Prosodic Syntax

Joost Kremers

Voor mijn vader

Contents

Introduction v

1 An architecture for grammar 1
- 1.1 Some remarks on the purpose of the generative enterprise 2
  - 1.1.1 The linguistic divide 3
  - 1.1.2 Symbolic vs. subsymbolic models 4
- 1.2 The primitives of linguistic structure 11
  - 1.2.1 A Bantu perspective on the German nominal system 13
  - 1.2.2 The status of features 19
  - 1.2.3 Representation of features and constraints 25
  - 1.2.4 The past tense debate 28
- 1.3 Deacon's hierarchy of reference 38
- 1.4 A parallel architecture 46

2 Hierarchical and linear structure 49
- 2.1 Introduction 49
- 2.2 The Exclusivity and Nontangling Conditions 50
- 2.3 The Separation Hypothesis 53
- 2.4 Autosegmental phonology 55
- 2.5 Bare phrase structure 56
- 2.6 Valued features 59
- 2.7 Phonological composition 61
  - 2.7.1 Arabic verbal nouns 61
  - 2.7.2 Mapping and linearisation 69
- 2.8 Concluding remarks 80

3 Phonology-syntax interaction 81
- 3.1 Simultaneity 81
  - 3.1.1 Negation 82
  - 3.1.2 Adverbials 85
  - 3.1.3 Interrogative marking 91
  - 3.1.4 Topics 92
  - 3.1.5 Simultaneity in spoken languages 94
  - 3.1.6 Concluding remarks 95

3.2 Head movement . . . 98
3.2.1 Introduction . . . 98
3.2.2 Motivating movement . . . 100
3.2.3 Morphologically motivated head movement . . . 102
3.2.4 An alternative analysis . . . 103
3.2.5 Head placement . . . 104
3.2.6 Concluding remarks . . . 111
3.3 Affixes and clitics . . . 112
3.3.1 Introduction . . . 112
3.3.2 Affixes . . . 112
3.3.3 Clitics . . . 116
3.3.4 Affix and clitic types . . . 130
3.4 Concluding remarks . . . 132

4 Syntax vs. morphology **133**
4.1 Introduction . . . 133
4.2 General aims . . . 135
4.3 Syntax-morphology synthesis . . . 137
4.3.1 Form-meaning pairing . . . 138
4.3.2 Distinguishing structures . . . 139
4.3.3 Making sense of the intuition . . . 143
4.4 Data . . . 147
4.5 Counterarguments . . . 151
4.5.1 Syntactic vs. morphological word formation . . . 152
4.5.2 Lack of morphological movement . . . 154
4.5.3 Lexical Integrity . . . 156
4.6 Morphology by itself . . . 162
4.6.1 Introduction . . . 162
4.6.2 The Romance N-pattern . . . 163
4.6.3 English past participles . . . 168
4.6.4 Templatic morphology . . . 175
4.7 Concluding remarks . . . 190

5 Competition-based interaction **193**
5.1 The Elsewhere Condition . . . 193
5.2 Serbo-Croatian clitics reprised . . . 203
5.3 Scrambling in German . . . 207
5.3.1 Stress placement . . . 208
5.3.2 The order of constituents . . . 213
5.4 Concluding remarks . . . 224

6 Summary and conclusions 225

Introduction

The present study has two main goals. The first of these is linguistic, while the second is more of a programmatic nature. The linguistic goal is to put forward and argue for the idea that prosodic morphology, first proposed by McCarthy (1981) and McCarthy and Prince (1995, 1996) and based on Goldsmith's (1976, 1990) model of autosegmental phonology, should be extended to syntax. By doing so, I believe we can achieve certain goals of theoretical linguistics that have so far remained elusive. These goals are less need to rely on "invisible things", specifically phonologically 'null' heads and string-vacuous movements, a clearer view on the relation between syntax and morphology, and a formal way of describing interactions between syntax and phonology that shape linguistic forms. It is explicitly not my claim that this study achieves all these goals, but I hope to bring us at least a little bit closer to them.

The present study has a background in the minimalist framework that will be obvious to even the casual reader. However, I believe that the basic ideas of what I am arguing for are not dependent on any particular syntax framework. Furthermore, I will argue that certain aspects of minimalist theories are unlikely to be correct and will try to refrain from using them as fundamental points in the argumentation. In this manner, I hope that the present work will remain compatible with non-minimalist frameworks.

Two aspects of minimalist frameworks are particularly problematic in my view and will have to be given up. The first of these is the idea that the syntax of any natural language is based on an innate set of features. As I discuss in chapter 1, such a set of features is unlikely to have arisen through normal processes of evolution and is therefore unlikely. To the hardcore minimalist, this assumption should not even be all that disturbing, since it appears that Chomsky's latest views on UG do not necessarily include an innate set of features at all, if we follow Hauser, Chomsky, and Fitch (2002) suggestion that the Faculty of Language in the Narrow sense (FLN) is indeed restricted to just recursion.

Rather than an innate set of features, syntax, I believe, operates with a set of features that have been extracted from the input during language acquisition. This means that different languages may (and in fact do) have different feature sets. Specifically, a language such as German has gender features on nouns, but a language such as English, in which gender features are morphologically absent,

does not. On the other hand, I do not wish to claim that only those features that are morphologically visible are viable. English likewise does not have overt case morphology (save in the pronoun system, to some extent), but there is clearly a syntactic distinction between subjects and objects, which should constitute sufficient grounds for assuming a formal feature that we may call 'abstract Case'.

Crucially, features, even if not innate, are nonetheless *formal*, in the sense that they are not directly linked to specific functional or semantic properties. Nominative case is the feature that marks the syntactic subject, while accusative (or oblique) case marks the syntactic object. It does not follow, however, that elements so marked have specific functions or semantic roles. It is true that subjects tend to correspond to agents and objects to patients, but these are just tendencies. (Actually, the link is stronger in the other direction: agents are generally subjects.) This is an obvious truth that few would dispute, but it has an important consequence: there may be features in syntax that cannot be linked to semantics, not even in an indirect way.

This leads to the second aspect that I think should be given up: we must allow a certain amount of arbitrariness in syntax. This, too, should not be too much of an issue, since I follow Chomsky (1995b) in the assumption that the arbitrariness is mediated by the lexicon. The difference is that I take this principle a step further: the syntactic structure-building component is not only responsible for constructing a syntactic tree. It also constructs the morphological structure of the elements involved. More precisely, I do not believe there is a fundamental difference between syntactic and morphological structures. Both are the result of the same structure-building operation, which Chomsky has termed *merge* and which in its most basic form amounts to 'take two elements and construct a new, larger element out of them'.

Some methodological remarks are in order here. My aim in the analysis is to remain as conservative as possible. I especially aim to refrain from assuming string-vacuous movement operations and phonologically empty heads. This, I believe, is a necessary strategy when one accepts the assumption that features cannot be innate. If features cannot be innate, there can also not be an innate clause structure. If clause structure is not innate, a language-learning child will have no reason to assume invisible movements and invisible heads. Granted, there may be situations in which such "invisible things" can be plausibly argued for (in fact, I will do so myself on one or two occasions), but they should remain the exception rather than the rule. A head that is invisible (i.e., has no phonological form), should at least be visible in *some* cases. As an example, I argue in section 3.2.5 that Arabic can have an empty D° head in certain circumstances, but since Arabic has an overt definite article *al*, it does not appear to be implausible that a language-learner assumes it can be empty (or, more specifically, instanti-

ated by a moved bare N°) in certain cases.

