving shared intentionality" (Moll & Tomasello 2007: 646), they touch on aspects of the biological side of evolutionary processes. In other words, also scholars who think that language is purely a cultural phenomenon still have to account for the evolution of human culture, and here, not only processes of cultural transmission but also mechanisms of biological evolution come into play. Given this assumption, it does not suffice, in my opinion, to argue that, in order to develop the capacity of recursive mindreading, "some early humans had to become less aggressive/competitive and more tolerant/friendly with one another" (Moll & Tomasello 2007: 646). From an experimentalist view that is committed more to standards of molecular biology or neuroscience, those "'[e]volutionarios' are entertaining but typically offer experimentalists little to work with" (Fitch 2009: 24). As we indicated in the last section, the pro-social component of becoming more 'friendly,' of helping and sharing, indeed lets some problems arise from the perspective of approaches that are committed to a Popperian methodology according to which only "a falsifiable hypothesis that repeatedly resists falsification is likely to be true" (Fitch et al. 2005: 193). Lest I will be misinterpreted here: The numerous studies of Tomasello et al. have led to insights that I do not wish to undermine in any way. However, if one assumes some biological preparation like recursive mindreading, as Tomasello et al. do, it is quite telling that even in recent work of this approach, there is not one reference to neuropsychological work that explores, for instance, whether so-called "we-centric neural representations repre-

sent the essential neural condition for collective intentional behaviour" (Becchio & Bertone 2004: 132). In light of such studies and given the more and more refined procedures for testing processes of hierarchical embedding in the context of artificial grammar, the approach of Tomasello et al., from my point of view, could profit a lot from paying attention to the biological side of the issues they are dealing with.

As for the computational view on language evolution sketched in section 6.1, I have already indicated that the further clarification of the relation between artificial grammar and natural language is crucial. What can be learned from the more communicative, 'cultural' view of Tomasello and others is that, while there is a biological preparation to process recursive embedding at all, the variety of hierarchically-embedded structures in natural language may be due to more cultural factors. So, it could be that, for instance, "writing, as an external representation of linguistic utterances, provides individuals with an extension of their memory, and thus also facilitates the recursive use of grammatical patterns" (Verhagen 2010: 108). In other words, also fundamental capacities like recursive embedding are "not themselves the underlying explanatory principles they have been taken to be within the Chomskian paradigm; instead they are facts about the structure that are to be explained as arising from the cumulative impact of the processes that shape each language" (Bybee & McClelland 2005: 406).

To conclude this section, there are some, if only rather abstract, points of convergence between approaches to language evolution committed to recent derivational syntactocentrism and accounts associated with Cognitive Linguistics. Unfortunately, however, for the most part, both approaches are steadfastly ignoring each other. Given that the problems all approaches to language evolution are facing are certainly deep ones, more collaboration and cross-fertilization regarding particular aspects, like the issue of recursion, may be of great help to cope with the conundrum all scholars concerned with language are in: How did language evolve?

7. Conclusion: Rethinking Syntactocentrism

In this thesis, I have shown that Jackendoff's notion of syntactocentrism is misguided, since characterizing the concept this term refers to as a dogmatic assumption that does not lend itself to participate in the general field of mentalist linguistics does not do justice to both recent models of mainstream generative grammar and to the conceptual consequences of applying these concepts to specific phenomena like the pragmatics of LP-Movement in German.

As a first step, chapter 1 approaches the notion of syntactocentrism by sketching its initial conception in the 1960s. In particular, I sketched the early model of the syntactic component as a system of rewrite rules and then, I illustrated in what sense both the phonological and the semantic component are regarded as 'purely interpretive.' By highlighting the main motivation to describe a single system of generating rules that can both account for the ability to produce (phonetically) and to understand (semantically) an infinite range of sentences, I clarified why even Jackendoff considers this nascent perspective on grammatical knowledge to be a quite reasonable view, given the by-then available approaches to phonology and semantics.

In chapter 2, I amended Jackendoff's claim that recent syntactocentrism ignores progress in both phonology and semantics and entirely dispenses with formal accounts of these components. To arrive at this conclusion, I first sketched the recent generative conception of syntax and then illustrated approaches to phonology and semantics that, although explicitly situated within the framework of syntactocentrism, both offer detailed formal accounts and cover crucial insights from research in phonology and semantics gained since the 1960s. In particular, I first showed the approach of cyclic prosodic mapping and thereby qualified Jackendoff's assertion that the relation between syntax and phonology is a 'no-man's-land' within recent mainstream generative grammar. After that, I outlined a recent semantic approach that does not only provide a detailed formal account of the semantic component but also adopts a cognitive view on meaning and thus concurs with Jackendoff's rejection of semantics as a purely logical enterprise.