The case for movement is essentially the same: if a (type of) element can appear overtly in more than one position, it is plausible to assume that it can "move". Such cases are generally restricted to those cases where non-transformational syntax frameworks assume *long-distance dependencies*. Put differently, if an element appears in one position but is (also) understood to have some syntactic function in another position, and if additionally there are cases where a functionally similar element actually occupies this other position, then it is safe to assume that the element in question is somehow linked to both positions. The typical case is *wh* movement: although a language such as English does not allow *wh in situ* (except for echo questions), a *wh* object can safely be assumed to have a link to a post-verbal position, because that is the position that non-*wh* objects occupy. In minimalism, and in the present study, this link is called "movement", but I believe this is just a label. My aim is to refrain from using movement as much as possible in the analysis, limiting my uses of it to cases that are fairly uncontroversial and to those where assuming movement seems inevitable.

The syntactic model to be proposed in the next chapters will rely on the phonological system to a much greater extent than is usual in syntactic theory. The entire burden of linearisation, affixation and related processes is put squarely in the phonology. Crucially, and interestingly, however, this does not mean that phonology needs to be much richer than it is. I explicitly refrain from adding any properties or features to phonology that have not already been proposed in the phonological literature and that most phonologist will agree on must exist in some form or other. I do rely heavily on autosegmental and prosodic phonology and I assume that phonology is able to integrate various "chunks" of phonological structure into a phonologically well-formed string. The details of this process will remain unclear, but suffice it to say that such a process is necessary in prosodic morphology as well, since phonology must be able to construct a complex word by combining several autosegmental morphemes. The current proposal employs the same principles, primarily Left-to-Right Association, that are used in autosegmental phonology and prosodic morphology.

The second —programmatic— goal of the current study is mainly concentrated in chapters 1 and 5. The gist of the argument is that theoretical linguistics cannot afford to ignore the fact that any complex system that must be implemented in some sort of "hardware" is constrained by the architectural structure of that "hardware." I put the term "hardware" in scare quotes here, because it should be obvious that what is meant is the brain, which is not really hardware in the computer sense (and certainly not hardware in the traditional sense). I use the

term here to indicate that language, which is an abstract system, is ultimately implemented in a physical medium.

Any complex system that is implemented in a physical medium bears the signs of the way this physical medium is structured. This means that it is not possible for theoretical linguistics to just assume whatever seems more economical, more principled or more plausible from the point of view of the theoretical model alone. To put it in a slightly misleading terminology: design decisions at the lower level have implications for the design options at the higher level. (The misleading aspect of this formulation lies in the agentive nature of the words *design* and *decision*. No such agentivity is intended here.) Ignoring these architectural constraints may lead to models that are incorrect, in the sense that they may make assumptions, state generalisations and define processes and structures that do not capture what is actually going on in the physical medium.

A commonly heard counterargument to this is that what we know about the physical medium that implements language, i.e., the brain, is too far removed from what a linguistic analysis requires in order to be at least descriptively adequate. This is to some extent true, but not so much that it should stop us from attempting to bridge this gap. As I argue in chapter 1, what we know about the hardware's architecture is more informative than sometimes assumed in generative circles. No theoretical linguist would deny the need to eventually bridge the gap between theoretical linguistics and neuroscience. What I wish to argue is that we can and should start doing so now.

Another way to put this is that we cannot ignore the advances in language modelling that are being made in fields that use so-called *deep learning* methods and connectionism (see, e.g., Golosio et al. 2015 for a recent example). Such models allow us to study how certain linguistic processes function at a very low level. The common counterargument that such models cannot really tell us anything about how the brain works because they do not *actually* model the brain is true, but only to a certain extent. While it is true that they do not model the brain directly, they are built using similar building blocks and they are therefore more informative than anything else we have regarding the way the brain *could* handle language. A connectionist model does not show us how the brain handles language, but it does show us ways that a neural network can employ in order to do so, and thus may help us constrain our theories such that they are likely to be more compatible with the brain than theories without such restrictions.

Although advances in connectionist models of language should not be ignored, they should not be overstated either. Golosio et al.'s claim that their model "produced 521 output sentences" (p. 1) is fascinating, but it does not lead to a deeper understanding of the cognitive ability that we call language. For that, we need a detailed analysis of the model, informed by a thorough understanding of language from a high-level, theoretical point of view. Basically, what I am arguing

for is more interaction: theoretical linguistics can aid in understanding connectionist models better and in suggesting phenomena to try and model, while connectionism can help improve theoretical models of language.

I would like to note (perhaps needlessly) that taking a more serious look at the results of connectionist language models does not entail a revival of behaviourism. Chomsky's comment in his review of Skinner's book *Verbal Behavior* is still very true:

> One would naturally expect that prediction of the behavior of a complex organism (or machine) would require, in addition to information about external stimulation, knowledge of the internal structure of the organism, the ways in which it processes input information and organizes its own behavior. These characteristics of the organism are in general a complicated product of inborn structure, the genetically determined course of maturation, and past experience. (Chomsky 1959, p. 27)

To me, there is no doubt that "the internal structure of the organism" and "the ways in which it processes input information and organizes its own behavior" constitute what is interesting about language, indeed about *any* cognitive ability. It is these aspects that we need to understand if we really want to understand how language works. Connectionist modelling is a tool that helps us do this, even if it does not provide the definitive answer.

These considerations, discussed more elaborately in chapter 1, lead me to adopt the assumptions about syntax that I outlined above: features are not innate, and should therefore be visible in the linguistic input. Furthermore, I argue that features should not be seen as objects in their own right. Each feature is essentially just a short-hand for all the elements that possess this feature and whenever a feature acts in some fashion in a syntactic computation, it is actually the bearer of the feature that acts. This may seem to contradict developments in theoretical linguistics since at least Beard's (1988) *Separation Hypothesis*, which states that syntactic features should be separated from the forms that express them, but it is not. A theoretical grammar model *should* make this distinction and be formulated in terms of features. Features are a way of expressing generalisations about linguistic forms; without them, there would be no way to formulate a grammar theory. Yet it should be kept in mind that they are just that: generalisations over linguistic forms. They are not entities in their own right and cannot be subject to syntactic computations without any consideration of the forms they generalise over.

In chapter 2 I discuss the grammar model that I assume, which includes an overview of the theory I am proposing, and which I dub *prosodic syntax*. The

central idea is that linear order is a requirement of the phonological system and it is ultimately this system that determines where elements end up in the linear string. This idea is less remarkable than it may initially appear. Phonology has no way of reorganising strings, i.e., of doing "phonological movement", "prosodic inversion", or any such operation that rearranges items in the phonological structure. What it does have is an array of autosegmental tiers, each of which contains elements that are sequentially ordered within the tier. These elements need to be associated with elements on other tiers, ultimately with elements on the segmental tier. Syntactic heads can be associated with elements on autosegmental tiers, and as a result of this they may end up in a position in the phonological string where one might not expect them given their syntactic position. The fundamental relation between syntax and phonology can still be described as 'c-command corresponds to precedence' (in fact, the model is compatible with a Kayne-style LCA, but also with a system that allows linearisation parameters), but this correspondence breaks down when the phonological form of the element in question is autosegmental.

The model that arises can provide a more elegant description of certain phenomena that are difficult to capture in a more standard model. In chapter 3, I discuss several such phenomena, which include *simultaneity* (the simultaneous expression of multiple meaningful elements), certain kinds of head movement, and the affix-clitic distinction. Each of these is illustrated with several examples, and the picture that eventually emerges suggests that phonology has a more important role to play in establishing the form that linguistic utterances take than usually assumed.