Chapter 3 demonstrated that Jackendoff misses another crucial point in his discussions of syntactocentrism by marginalizing and sometimes even ignoring significant changes involved in the recent shift from representational to derivational syntactocentrism. In order to develop this argument, I first illustrated the representational

conception, implying an enriched conception of syntax with several levels of representation. Then, I turned to the development that has abandoned this view stepwise and sketched how prominent empirical phenomena originally motivating the postulation of D-Structure and S-Structure can be accounted for by assuming alternative descriptive tools that do not require these 'extra' levels of representation. I ended by demonstrating that even the interface levels LF and PF can be questioned within recent derivational syntactocentrism – an aspect not incorporated by Jackendoff at all.

Based on this up-to-date notion of syntactocentrism, chapter 4 showed two theoretical alternatives to syntactocentrism: Cognitive Linguistics and the Parallel Architecture. Concerning Cognitive Linguistics, I first concentrated on the construction grammar model and highlighted its rejection of the componential model of grammatical knowledge in favor of a uniform representation in terms of constructions. After that, I sketched an approach within this paradigm that claims that every aspect of language can be derived from semantics. Then, I turned to Jackendoff's own proposal, which I classified, compared to this 'semanticocentric' view, as a less extreme alternative to syntactocentrism because Jackendoff does not reject the componential model of grammar. However, since he claims that both the phonological and the semantic component must be regarded as generative on a par with syntax, I concluded that the Parallel Architecture can be regarded as an intermediate position between the syntactocentric view and Cognitive Linguistics. In the remainder of this chapter, I adopted this mediating perspective and pointed out that bridging the gulf between syntactocentrism and its theoretical alternatives is not inconceivable but a quite realistic enterprise. I finished this chapter by arguing in favor of a comparison of the syntactocentric view with its alternatives in light of the amended notion of syntactocentrism developed in chapters 2 and 3.

Chapter 5 undertakes such a comparison with respect to the Parallel Architecture and shows that, once the consequences of the recent changes within syntactocentrism are taken seriously, some points of convergence between recent syntactocentrism and the Parallel Architecture can be demonstrated. To show this, I brought down the comparison to a controllable size and concentrated on the analysis of one specific phenomenon, namely the pragmatics of LP-movement in German. To arrive at the derivational analysis of this subject, I first illustrated the strong representational view of the cartographic approach on this phenomenon, which postulates that

syntactic structures have to provide a transparent representation of interpretive aspects, including pragmatic properties associated with information structural notions like focus or contrast. After that, I pointed out both conceptual and empirical problems for cartographic accounts that assume a specific position in the German left periphery that is associated with a contrastive interpretation. Specifically, I demonstrated that the analysis contradicts some general concepts of minimalist syntax and cannot account for certain empirical cases of so-called pars-pro-toto-movement in German. Having sketched these problems, I outlined an analysis of LP-movement in German that adopts the phase-based approach within recent syntactocentrism and assumes the concept of cyclic linearization, which captures the linear ordering of words by referring to different cycles, different domains of Spell-Out. I showed that this analysis abandons cartographic, representational devices and claims that information structural properties play no role in syntactic operations and are only linked to freely assigned accents, hence, they only interface with phonology. Based on this recent analysis, I argued that the strong derivational view of syntactocentrism implies significant points of convergence with conceptual aspects of the Parallel Architecture. First, both approaches share the assumption that a direct interaction between phonology and semantic interpretation is required in order to deal with phenomena like prosodically-expressed focus or contrast. Second, both models imply that this interaction is established by pragmatic rather than by syntactic factors and, third, both accounts regard pragmatics as an independent component, which is described in the Parallel Architecture and omitted in the descriptive apparatus of mainstream generative linguistics but not denied as a component crucial for human language. I concluded by hypothesizing that, within recent syntactocentrism, a crucial criterion for determining what belongs to 'pure syntax' and what can be excluded may be the property of showing recursive embedding, which can be argued to be absent in information structural notions.

In order to look also for convergence between recent syntactocentrism and Cognitive Linguistics, chapter 6 compared an approach to language evolution that is based on the general conception of syntax exemplified in chapter 5 with an account that is associated with concepts of Cognitive Linguistics. To make this comparison, I first presented a computational view according to which the crucial factor in the evolution of human language is the capacity for syntactic recursion, as conceived of in minimalist syntax. I argued that this approach has led to deepen both the theoretical