Chapter 4 turns to a different question that is also amenable to a more enlightening answer in light of the prosodic syntax model. The question whether morphology constitutes a separate module in the grammar of human languages or not is an often debated topic. With the framework of Distributed Morphology (DM) having become mainstream in generative grammar, the accepted answer to this question, at least in parts of the linguistic community, in "no". It should be pointed out, however, that even DM reserves a special corner of its grammar model for typical morphological operations (e.g., fission, fusion, impoverishment, etc.) Such operations only apply at a single point in the model, specifically at the end of the syntactic derivation, right before Vocabulary Insertion. Such a model does not differ from a model that assumes a separate morphology module into which fully-derived syntactic structures are fed, and whose output constitutes the input to Vocabulary Insertion. This assumption, in my view, is actually at odds with the stated goal of DM, which is to have *syntax all the way down*.

Prosodic syntax offers a different way of looking at the syntax-morphology distinction. Although it is still the case that there are certain operations whose

effects are intuitively morphological, they can apply anywhere in the syntactic tree, or, assuming a derivational model, anywhere during the derivation. Furthermore, the *syntactic* part of such operations is not distinguishable from other syntactic operations. The effects that give them their 'morphological' feel happen in phonology. For example, the addition of a suffix such as the plural marker *-s* to a nominal root is syntactically identical to the addition of the determiner *the* to a noun phrase, despite the fact that the suffix is a bound morpheme while the determiner is a free-standing form. This difference between the two elements is located in the lexicon and is actually encoded in their phonological forms. Syntactically, they are both just heads c-selecting a nominal complement (though selecting slightly different nominal categories, e.g., N vs. *n* or $\bar{\text{N}}$).

The final chapter, chapter 5, addresses another question, one that I believe finds an interesting solution in the prosodic syntax framework developed here. The question is how competition can be incorporated into the framework. One common way of dealing with competition is to adopt an Optimality Theory framework, but OT models are not easily incorporated into grammar models such as minimalism or HPSG. Yet, the essential role that the *Elsewhere Condition*, or its DM equivalent underspecification, plays in grammar is an indication that no grammar model can really forego dealing with competition.

The prosodic syntax model that is developed in chapters 2– 4 can be seen as a model in which small pieces of structure, a semantic concept, a syntactic head, a prosodic word, are linked to each other, with each type of element (semantic, syntactic, or phonological) being subject to rules of composition that are specific to the relevant module (semantics, syntax, or phonology). The links between these pieces of structure can be seen as excitatory relations: if element X is present, element Y linked to X must also be present.

In chapter 5 I argue that this model should be extended with two additional assumptions. The first is that the pieces of structure that are linked do not have to be atoms. They can be larger chunks of structure, e.g., a syntactic structure combining two heads, or a syllable and the context in which it must be inserted. This is not a new assumption, it has firm roots in construction grammar, and is also adopted by Culicover and Jackendoff's (2005) *Simpler Syntax* approach and in *nanosyntax* frameworks. I should perhaps point out that I do not advocate a construction grammar approach: syntactic structures are construed out of syntactic atoms (although I consider it possible that stored constructions are used in speech production and perception). Nonetheless, the lexicon can store pieces of structure that are larger than atoms and these, like all other pieces of structure, can be linked to. I call such pieces of structure *schemata*, a term borrowed from representational grammar frameworks, although it should be noted that I use the term in a different meaning. For me, any piece of structure stored in the lexicon, even atomic heads, are schemata.

The second assumption with which I extend the model is the idea that links between schemata are not necessarily excitatory, but can also be inhibitory. This aspect is crucial for the analysis of competition. The basic idea is that if two morphemes A and B (which in the model to be proposed are two phonological schemata) compete for the same position, but with B requiring a more specific syntactic context, B has an inhibitory relation to A. This means that if the syntactic context for B is present, this context activates both A and B (because A's context is a subset of B's), but B then inhibits A. This allows a fairly low-level implementation of the *Elsewhere Condition*, without the need to actually compare the contexts of A and B *as part of the derivation*.

Adopting the idea of excitatory and inhibitory connections between schemata should not be taken as an attempt to turn the grammar into a connectionist model. Although connectionist models, and indeed neural networks, also employ such connections, I do not wish to make the claim that a one-to-one mapping can be made between schemata and their connections in the proposed model on the one hand and units / neurons and their connections in connectionist / neural networks on the other. The two models are still too far removed from each other to even consider such an idea.

In my opinion, it is merely an interesting coincidence that when one attempts to construct a model of linguistic structure that is as simple as possible (by measures that are discussed in chapters 1 and 5.1) one arrives at a model that is in some ways similar to a connectionist network. It is a coincidence, because it has not been my specific aim to make the model look like a connectionist network. It is interesting, because it raises the question whether a network of excitatory and inhibitory relations between pieces of structure is a good way of describing the function and structure of a connectionist (or neural) network *at a higher level*. If it is, then the current proposal may be on the right track. Whether this is indeed the case, however, will have to await future research.

1

An architecture for grammar

The main purpose of this chapter is to sketch a broad outline for an architecture of grammar, although not by specifying a specific set of modules and the manner in which they interact —that is partially the purpose of the rest of this book— but by specifying some of the limits that should be imposed on a theoretical grammar model. Crucially, however, these limits do not follow from grammatical theory itself. Rather, they are suggested by domains of research that also deal with language, albeit from a different point of view.

This in fact constitutes an additional purpose of this chapter — linguistic theory, I believe, should interact more with fields that also study language, but which do so at a different level of analysis, i.e., with a different granularity and with different primitives, than theoretical grammar. The key argument is that the structure of the system in which language is implemented, i.e., the brain, has a decisive impact on the structure of language and the language faculty.

While this point seems obvious enough, the general tendency in generative linguistics seems to be to consider this fact inconsequential as far as linguistic theory is concerned. To put it somewhat more provocatively, generative linguistics suffers from a strong *not invented here* syndrome. The common assumption is that linguistic theory is so far removed from neurological facts that there is no point in trying to bring the two together.

At first sight, the fundamental notions of neuroscience and linguistics do indeed seem too disparate to *ever* be connected. The workings of neurons, of synapses and axons are hardly the kinds of primitives that would allow us to formulate any kind of linguistic theory, in the same way that studying how a logic gate works would not tell us very much about how a computer handles a web search. However, one is undoubtedly at the basis of the other and there is a finite number of intermediate levels that connect the two.

Linguistic theory therefore cannot be connected directly to the neurological level, but it can be connected to the level right below it. This is a level that abstracts away over the details of neurons, synapses and axons and presents a model of what is sometimes called *neural computation*, offering primitives that come closer to what a linguistic theory needs.

The first section of this chapter discusses this point in somewhat more detail. This section is followed by two sections that attempt to put the ideas into practice, by discussing two fundamental questions of linguistic theory in light of what we know about neural computation. The final section of this chapter takes a different perspective. Instead of looking at the next lower level, it looks at linguistic theory from a higher level, which involves the structure of the building blocks of communication.

The conclusions of this chapter are that in order for linguistic theory to be compatible with the analytic levels surrounding it, we must put specific restrictions on certain aspects of our models — particularly with respect to the primitives of syntactic structure, the nature of features and the status of phonologically empty heads.

It should be noted that the discussion in this chapter is necessarily limited in scope. A thorough discussion of the questions raised here could easily fill a book of its own, and probably more than one. The main aim of this chapter is to provoke questions that I believe any theoretical linguist should ask, or at least keep in the back of their head.