notion of recursion in language and the methodological discussion concerned with creating adequate experimental designs to test the processing of hierarchically-embedded structures. After that, I illustrated an approach that is committed to the communicative view on language evolution and rests on concepts of Cognitive Linguistics. I showed that the unique feature that is hypothesized in this approach is the capacity for shared intentionality, which implies both the generation of joint intentions and the existence of pro-social motives. I then argued that these pro-social motives are questionable on methodological grounds, inasmuch as there seem to be several problems to define motivation in an experimentally falsifiable way. Given these problems, I abstracted away from these motivational aspects and focused on the cognitive component of generating joint intentions. Having shown that this component involves the capacity for recursive mindreading, I highlighted that both the computational and the communicative view regard recursive operations of the mind as a sine qua non for the emergence of grammar and that both accounts postulate a representational format that contains these recursive operations. Based on this convergence, I suggested some points at which these two accounts could cross-fertilize each other. As for the communicative view, I argued that this approach could profit from paying attention to biological approaches to shared intentionality and from looking at the refined procedures for testing processes of hierarchical embedding within research on syntactic recursion. As for the computational view on language evolution, I pointed out that further clarification of the relation between the stimuli used in experimental studies and natural language is crucial and that at least some aspects of hierarchically-embedded structures in natural language may find better explanation by taking into account more cultural factors.

To conclude, although I demonstrated that Jackendoff misses some significant aspects of recent mainstream generative grammar and thus shifted all responsibility on him, it might be that some confusion is caused by generative linguistics itself, since, as we saw in chapter 5, in the context of analyzing specific language data, both the representational and the more recent derivational view still exist side by side. Regarding interdisciplinary alignment, however, the more recent concept seems more promising, as ongoing research on language evolution demonstrates. Moreover, as this thesis indicates, the derivational view also lends itself better to convergence, at least in some aspects, with other branches of mentalist linguistics. Given these observations and given the general conclusion that the broad-brush

characterization of recent mainstream generative grammar as a fixed ideology, as an 'ism,' has proven to be misguided, the reader who accepts the arguments made in this thesis may "throw away the ladder, after he has climbed on it" (Wittgenstein 1961 [1921]: 189) and thus abandon the notion of syntactocentrism altogether, since labeling competing approaches as 'isms' may impede "to fight fair in the interests of deeper understanding" (Jackendoff 2002: xiii).

References

Austin, John L. (1962). *How to Do Things with Words*. Oxford: Oxford University Press.

Bahlmann, Jörg, Ricarda I. Schubotz & Angela D. Friederici (2008). Hierarchical artificial grammar processing engages Broca's area. *NeuroImage* 42, 525-534.

Baker, Mark C. (2003). *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.

Bar-Hillel, Yehoshua (1953). On recursive definitions in empirical science. *11th International Congress of Philosophy* 5, 160-165.

Becchio, Cristina & Cesare Bertone (2004). Wittgenstein running: Neural mechanisms of collective intentionality and we-mode. *Consciousness and Cognition* 13, 123-133.

Bhatt, Rakesh M. (1999). *Verb Movement and the Syntax of Kashmiri*. Dordrecht: Kluwer.

Boeckx, Cedric (2006). *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. Oxford: Oxford University Press.

Boeckx, Cedric (2008). *Bare Syntax*. Oxford: Oxford University Press.

Boeckx, Cedric & Massimo Piattelli-Palmarini (2007). Linguistics in cognitive science: The state of the art amended. *The Linguistic Review* 24, 403-415.

Büring, Daniel (2007). Semantics, intonation, and information structure. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford Handbook of Linguistics Interfaces*, 445-473. Oxford: Oxford University Press.

Büring, Daniel (2008). Been there, marked that: A theory of second occurrence focus. Ms., University of California, Los Angeles <http://www.linguistics.ucla.edu/people/buring> 2010-06-10.

Bybee, Joan L. (1988). The diachronic dimension in explanation. In John A. Hawkins (ed.), *Explaining Language Universals*, 350-379. Oxford: Blackwell.

Bybee, Joan L. & James L. McClelland (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review* 22, 381-410.

Call, Josep & Michael Tomasello (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences* 12, 187-192.

Castillo, Juan C., John Drury & Kleanthes K. Grohmann (1999). Merge over move and the extended projection principle. *University of Maryland Working Papers in Linguistics* 8, 63-103.

Chafe, Wallace L. (1968). Idiomaticity as an anomaly in the Chomskyan paradigm. *Foundations of Language* 4, 109-127.

Chomsky, Noam (1957). *Syntactic Structures*. 'S-Gravenhage: Mouton.

Chomsky, Noam (1964). *Current Issues in Linguistic Theory*. The Hague: Mouton.

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.

Chomsky, Noam (1966). *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. New York & London: Harper & Row.

Chomsky, Noam (1970a). Deep structure, surface structure, and semantic interpretation. In Roman Jakobson & Shigeo Kawamoto (eds.), *Studies in General and Oriental Linguistics Presented to Shirô Hattori on the Occasion of his Sixtieth Birthday*, 52-91. Tokyo: TEC Company.