1.1 Some remarks on the purpose of the generative enterprise

I believe it is not too dramatic to say that the discipline of linguistics suffers from something that might be called the *linguistic divide*: the fact that there often seem to be two completely separate disciplines of linguistics, differentiated by the question, to put it simply, whether or not they accept "Chomsky". Although it will obviously not be possible to resolve this matter in a single introductory section, I nonetheless want to suggest that this divide is unnecessary and even detrimental to the field of linguistics.

I should point out from the start that I do not want to argue that either side of the divide is misguided, nonsensical or simply unnecessary. Both sides, which I will refer to as the generativists vs. the empiricists,[1] expand our knowledge of

[1] Primarily for convenience' sake. I do not wish to imply too much with these terms. Specifically, it is not my intention to claim that people working in a generativist framework in a broad sense completely ignore empirical data. A lot of experimental work is being done on the "generativist" side.

language in meaningful ways, each in their own way. The field, I believe, would benefit if both sides took each other seriously and actively talked to each other.

To summarise the discussion that is to follow, it helps to always keep Box and Draper's (1987) famous quote in mind: "Essentially, all models are wrong, but some are useful" (p. 424). The measure of the usefulness of a linguistic model is the extent to which it helps us understand language as a cognitive system. High-level models that completely ignore the lower-level architecture on the assumption that it is irrelevant, are less likely to be able to do so. Theoretical linguistics should therefore be aware of what is going on in research areas that focus on lower levels of analysis and should consider whether lower-level results can be integrated into its higher-level theories.

1.1.1 The linguistic divide

Ever since its inception, the generative enterprise has been concerned with finding an explicit, formal description of the rules underlying the syntax of natural languages. This has been the explicit goal of Chomsky (1955, 1957) and remains an important goal of current generative linguistics. In his influential 1965 book *Aspects of the Theory of Syntax*, Chomsky calls this goal *descriptive adequacy*: a complete, formal description of the knowledge a speaker has of his or her language, which tells us which sentences it can generate and which it cannot.

In that same book, Chomsky formulates another goal that generative grammar should strive toward: apart from describing the linguistic knowledge of a speaker, the theory should also explain how the speaker came by this knowledge. This goal, dubbed *explanatory adequacy* by Chomsky, led, in combination with the familiar *Poverty of the Stimulus* argument (e.g., Gold 1967) to the idea that certain aspects of grammatical knowledge must be innate: the well-known Universal Grammar (UG) hypothesis.

The UG hypothesis has resulted in a formidable amount of criticism. The Poverty of the Stimulus argument has been taken apart (Pullum and Scholz 2002) and it has been argued on many occasions that the learnability theory underlying Gold's argumentation is not realistic (cf. Johnson 2004). For example, Elman (1993) argues that the key to learning language may lie in the limited processing capacity of infants. As a result, infants do not try to analyse the language of their environment as a whole, but focus only on particular aspects first,[2] and turn to other aspects later. This step-wise focusing on different aspects does not come about because of some innate bias, but is caused by the fact that young children simply do not have the ability to even perceive and process all aspects of human

[2] Which is often argued to be prosodic structure, cf. Christophe et al. (2008), Morgan and Demuth (1996), Nespor, Shukla, et al. (2008), and Soderstrom et al. (2003), etc.

language. Language, on this view, is especially well adapted to be learnt by an infant brain, because those aspects that a child is forced to focus on first help it to bootstrap acquisition (see also Deacon 1997 for this idea).

In spite of the criticism, the UG hypothesis is still going strong. One reason for this is that despite the —in my opinion— undeniable problems of the UG hypothesis, the alternative suffers from similar shortcomings. As Lasnik and Uriagereka (2002) scathingly put it in their reaction to Pullum and Scholz's criticism of Poverty of the Stimulus arguments: "[...] it is pointless to engage in mathematical analyses of part of the data. What could that possibly decide if there is only one theory under scrutiny?" (p. 150).[3] The point is that although the empiricist approach offers clear impressions of how language acquisition might take place, it does not (to my knowledge) offer a worked-out theory of language acquisition with the precision that generative syntacticians are used to from their analyses of syntactic phenomena.

On the other hand, the quip that the empiricist approach is not even a theory is not entirely fair, given that the UG hypothesis is nothing more than what its name implies: an hypothesis. Since there is no worked-out proposal as to the contents of UG, it cannot sensibly be claimed that it constitutes a *theory* of language acquisition. A theory must provide verifiable hypotheses, which the UG hypothesis in its current form does not.[4]

In essence, the field is divided into two camps that are in a way engaged in a cold war. Among generative linguists, there is generally little inclination to listen to "the other side", and the empiricists have generally stopped listening to generativists when they started exploding their functional structures.[5] This, I believe, to the detriment of the field.

1.1.2 Symbolic vs. subsymbolic models

The controversy between the two camps is sometimes phrased in terms of symbolic vs. subsymbolic models. In fact, the fight between these two factions has been decided for several years now and the result is a resounding tie. As Anderson, Bothell, et al. (2004) make clear, the symbolic and subsymbolic approaches

[3] The example that Lasnik and Uriagereka (2002) focus on is English auxiliary inversion in questions. This is hardly a convincing example, however, because it should be obvious that the crucial factor that tells a child which auxiliary to front has to do with meaning: the auxiliary that is fronted is the auxiliary in the proposition that is being questioned. This does not answer the question how exactly the child arrives at the correct grammatical analysis, but it seems quite possible that it does not require innate knowledge. Furthermore, Bod (2009) shows that a machine can actually extract the relevant rule without reference to meaning.

[4] In principle, the UG hypothesis is at least itself a verifiable hypothesis, although we cannot simply set up an experiment to test it.

[5] I am obviously exaggerating here, but I believe the general sentiment is true.

are not at odds with each other. Rather, they are two sides of the same coin; they complement rather than exclude each other. The essential difference between them is the level of analysis: symbolic models are generally formulated at a higher level of analysis than subsymbolic models. Crucially, however, one cannot get a full picture of the phenomenon one is studying without considering *all* levels of analysis (cf. also Anderson 2007 for some discussion of this point).

I would like to follow Anderson, Bothell, et al.'s lead and argue that both sides of the linguistic divide are contributing to our understanding of language in meaningful ways. Doing generative linguistics is in essence studying the brain at a high level of analysis. The job of generative linguistics is to understand the system at a high level of abstraction, to see how the various components of linguistic computation (semantics, syntax, phonology, but also pragmatics, information structure, etc.) cooperate to produce utterances.

This does not mean that generative linguistics should not care about the details. It should, of course, since the details tell us a lot about how the system works. The point is, however, that this research should be complemented by research at lower levels of analysis. An often heard claim among generativists is that their models are descriptions of the language knowledge that an idealised speaker possesses, and that they are in no way attempting to describe the psychological or neurological reality of language processing (see, e.g., Struckmeier 2011 for a recent formulation of this idea). In other words, generativists often completely detach themselves from the psychological and neurological aspects of language as a matter of principle.

Marantz (2005) argues that this is not actually the case. Basically, Marantz argues that only a "Platonist" view of linguistics would be compatible with a strict separation between competence and performance theories, and states that "[…] generative grammar is not and has never been Plantonist (*sic*) in theory or practice" (p. 431–432). I tend to disagree with Marantz on this point, especially on the "in practice" part: the competence/performance distinction is often referred to as an argument for restricting research data to grammaticality judgements. Marantz' own experimental psycholinguistic work, however, shows that he is at least partially right. Similarly, Poeppel and Embick (2005) argue that linguistic research should not be limited to abstract analyses of grammar but should also strive to establish a bridge to neurolinguistic research.