Chomsky, Noam (1970b). Remarks on nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, 184-221. Waltham, Mass.: Ginn.

Chomsky, Noam (1973). Conditions on transformations. In Stephen R. Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 232-286. New York: Holt, Rinehart and Winston.

Chomsky, Noam (1975) [1955]. *The Logical Structure of Linguistic Theory*. New York & London: Plenum.

Chomsky, Noam (1975). *Reflections on Language*. New York: Pantheon.

Chomsky, Noam (1980). *Rules and Representations*. New York: Columbia University Press.

Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam (1986a). *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, Noam (1986b). *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.

Chomsky, Noam (1991). Some notes on economy of derivation and representation. In Robert Freidin (ed.), *Principles and Parameters in Generative Grammar*, 417-454. Cambridge, Mass.: MIT Press.

Chomsky, Noam (1993). A minimalist program for linguistic theory. In Kenneth Hale & Samuel J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 1-52. Cambridge, Mass.: MIT Press.

Chomsky, Noam (1995). The Minimalist Program. Cambridge, Mass.: MIT Press.

Chomsky, Noam (2000). Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89-155. Cambridge, Mass.: MIT Press.

Chomsky, Noam (2004). Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and Beyond*, 104-131. Oxford: Oxford University Press.

Chomsky, Noam (2005). Three factors in language design. *Linguistic Inquiry* 36, 1-22.

Chomsky, Noam (2008). On phases. In Robert Freidin, Carlos P. Otero & Maria L. Zubizarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, 133-166. Cambridge, Mass.: MIT Press.

Chomsky, Noam (2010). Some simple evo devo theses: How true might they be for language? In Richard K. Larson, Viviane Déprez & Hiroko Yamakido (eds.), *The Evolution of Human Language: Biolinguistic Perspectives*, 45-62. Cambridge: Cambridge University Press.

Chomsky, Noam, Morris Halle & Fred Lukoff (1956). On accent and juncture in English. In Morris Halle, Horace G. Lunt, Hugh McLean & Cornelis H. van Schooneveld (eds.), *For Roman Jakobson: Essays on the Occasion of his Sixtieth Birthday*, 65-80. The Hague: Mouton.

Chomsky, Noam & Morris Halle (1968). *The Sound Pattern of English*. New York: Harper & Row.

Chomsky, Noam & Howard Lasnik (1977). Filters and control. *Linguistic Inquiry* 8, 425-504.

Chomsky, Noam & Howard Lasnik (1993): The theory of principles and parameters. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax: An International Handbook of Contemporary Research*, vol. 1, 506-569. Berlin & New York: de Gruyter.

Christiansen, Morten H. & Simon Kirby (2003). Language evolution: The hardest problem in science? In Morten H. Christiansen & Simon Kirby (eds.), *Language Evolution*, 1-15. Oxford: Oxford University Press.

Chung, Sandra & William A. Ladusaw (2004). *Restriction and Saturation*. Cambridge, Mass.: MIT Press.

Cinque, Guglielmo & Luigi Rizzi (2010). The cartography of syntactic structures. In Bernd Heine & Heiko Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*, 51-65. Oxford: Oxford University Press.

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

Clark, Herbert H. & Catherine R. Marshall (1981). Definite reference and mutual knowledge. In Aravind K. Joshi, Bonnie L. Webber & Ivan A. Sag (eds.), *Elements of Discourse Understanding*, 10-63. Cambridge: Cambridge University Press.

Croft, William (2007). Construction grammar. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*, 463-508. Oxford: Oxford University Press.

Croft, William & D. Alan Cruse (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Culicover, Peter W. & Ray Jackendoff (2005). *Simpler Syntax*. Oxford: Oxford University Press.

Davidson, Donald (1967). The logical form of action sentences. In Nicholas Rescher (ed.), *The Logic of Decision and Action*, 81-120. Pittsburgh: Pittsburgh University Press.

Descartes, Réne (2003) [1637]. Discourse on method. In Elizabeth S. Haldane & George R. Thomson (trans.), *Discourse on Method and Meditations* [Dover Philosophical Classics], 1-52. Mineola, N.Y.: Dover.

Elordieta, Gorka (2007). Segmental phonology and syntactic structure. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford Handbook of Linguistics Interfaces*, 125-177. Oxford: Oxford University Press.

Fanselow, Gisbert (2003). Surprising specifiers and cyclic spellout. *Generative Linguistics in Poland* 5, 29-46.

Fanselow, Gisbert (2004). Cyclic phonology-syntax-interaction: Movement to first position in German. In Shinichiro Ishihara, Michaela Schmitz & Anne Schwarz (eds.), *Interdisciplinary Studies on Information Structure*, 1-42. Potsdam: Potsdam University Press.