However, both Marantz and Poeppel and Embick focus on a specific kind of lower-level research: that of psycholinguistic and neurolinguistic experimental approaches. Poeppel and Embick argue that the fundamental elements of representation and the fundamental operations on primitives in neuroscience and in linguistics are incommensurable and the granularity of the analyses in both domains is vastly different. In order for future cooperation between the fields to be fruitful, they claim, it is necessary for psycho- and neurolinguistic approaches

to focus on identifying the neural mechanisms that could support the kinds of computations and representations that linguistics assumes. Crucially, they state that doing it the other way around, i.e., restricting the operations and primitives of linguistics to those that can be explained by neuroscience, would be baseless (and would in fact stifle the field):

> An explanatory theory of linguistic computation in the brain should employ linguistic categories as a means of exploring neural computation; but the failure to detect distinctions in any particular case does not necessarily imply that the linguistic distinctions are incorrect. The latter type of inference might be possible in the context of an articulated theory of neurolinguistic computation; but we have nothing like that at present (Poeppel and Embick 2005, p. 108).

Poeppel and Embick are correct when they say that linguistic research should strive to establish a bridge to neurolinguistic research. However, this bridge cannot be established without piers: one or more intermediate levels of analysis that help bridge the gap. Poeppel and Embick essentially claim that such an intermediale level of analysis does not exist, but their claim that we have nothing like an "articulated theory of neurolinguistic computation" is not entirely correct. Although there is no *fully* articulated theory, there is certainly work being done toward reaching that goal. This work, however, involves an area that generativists generally do not take seriously: connectionist modelling. The reasons for this negative attitude among generativists toward connectionism (and vice versa, by the way) are largely sociological and historic, I believe, even though it is of course motivated with scientific arguments. One of the main generativist arguments against connectionism is that this approach is subsymbolic, while linguistic computation is "obviously" symbolic. The other way around, one important argument against generativism from connectionist empiricists is that the brain does not employ symbolic computation and that representations of information and categories are distributed, not localist, as would be required by symbolic computation.[6]

It is exactly this kind of research that I have in mind, and that Anderson, Bothell, et al. (2004) (seem to) have in mind when they claim that symbolic and subsymbolic analyses (should) complement each other. Connectionist modellers often do not generalise over their results in order to create a linguistic *architecture* in the sense of Anderson's definition of the term. This, I believe, is the task

[6] This is somewhat of a simplification. In actual fact, the discussion between localised and distributed representations is far from resolved in neural network and brain research; cf. Thorpe (2003) for discussion.

of generative linguistics.[7] In fact, Anderson's discussion is very informative in this respect and so deserves some attention here.

Anderson defines a *cognitive architecture* as follows:

> A *cognitive architecture* is a specification of the structure of the brain at a level of abstraction that explains how it achieves the function of the mind. (p. 7)

The term *function of the mind* is admittedly vague, but Anderson states that it "can be roughly interpreted as referring to human cognition in all of its complexity" (*ibid*). Basically, then, it refers to everything that the brain achieves that goes beyond simple, automated processes such as heart beat and breathing regulation, blinking, etc. This, of course, does not really make the definition more precise, since there is no clear line separating "simple" processes from more complex ones, but it should be intuitively clear what sort of processes are intended. We can proceed on the assumption that it is currently not necessary (or even possible) to give a more precise definition of *function of the mind* and that a more precise definition should become available once we understand the brain and the mind better.

Anderson's definition establishes the notion of *cognitive architecture* as a way to model the link between the brain and the mind. It assumes a low-level model of the brain, describing its building blocks (neurons, synapses) and how they interoperate. It also assumes a model of higher cognitive functions, describing their properties in a principled way. The cognitive architecture then explains how these higher functions can be achieved with the means that the low-level model provides.

Anderson then goes on to note that in the history of cognitive science, there have been three approaches that each failed to address one aspect of this definition. The first approach is what he calls *classic information processing psychology*, which ignores the brain. This approach attempts to explain the mind by creating models of information processing without considering how such models could be implemented by the brain. The second approach, which Anderson calls *eliminative connectionism*, ignores the mind. It attempts to model the neural processes going on in the brain during information processing, without looking at the larger picture, i.e., answering the question what the larger structure of the brain is and why it is structured this way. Finally, the third approach is dubbed *rational analysis* by Anderson and, according to him, ignores the architecture. The modern variant of this is the interest in statistical and Bayesian modelling.

[7] Or rather, of *formal* linguistics. The relevant architecture would only (have to) be generative in the broadest sense of the word.

Although Anderson writes about the field of cognitive science, his remarks hold true for the field of linguistics as well. Statistical and Bayesian methods are especially popular in corpus linguistics and in certain areas in computational linguistics. Connectionist models of linguistic phenomena are often developed in the tradition of connectionist research in psychology and cognitive science and share the basic assumptions and approaches of this field. And, most relevant to the current discussion, standard generative linguistics employs the same methods as *classic information processing psychology*: generative linguists generally ignore the brain. The assumption is that language should be studied as a domain in its own right, abstracted away from the "hardware" that implements it. Once linguists have established the primitive operations and functions that are needed for language, the field of neuroscience should then try and figure out how these operations and functions are implemented in the brain (cf. specifically Poeppel and Embick 2005, discussed above).

The idea behind this stance is that the brain, as a complex system, must be studied at different levels in order to understand it completely (which is correct) and, more importantly, that these different levels are phenomenologically *independent* of each other (which is not correct). That is, standard generative grammar assumes that although the properties present at one level must be implemented in terms of properties of the next-lower level, the latter do not impose *restrictions* on the former. Seen in that light, it becomes (almost) understandable that many generativists are not interested in the results of empiricists, especially when it is clear that language has properties that are difficult to implement in, e.g., standard artificial neural network approaches (cf. Hurford 2002; Jackendoff 2007).[8]

However, as just remarked, this view is incorrect. In a complex system, *any* complex system, the properties of one level *are* restricted by the properties of the levels beneath it.[9] Of course, once you have a system that is Turing-equivalent (which the brain arguably is), it is possible to implement any computational algorithm. However, some algorithms will be more difficult to implement than others, and it depends on the architecture of the system which these are. Again, this applies to *any* complex system, including the brain.

To make this point more concrete, it is helpful to look at another type of complex system: the electronic computer. Anyone who has done some computer programming, especially when that involved calculations, is familiar with the

[8] But see Pulvermüller (2003) for proposals on implementing syntactic operations in a neural architecture.

[9] As an important side note: it should be kept in mind that these "levels" are purely descriptive. A complex system such as the brain can be studied at different levels, but it is not the case that these are neatly separable or separated in the anatomy of the brain. (They do not, for instance, coincide with the hierarchical structure of the brain, which is messy in its own right).

so-called *floating point error*:[10] when doing numeric calculations, small errors sometimes pop up. For example, on the computer I am writing this on, the Python interpreter does the following when adding the numbers 0.1 and 0.2:

```
>>> 0.1 + 0.2
0.30000000000000004
```

This is not a quirk of the programming language being used: Elisp, running on the same machine, has the same problem:

```
Elisp> (+ 0.1 0.2)
0.30000000000000004
```

The cause of such floating point errors lies in the architecture of the underlying system: standard electronic computers are implemented as binary machines. In a binary architecture, it is not possible to represent decimal fractional numbers accurately. The problem is similar to the problem that arises when representing a number such as 1/3 in a decimal system; 1/3 is not entirely equal to 0.3, nor to 0.33, nor to 0.333, etc. Representing decimal fractional numbers in a binary system is equally problematic.