Fanselow, Gisbert (2006). On pure syntax. In Patrick Brandt & Eric Fuss (eds.), *Form, Structure and Grammar*, 137-157. Berlin: Akademie Verlag.

Fanselow, Gisbert (2008). In need of mediation: The relation between syntax and information structure. *Acta Linguistica Hungarica* 55, 397-413.

Fanselow, Gisbert & Denisa Lenertová (2008). Left peripheral focus: Mismatches between syntax and information structure. Ms., University of Potsdam. <http://www.ling.uni-potsdam.de/~fanselow/files/LeftPeripheralFocus2.pdf> 2009-08-01.

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

Fitch, W. Tecumseh (2009). Prolegomena to a science of biolinguistics. In Louise S. Röska-Hardy & Eva M. Neumann-Held (eds.), *Learning from Animals? Examining the Nature of Human Uniqueness*, 15-44. Hove & New York: Psychology Press.

Fitch, W. Tecumseh (2010). Three meanings of 'recursion': Key distinctions for biolinguistics. In Richard K. Larson, Viviane Déprez & Hiroko Yamakido (eds.), *The Evolution of Human Language: Biolinguistic Perspectives*, 73-90. Cambridge: Cambridge University Press.

Fitch, W. Tecumseh & Marc D. Hauser (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science* 303, 377-380.

Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky (2005). The evolution of the language faculty: Clarifications and implications. *Cognition* 97, 179-210.

Fodor, Jerry A. (1975). *The Language of Thought*. Cambridge, Mass.: Harvard University Press.

Fox, Danny & David Pesetsky (2005). Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31, 1-45.

Frege, Gottlob (1980) [1914]. Letter to Jourdain. In Gottfried Gabriel (ed.), *Philosophical and Mathematical Correspondence*, 78-80. Chicago: Chicago University Press.

Frey, Werner (2004a). The grammar-pragmatics interface and the German prefield. *Sprache & Pragmatik* 52, 1-39.

Frey, Werner (2004b). A medial topic position for German. *Linguistische Berichte* 198, 153-190.

Frey, Werner (2006). Contrast and movement to the German prefield. In Valéria Molnár & Susanne Winkler (eds.), *The Architecture of Focus*, 235-264. Berlin & New York: de Gruyter.

Frey, Werner (2010). Ā-Movement and conventional implicatures: About the grammatical encoding of emphasis in German. *Lingua* 120, 1416-1435.

Friederici, Angela D., Jörg Bahlmann, Stefan Heim, Ricarda I. Schubotz & Alfred Anwander (2006). The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *PNAS* 103, 2458-2463.

Geeraerts, Dirk & Hubert Cuyckens (2007). Introducing cognitive linguistics. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*, 3-21. Oxford: Oxford University Press.

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

Goldberg, Adele E. (1996). Jackendoff and construction-based grammar. *Cognitive Linguistics* 7, 3-19.

Goldberg, Adele E. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Haegeman, Liliane (1991). *Introduction to Government and Binding Theory*. Oxford: Blackwell.

Halle, Morris & Jean-Roger Vergnaud (1987). *An Essay on Stress*. Cambridge, Mass.: MIT Press.

Halliday, Michael A. K. (1967). Notes on transitivity and theme in English: Part 2. *Journal of Linguistics* 3, 199-244.

Hare, Brian, Josep Call, Bryan Agnetta & Michael Tomasello (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour* 59, 771-785.

Hare, Brian, Josep Call & Michael Tomasello (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour* 61, 139-151.

Harris, Randy A. (1993). *The Linguistics Wars*. Oxford: Oxford University Press.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569-1579.

Hauser, Marc D., David Barner & Tim O'Donnell (2007). Evolutionary linguistics: A new look at an old landscape. *Language Learning & Development* 3, 101-132.

Hornstein, Norbert, Jairo Nunes & Kleanthes K. Grohmann (2005). *Understanding Minimalism*. Cambridge: Cambridge University Press.

Hornstein, Norbert & Paul M. Pietroski (2009). Basic operations: Minimal syntax-semantics. *Catalan Journal of Linguistics* 8, 113-139.

Horvath, Julia (2007). Separating 'focus movement' from focus. In Simin Karimi, Vida Samiian & Wendy K. Wilkins (eds.), *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation*, 108-145. Amsterdam & Philadelphia: John Benjamins.

Humboldt, Wilhelm von (1999) [1836]. The diversity of human language-structure and its influence on the mental development of mankind. In Peter Heath (trans.), *Wilhelm von Humboldt: On Language*, 1-287. Cambridge: Cambridge University Press.

Jackendoff, Ray (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, Mass.: MIT Press.

Jackendoff, Ray (1983). *Semantics and Cognition*. Cambridge, Mass.: MIT Press.

Jackendoff, Ray (1995). The boundaries of the lexicon. In Martin Everaert, Erik-Jan van der Linden, André Schenk & Rob Schreuder (eds.), *Idioms: Structural and Psychological Perspectives*, 133-165. Hillsdale, N.J.: Lawrence Erlbaum.

Jackendoff, Ray (1996). Conceptual semantics and cognitive linguistics. *Cognitive Linguistics* 7, 93-129.

Jackendoff, Ray (1997). *The Architecture of the Language Faculty*. Cambridge, Mass.: MIT Press.

Jackendoff, Ray (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jackendoff, Ray (2003). Précis of "Foundations of Language: Brain, Meaning, Grammar, Evolution." *Behavioral and Brain Sciences* 26, 651-665.

Jackendoff, Ray (2007). *Language, Consciousness, Culture: Essays on Mental Structure*. Cambridge, Mass.: MIT Press.

Jackendoff, Ray (2010). *Meaning and the Lexicon: The Parallel Architecture 1975-2010*. Oxford: Oxford University Press.

Janssen, Theo M. V. (1997). Compositionality. In Johan van Benthem & Alice ter Meulen (eds.), *Handbook of Logic and Language*, 417-473. Amsterdam: Elsevier.

Katz, Jerrold J. & Jerry A. Fodor (1963). The structure of a semantic theory. *Language* 39, 170-210.

Katz, Jerrold J. & Paul Postal (1964). *An Integrated Theory of Linguistic Descriptions*. Cambridge, Mass.: MIT Press.

Kayne, Richard S. (1994). *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.

Kuroda, Shige-yuki (2005). Focusing on the matter of topic: A study of *wa* and *ga* in Japanese. *Journal of East Asian Linguistics* 14, 1-58.

Ladd, D. Robert (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

Langacker, Ronald W. (1987). *Foundations of Cognitive Grammar – Vol. 1: Theoretical Prerequisites*. Stanford, Cal.: Stanford University Press.

Lasnik, Howard (2000). *Syntactic Structures Revisited*. Cambridge, Mass.: MIT Press.

Lasnik, Howard (2005). Grammar, levels, and biology. In James McGilvray (ed.), *The Cambridge Companion to Chomsky*, 60-83. Cambridge: Cambridge University Press.

Lasnik, Howard, Juan Uriagereka & Cedric Boeckx (2005). *A Course in Minimalist Syntax*. Oxford: Blackwell.

Lyons, Derek E., Webb Phillips & Laurie R. Santos (2005). Motivation is not enough. *Behavioral and Brain Sciences* 28, 708.

Maienborn, Claudia (1996). *Situation und Lokation: Die Bedeutung lokaler Adjunkte von Verbalprojektionen*. Tübingen: Stauffenburg.

May, Robert (1977). The grammar of quantification. Doctoral dissertation, MIT.

McClelland, James L. & Joan L. Bybee (2007). Gradience of gradience: A reply to Jackendoff. *The Linguistic Review* 24, 437-455.

Moll, Henrike & Michael Tomasello (2007). Cooperation and human cognition: The Vygotskian intelligence hypothesis. *Phil. Trans. R. Soc. B* 362, 639-648.

Müller, Gereon (2007). Towards a relativized concept of cyclic linearization. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*, 61-114. Berlin & New York: de Gruyter.

Muysken, Pieter (1982). Parametrizing the notion 'head.' *Journal of Linguistic Research* 2, 57-75.

Neeleman, Ad & Kriszta Szendrői (2004). Superman sentences. *Linguistic Inquiry* 35, 149-159.

Newman, Stanley S. (1946). On the stress system of English. *Word* 2, 171-187.

Perruchet, Pierre & Arnaud Rey (2005). Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? *Psychonomic Bulletin & Review* 12, 307-313.

Pietroski, Paul M. (2007). Systematicity via monadicity. *Croatian Journal of Philosophy* 7, 343-374.

Pietroski, Paul M. (2008a). Minimal semantic instructions. Ms., University of Maryland, College Park. <http://terpconnect.umd.edu/~pietro/research/papers/MinimalInstructionsFinal.pdf> 2010-01-29.

Pietroski, Paul M. (2008b). Minimalist meaning, internalist interpretation. *Biolinguistics* 2, 317-341.

Pijper, Jan R. de & Angelien A. Sanderman (1994). On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *The Journal of the Acoustical Society of America* 96, 2037-2047.

Pinker, Steven & Paul Bloom (1990). Natural language and natural selection. *Behavioral and Brain Sciences* 13, 707-727.

Pollock, Jean-Yves (1989). Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20, 365-424.

Raible, Wolfgang (2009). Zur Realität von Tiefenstrukturen. In Angelika Linke & Helmuth Feilke (eds.), *Oberfläche und Performanz: Untersuchungen zur Sprache als dynamischer Gestalt*, 77-97. Tübingen: Niemeyer.