In other words, the architecture of the underlying system, which is binary in nature, puts restrictions on the properties of the system at higher levels. The simplest possible algorithms for implementing decimal fractional computation on a computer cannot represent decimal fractions accurately. Usually, the deviations are so small that they do not influence the final result, but there are cases where absolute precision is required. Because computers are Turing-equivalent machines, it is possible to implement a decimal data type that gives accurate results in computations, but such algorithms are more complex and more resource-intensive.

A similar example is provided by the GPU (*Graphical Processing Unit*) chips common in modern computers. These processors are designed in such a way that they can efficiently handle specific computations needed in computerised graphics manipulations. Although conventional CPUs (*Central Processing Unit*, the "standard" processor in a computer that handles most computations) can be programmed to do the same kinds of computations, the architecture of a GPU allows these computations to be executed much more efficiently.[11]

Obviously, this digression into computer architecture should not imply that computers are an accurate metaphor for the human brain. It has been pointed out

[10] See `http://floating-point-gui.de/` for extensive information and discussion.

[11] In fact, much of the graphical effects that are common in modern operating systems, transparency, translucency, drop shadows, animations, etc., would not be possible without the use of GPUs; executing them on the CPU would simply be too resource-intensive.

often enough that modern digital computers are a bad analogy for the way that brains work. The point of comparison here is the fact that both computers and brains are complex systems with a hierarchical structure that can (and must) be studied at various levels of analysis. These levels, however, are not independent. Each level depends on the level below it and is structurally constrained by it.

In a Turing-equivalent system, such limitations may be overcome by implementing more complex algorithms. The question one needs to ask, however, is whether it makes sense to do so. More complex algorithms are more resource-intensive and may therefore be inefficient (i.e., too slow, consuming too much energy, etc.) If the simpler algorithm is good enough, it makes more sense to use it. In a complex biological system such as the brain, which is shaped by evolution, "good enough" is an important measure, because once evolution has resulted in a system that is good enough for a certain task, selection pressure decreases below a critical point, so that no further adaptation takes place.

This is something that we need to keep in mind when studying language. Studying language in the Chomskyan sense is a study of the brain at a high level of description, but that does not mean that we may completely ignore the lower systems. Rather, we should be aware of the theories and advances made at lower levels of analysis and consider if and how these can be integrated in our own higher-level analyses. At the same time, those working at lower levels of analysis should be aware of what is going on at higher levels and consider the same questions.

The answers to these questions are not straightforward; we cannot know in advance what lower-level aspects we must consider in developing higher-level models and vice versa. The answers can only emerge through constant interaction between research done at various levels of description, by constantly adjusting hypotheses at one level in an attempt to incorporate results at another. Obviously, to incorporate does not mean to adopt blindly: results and conclusions at other levels of description of the system one is studying are not the only source of data and the only observations that one needs to model and explain. However, if one can model the phenomena one is interested in in such a manner that the analysis is (partly) compatible with analyses at a lower level, such an analysis should be preferred over an incompatible one.

This approach is different from what Poeppel and Embick (2005) argue for. They claim that "[a]n explanatory theory of linguistic computation in the brain should employ linguistic categories as a means of exploring neural computation;" (p. 108). For Poeppel and Embick results flow in one direction: from theoretical linguistics (high-level) to neural computation (low-level). In their view, the absence of an intermediate level (the "articulated theory of neurolinguistic computation" that they refer to) makes the reverse pointless. If we accept that we do have something that goes toward an explanatory theory of linguistic com-

putation, we should also consider whether the results achieved at that level of research should impact theoretical models of language. In my opinion, it should, although, as just discussed, this does not make it any easier to decide *which* aspects of these results we should adopt into a theory of grammar.

However, it is important that we try. In the next section, I discuss a fundamental issue in linguistic theory in this light: the issue revolves around the question what the primitives of linguistic structure are, i.e., the building blocks out of which linguistic (syntactic) structures are constructed. It will become clear that if we take the results from lower-level analyses seriously, we must assume that the primitives of linguistic structure contain not just morphosyntactic features (a common assumption in minimalist theories), but also phonological features.[12] The discussion also touches upon the issue of features, what they are and what they cannot be.

1.2 The primitives of linguistic structure

Government & Binding Theory and Minimalism standardly assume a grammar architecture that resembles an (inverted) capital Y, as demonstrated in figure 1.1.

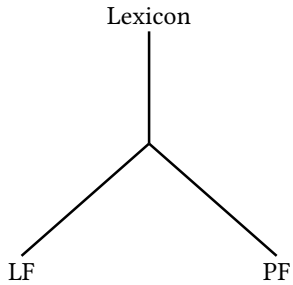


Figure 1.1: Y-model

The fundamental assumption behind this grammar model is that the construction of a sentence starts out with the derivation of its syntactic structure. Only after this structure has been completed (or, in current variants of minimalism, when a phase is completed; cf. Chomsky 2001 and subsequent work), are the semantic and phonological structures of the sentence constructed.

In more recent work, Chomsky seems to be silent on the question whether the primitives of structure, i.e., the terminal elements in the syntactic tree, are lex-

[12] I assume that a similar reasoning can be applied to semantic features, but that is beyond the scope of this chapter.

ical items with a phonological matrix or not, but the most influential proposals in minimalism assume that they are not. Probably the most explicit formulation of this idea is found in the framework of *Distributed Morphology* (DM; cf. Halle and Marantz 1993 and much subsequent work). The DM model, displayed in figure 1.2, is still a standard Y-model: the derivation starts out from the morphosyntactic feature store[13] (the top oval in the figure) and then proceeds toward the phonological and logical forms. However, DM explicitly incorporates the assumption that the morphosyntactic store contains only morphosyntactic features, which are combined into a syntactic structure during the derivation. Before the resulting structure is sent to the phonological component, it undergoes a set of morphological operations, which are operations defined on the syntactic tree that can change local syntactic structures in ways that ordinary syntactic operations cannot.

Only after these adjustments to the syntactic structure have been made is the structure given phonological content, using elements from the *Vocabulary*. This process is known as *Vocabulary Insertion* and associates *Vocabulary Items* (VIs) with elements in the syntactic tree. A VI is an element that specifies a phonological form and a morphosyntactic context in which this form is inserted. Obviously, VIs are primitives in some sense of the word, but they are not the building blocks of structure.

In fact, DM forces us to provide a precise definition of the term "primitive of linguistic structure". The term as understood here is meant to convey the notion of the smallest possible element that is combined by a speaker's computational system to build larger structures. Intuitively, one is inclined to think that such an element has a meaning or (at least) a function. But obviously, it is exactly this notion that DM denies: DM argues that the meaning-bearing elements are *not* the building blocks of syntactic structure. The building blocks of syntactic structure are morphosyntactic features. Bundles of such features may (or may not, in the case of underspecification) be associated with phonological features and with meaning *at a later stage in the derivation*, after (a relevant part of) the structure has been completed.

Note that this view of things leads to a very strong conclusion: the morphosyntactic features that are the building blocks of syntactic structure must be innate. The reason for this is that they are not guaranteed to be observable from the input a language-learning child receives. A child's input consists of phonetic and prosodic features. Abstracting morphosyntactic features from these is a complicated process with an unpredictable outcome and is therefore not guaranteed to result in the same set of features for every speaker. How unpredictable this

[13] There is no 'lexicon' in DM; rather, the information stored in the lexicon in other theories is divided between the morphosyntactic store, the Vocabulary and the Encyclopedia.

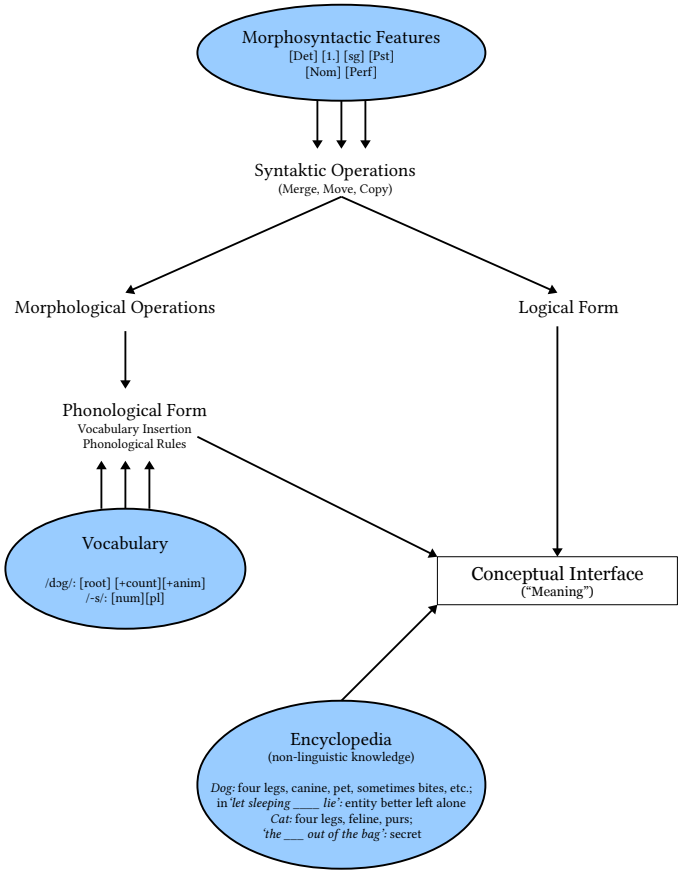


Figure 1.2: Inverted Y-Model in DM (adapted from Harley and Noyer 1999)

outcome is can be demonstrated with a small thought experiment: a look at the nominal system of German through Bantu eyes.

1.2.1 A Bantu perspective on the German nominal system

In the German nominal system, we can distinguish several distinct noun classes, some with subclasses. Table 1.1 lists six classes, two of which have subclasses, (which, however, will be eliminated in the following discussion). Although German has a prefix system, it is extraordinary in several ways. Most importantly, it is not possible to define classes purely on the basis of prefix forms: the prefix *die*,

for example, occurs in classes I, IV, V and VI; the null prefix ∅ occurs in several classes. Most strikingly, all three prefixes *die, der* and *das* occur in class V. For this reason, classes are distinguished not only on the basis of their prefixes, but also on the basis of suffixes appearing on the nominal.[14]

Class		Prefix	Suffix	Function
I		*die, eine*		
II		*der, ein*		
III		*das, ein*		
IV		*die,* ∅		plural of I, II, III
V	*a*	*die, der, das*	sec	
	b	∅	prm	
	c	*ein-*	mix	
VI	*a*	*die*	sec	plural of V*a*
	b	∅	prm	plural of V*b,c*

Table 1.1: German noun classes (preliminary)

Another important thing to note is that the prefixes are *phrasal*. That is, they occur only once in each nominal phrase. So whereas for example in Swahili, a phrase such as *mtoto mdogo* contains two occurrences of the class I prefix *m-*, its equivalent in German, *das kleine Kind*, only contains one occurrence of the class III prefix *das*, whereby the prefix is always the very first element of the nominal phrase.

A further complication with the prefixes is that they are complex. They consist of a stem *d-* or *ein-*,[15] combined with a suffix that indicates the grammatical function of the nominal phrase.[16] Four grammatical functions are distinguished, which I simply number 1–4 (for clarity's sake, I will generally refer to them as GF1–GF4 in the text). GF1 is the argument with which the verb agrees, and often (though not necessarily) expresses the agent or cause role of the verb. GF2 is generally the second argument, the theme or patient, though sometimes GF3 (and on rare occasions, GF4) is also used for the second argument. In most cases,

[14] The suffix system will be discussed below. For the moment, it is sufficient to note that there are three suffix groups, which I label *primary*, *mixed* and *secondary*, and that the nominals in classes I–IV are not characterised by any particular suffix, although some subclasses can be distinguished that are. These subclasses will not be discussed here, however.

[15] In fact, there are several other elements may occur as phrasal prefix as well. They will not be discussed here, as their behaviour is similar. The distinction between *d-* and *ein-* is functional: roughly speaking, *d-* marks nouns that are known in the discourse, *ein-* marks nouns that are new. Note that *ein-* cannot combine with class IV nouns; its functional equivalent is the empty prefix.

[16] The empty prefix does not take a suffix, for obvious reasons.

GF3 is the third argument of the verb and usually expresses the recipient, bene- or malefactive. GF4 is rarely a verbal argument; rather, it expresses the nominal dependent of another noun. The thematic relation between the two nouns is not restricted. Often, it expresses some form of possession, but this is not necessarily the case: any (contextually relevant) relation is possible.

In classes I–IV, the suffixes that appear on *d-* and *ein-* correlate with the noun class. For ease of reference, I will refer to these as *suffix classes A–D*. Suffix class *A* co-occurs with noun class I, suffix class *B* with noun class II, etc. The full overview is given in table 1.2. The reason that I do not simply label suffix class *A* as class I is that they are also relevant for the noun classes beyond class IV, as discussed below.

GF	*A*	*B*	*C*	*D*
1	*-ie/-e*	*-er/-∅*	*-as/-∅*	*-ie/-e*
2	*-ie/-e*	*-en*	*-as/-∅*	*-ie/-e*
3	*-er*	*-em*	*-em*	*-en*
4	*-er*	*-es*	*-es*	*-er*

Table 1.2: Suffix classes

Some of the cells in table 1.2 have two forms. The first of these combines with *d-*, the second one combines with *ein-*. In the cells that only have one form, there is nonetheless a distinction related to the stem they combine with: when combined with *d-*, the suffixes are pronounced with a full vowel: either /e/ or /ɛ/, depending on the suffix. When combined with *ein-*, the vowel of the suffix is reduced to schwa. Orthographically, nothing changes. because the schwa is written «e», but the pronunciation is clearly distinct. We can describe this alternation with a phonological rule that reduces the vowel when it is unstressed.[17] We can extend this analysis to the forms *-ie/-e* of class *A*: here, the full vowel is /i/, which is also reduced to schwa when the ending is combined with *ein-*. The only anomalous cases are GF1 of class *B* and GF1 and 2 of class *C*, where *ein-* has no ending at all.

As mentioned, it is not possible to assign noun classes on the basis of phrasal prefixes alone in German. We also need to take the suffixes into account. Nominals of classes V and VI take suffixes that indicate grammatical function and that correlate with the prefix class. Traditionally, three sets of suffixes are assumed, although I argue below that we can reduce these to two. I label these sets *primary*, *mixed* and *secondary*, of which the mixed set is, in my opinion, superfluous.

[17] Unstressed vowels in German are often reduced to schwa, at least in words that are not of foreign origin.

Let us first look at the secondary suffixes. These suffixes appear on nominals that have a phrasal prefix *d-*.[18] Table 1.3 lists all forms. As can be seen from the table, this group has only two endings: *-e* and *-en*. The ending *-en* occurs everywhere, except in GF1 for classes *A*–*C* and GF2 for classes *A* and *C*, which have *-e*. The distribution of the *-e* ending will turn out to be typical: GF1 in the singular classes (*A*–*C*) and GF2 for classes *A* and *C* are grouped together in the mixed group as well, as demonstrated by the endings in table 1.4.