Ritter, Nancy A. (2005). On the status of linguistics as a cognitive science. *The Linguistic Review* 22, 117-133.

Rizzi, Luigi (1997). The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of Grammar: Handbook in Generative Syntax*, 281-337. Dordrecht: Kluwer.

Rizzi, Luigi (2004). On the cartography of syntactic structures. In Luigi Rizzi (ed.), *The Structure of CP and IP*, 3-15. Oxford: Oxford University Press.

Rosenbaum, Peter (1967). *The Grammar of English Predicate Complement Constructions*. Cambridge, Mass.: MIT Press.

Sauerland, Uli & Arnim von Stechow (2001). The syntax-semantics interface. In Neil J. Smelser & Paul B. Baltes (eds.), *International Encyclopedia of the Social & Behavioural Sciences*, 15412-15418. Oxford: Pergamon Press.

Schwarz, Monika (1996). *Einführung in die Kognitive Linguistik*, 2nd edn. Tübingen & Basel: Francke.

Searle, John R. (1990). Collective intentions and actions. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in Communication*, 401-415. Cambridge, Mass.: MIT Press.

Selkirk, Elisabeth (1986). On derived domains in sentence phonology. *Phonology Yearbook* 3, 371-405.

Shannon, Claude E & Warren Weaver (1949). *The Mathematical Theory of Communication*. Urbana, Ill.: Illinois University Press.

Southgate, Victoria, Catharine van Maanen & Gergely Csibra (2007). Infant pointing: Communication to cooperate or communication to learn? *Child Development* 78, 735-740.

Thiersch, Craig (1978). Topics in German syntax. Doctoral dissertation, MIT.

Tomasello, Michael (2008). *Origins of Human Communication*. Cambridge, Mass.: MIT Press.

Tomasello, Michael & Josep Call (1997). *Primate Cognition*. Oxford: Oxford University Press.

Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Behne & Henrike Moll (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28, 675-691.

Tomasello, Michael, Malinda Carpenter & Ulf Liszkowski (2007). A new look at infant pointing. *Child Development* 78, 705-722.

Tomioka, Satoshi (2007). Information structure as information-based partition. In Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds.), *Interdisciplinary Studies on Information Structure* 6, 97-107. Potsdam: Potsdam University Press.

Trotzke, Andreas (2008). Rekursive Syntax & FLN? Eine kritische Betrachtung neuerer Diskussionen zur menschlichen Sprachfähigkeit innerhalb des Generativismus. Master's Thesis, University of Freiburg. <http://www.freidok.uni-freiburg.de/volltexte/7187> 2010-08-01.

Truckenbrodt, Hubert (2007). The syntax-phonology interface. In Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*, 435-456. Cambridge: Cambridge University Press.

Vallduví, Enric & Maria Vilkuna (1998). On rheme and kontrast. In Peter W. Culi-cover & Louise McNally (eds.), *The Limits of Syntax*, 79-108. New York: Academic Press.

van der Hulst, Harry (2006). On the parallel organization of linguistic components. *Lingua* 116, 657-688.

Verhagen, Arie (2010). What do you think is the proper place of recursion? Conceptual and empirical issues. In Harry van der Hulst (ed.), *Recursion and Human Language*, 93-110. Berlin & New York: de Gruyter.

Wagner, Michael (2005). Prosody and recursion. Doctoral dissertation, MIT.

Wagner, Michael (2010). Prosody and recursion in coordinate structures and beyond. *Natural Language & Linguistic Theory* 28, 183-237.

Winkler, Susanne (2005). *Ellipsis and Focus in Generative Grammar*. Berlin & New York: de Gruyter.

Wittgenstein, Ludwig (1961) [1921]. *Tractatus Logico-Philosophicus*. Trans. David F. Pears and Brian F. McGuinness. London: Routledge & Kegan Paul.

Zubizarreta, Maria L. (1998). *Prosody, Focus and Word Order*. Cambridge, Mass.: MIT Press.

Zusammenfassung in deutscher Sprache

Der Linguist Ray Jackendoff hat zur Bezeichnung der grundlegenden Annahme der generativen Linguistik, die Syntax nehme die zentrale Stellung innerhalb der mentalen Architektur der Sprachfähigkeit ein, den Begriff ‚Syntaktozentrimus' eingeführt. In der vorliegenden Arbeit habe ich aufgezeigt, dass Jackendoffs Diskussion dieser Annahme sowohl hinsichtlich neuerer Modelle der generativen Grammatik als auch bezüglich der Anwendung dieser Modelle auf spezielle sprachliche Phänomene fehlgeleitet ist.

Hierzu sind in Kapitel 1 zunächst die grundlegenden Annahmen des syntaktozentrischen Ansatzes eingeführt worden, indem anhand früher Modelle der generativen Linguistik verdeutlicht worden ist, in welchem Sinne phonologische sowie semantische Aspekte von Sprache als Resultate der Interpretation syntaktischer Strukturen aufgefasst werden.

In Kapitel 2 habe ich sodann Jackendoffs Behauptung relativiert, neuere Ansätze innerhalb des Syntaktozentrismus ignorierten Fortschritte in den Feldern der Phonologie und Semantik und verzichteten auf jegliche formale Beschreibung von Laut- und Bedeutungsstrukturen. Hierzu habe ich zunächst die neuere Syntaxkonzeption innerhalb des Minimalistischen Programms skizziert und anschließend hiermit verbundene phonologische sowie semantische Theorien veranschaulicht, die wesentliche Argumente Jackendoffs entkräften.

In Kapitel 3 habe ich herausgestellt, dass in Jackendoffs Darstellungen der syntaktozentrischen Konzeption wichtige Aspekte der neueren Entwicklung der generativen Grammatik von einem repräsentationellen hin zu einem derivationellen Ansatz ausgespart werden. Hierzu ist skizziert worden, wie das repräsentationelle Modell einer angereicherten Syntax mit mehreren Repräsentationsebenen sukzessive abgeschafft und durch alternative Beschreibungsmittel im Rahmen des Minimalismus ersetzt worden ist.

Nachdem ich auf diesem Wege die neuere Konzeption innerhalb des Syntaktozentrismus herausgearbeitet habe, sind in Kapitel 4 zwei theoretische Alternativen zu diesem Ansatz aufgezeigt worden. Zunächst habe ich ein Modell innerhalb des

Paradigmas der Kognitiven Linguistik[21] dargestellt, das die komponentielle Sicht der generativen Linguistik durch das Repräsentationsformat der ‚Konstruktion' ersetzt. Anschließend habe ich Jackendoffs Theorie der Parallelarchitektur skizziert, die – ebenso wie der generative Ansatz – eine komponentielle Konzeption der Grammatik annimmt. Da Jackendoff indes eine von der Syntax unabhängige Phonologie sowie Semantik postuliert, habe ich geschlussfolgert, dass seine Position eine zwischen Kognitiver Linguistik und Syntaktozentrismus vermittelnde Sichtweise darstellt. Diese Perspektive aufgreifend, habe ich abschließend für einen Vergleich der syntaktozentrischen Sicht mit den genannten theoretischen Alternativen argumentiert, der die in Kapitel 2 und 3 herausgestellten neueren Konzeptionen berücksichtigt.

In Kapitel 5 habe ich sodann einen solchen Vergleich zwischen dem in Kapitel 3 eingeführten neueren derivationellen Ansatz und der Parallelarchitektur vorgenommen und mich hierbei, um ein von Jackendoff angeführtes Argument detailliert zu behandeln, auf ein spezielles Phänomen – die Pragmatik von Bewegungen in die linke Satzperipherie des Deutschen – beschränkt. Der Vergleich der derivationellen Analyse dieses Phänomens mit den diesbezüglichen Annahmen der Parallelarchitektur hat ergeben, dass beide Ansätze – anders als von Jackendoff behauptet – grundsätzliche Gemeinsamkeiten, in diesem Fall hinsichtlich der Beziehung zwischen Syntax und Pragmatik, aufweisen.

Um abschließend die Frage nach Annäherungspunkten zwischen Syntaktozentrismus und Kognitiver Linguistik zu behandeln, habe ich in Kapitel 6 einen auf neueren Konzepten der generativen Syntax beruhenden Ansatz zur Sprachevolution mit einer Theorie verglichen, die mithilfe grundlegender Konzepte der Kognitiven Linguistik die evolutionäre Entwicklung der Sprachfähigkeit beschreibt. Hierbei habe ich zunächst herausgestellt, dass beide Konzeptionen rekursive Operationen innerhalb einer repräsentationellen Theorie des Geistes annehmen, und sodann aufgezeigt, in welchen Punkten beide Ansätze voneinander profitieren könnten.

In Kapitel 7 sind die Resultate dieser Arbeit noch einmal zusammengefasst und eine abschließende Bewertung bezüglich Jackendoffs Diskussion des syntaktozentrischen Modells formuliert worden.

---

[21] Im Deutschen wird der Begriff ‚Kognitive Linguistik' zuweilen auch für die generative Linguistik verwendet (vgl. etwa Schwarz 1996); in der vorliegenden Zusammenfassung ist er jedoch als Bezeichnung derjenigen nicht-generativen Ansätze zu verstehen, die im englischen Sprachraum unter ‚Cognitive Linguistics' subsumiert werden und die ich in Abschnitt 4.1 dieser Arbeit behandelt habe.