GF	*A*	*B*	*C*	*D*
1	*-e*	*-e*	*-e*	*-en*
2	*-e*	*-en*	*-e*	*-en*
3	*-en*	*-en*	*-en*	*-en*
4	*-en*	*-en*	*-en*	*-en*

Table 1.3: Secondary suffixes

GF	*A*	*B*	*C*	*D*
1	*-e*	*-er*	*-es*	*-en*
2	*-e*	*-en*	*-es*	*-en*
3	*-en*	*-en*	*-en*	*-en*
4	*-en*	*-en*	*-en*	*-en*

Table 1.4: Mixed suffixes

The set of mixed suffixes occurs on Class V-VI nominals that have *ein-*[19] as phrasal prefix. Note that this group has the suffix *-en* everywhere where the secondary group has it. Furthermore, the five cases where the secondary group has the suffix *-e*, the mixed group has the same suffixes as the primary group, as table 1.5 shows.

When we compare the set of primary suffixes with the set of suffixes on *d-/ein-*, we notice that they are almost identical. In fact, given that all primary suffixes are pronounced with a schwa, we can conclude that they are in fact the exact same set. The only differences compared to the suffixes for *d-* are GF1 and 2 of classes *A* and *C*. We explained the *-ie/-e* alternation between *d-* and *ein-* on the basis of a phonological reduction, and we can do the same for the *-as/-es* alternation. Since the suffixes listed in table 1.5 occur on nominals with their own word stress and

[18] Or any of a small group of similarly behaving elements.

[19] As with *d-*, there are a number of elements that that trigger the same suffixes on nominals as *ein-*.

GF	*A*	*B*	*C*	*D*
1	*-e*	*-er*	*-es*	*-e*
2	*-e*	*-en*	*-es*	*-e*
3	*-er*	*-em*	*-em*	*-en*
4	*-er*	*-es*	*-es*	*-er*

Table 1.5: Primary suffixes

full vowel (or vowels), the phonological rule we formulated earlier will force the vowel of the suffix to be reduced.

Summarising the discussion so far, we can make the following observations:

(1) In any noun phrase consisting of a phrasal prefix plus one or more nominals, the following holds:

   a. The *d-* prefix has a primary suffix (with a full vowel); subsequent elements have a secondary suffix (if they belong to class V or VI).

   b. The *ein-* prefix has a primary suffix, except in GF1 of class *B* and GF1 and 2 for class *C*, which remain suffixless; subsequent elements have a mixed suffix (if they belong to class V or VI).

Take a closer look at the set of mixed suffixes. As mentioned, this set has the *-en* suffix of the set of secondary suffixes combined with the forms of the primary suffix in three cases: GF1 and 2 of class *C* and GF1 of class *B*. In GF1 and 2 of class *A*, both the primary and the secondary set of suffixes have *-e*, so we cannot determine where the *-e* in the mixed set comes from. For reasons that will become clear in a moment, I assume that it comes from the set of primary suffixes.

In other words, in the mixed suffixes the primary forms occur exactly there where *ein-*, the phrasal prefix associated with the mixed suffixes, has no (primary) suffix itself. This observation allows us to formulate the following generalisation:

(2) Each nominal phrase has a primary suffix on the first element that requires a suffix. All subsequent elements that require a suffix take a secondary suffix.

Elements that require a suffix are the following:

- The phrasal prefix *d-* plus a number of similarly behaving elements.
- The phrasal prefix *ein-* in all forms except for GF1 of class *B* and GF1 and 2 of class *C*, plus a number of similarly behaving elements.

- Nominals of classes V and VI.

With this generalisation in place, we can reduce the number of noun classes to six. Furthermore, having analysed the phrasal prefixes as consisting of a stem plus a suffix, the prefixes themselves turn out to be irrelevant for noun class membership. What is relevant is the suffix class of a nominal and whether the nominal itself is suffix-taking or not. Table 1.6 shows the noun classes on the basis of these criteria.

Class	Suffix class	Suffix-taking	Function
I	*A*	–	
II	*B*	–	
III	*C*	–	
IV	*D*	–	plural of I, II, III
V	*A*–*C*	+	
VI	*D*	+	plural of V

Table 1.6: German noun classes (final)

For completeness' sake, the primary and secondary suffixes of all four classes are listed in table 1.7.

	primary				secondary			
GF	*A*	*B*	*C*	*D*	*A*	*B*	*C*	*D*
1	*-ie*	*-er*	*-as*	*-ie*	*-e*	*-e*	*-e*	*-en*
2	*-ie*	*-en*	*-as*	*-ie*	*-e*	*-en*	*-e*	*-en*
3	*-er*	*-em*	*-em*	*-en*	*-en*	*-en*	*-en*	*-en*
4	*-er*	*-es*	*-es*	*-er*	*-en*	*-en*	*-en*	*-en*

Table 1.7: German suffix classes

One final remark needs to be made: although it is generally the case that the nouns of classes I–IV do not take primary or secondary suffixes, there are two exceptions to this generalisation. Most nouns of classes II and III take a primary suffix in GF4 (which is often reduced to *-s*) and nouns of class IV take a primary suffix in GF3 (which is always reduced to *-n*; the *-n* is not added when the noun ends in *-n* or *-s*). If the nominal phrase contains other nominals from classes V or VI, these take a secondary suffix. A phrasal prefix, however, always takes a primary suffix, even if the noun has one itself.

1.2.2 The status of features

The Bantu analysis of the German nominal system obviously resembles the traditional analysis in many ways. After all, the same forms are being analysed. However, a bias toward noun classes does lead to an analysis with a different internal structure. In the traditional analysis, the German nominal system is divided up according to gender with a subdivision in each gender according to case. This is not necessarily bad, but the division is obviously inspired by the analysis of Latin and Greek.

The Bantu-inspired analysis uses noun classes as its primary division. Gender distinctions are among the defining characteristics of the proposed noun classes, but they are on a par with number. In terms of the traditional system, one might say that [pl] is a possible value of the Gender feature, in the same way that [masc], [fem] and [neut] are. This makes sense in German, because there are no gender distinctions in the plural, unlike in Latin and Greek.[20] Put differently, the main difference between the Bantu and the Latin analysis of German is the fact that number defines a class in the former but not in the latter system.

As far as descriptive analyses go, the relatively straightforward analysis of the German nominal system presented in the previous section does not seem to be any worse (or better) than a more traditional analysis that relies on gender. If such formal linguistic analyses represent some sort of truth about how linguistic knowledge is stored in the brain,[21] then we have no way of knowing which analysis corresponds to the actual I-language of German speakers.[22]

The point of this little thought experiment and the crux of the argument here is that the E-language does not tell us which features underlie the E-language's grammar, i.e., the I-language that produces the E-language. In the German nominal system, [pl] may be on a par with [masc], [fem], and [neut] as values of Gender (or NounClass), or it may be on a par with [sg] as a value of Number. The former option would mean that there is no need for a [sg] feature and consequently no need for a Number feature, at least as a formal syntactic feature.

Therefore, if one assumes that features are the building blocks of structure, and if one assumes that for analysing language (not "*a* language") it is crucial to find out which features underlie grammar, then one must inevitably assume that those features are innate. They cannot be learnt because they are not reliably represented in the input.[23] This is indeed the common assumption in certain areas of generative linguistics (cf. Cinque and Rizzi 2010).

[20] In fact, the distinctions in Latin and Greek are not truly gender distinctions but declension distinctions, which are for all intents and purposes the same as noun classes in Bantu.

[21] And there is no guarantee that they do; cf. the discussion in section 1.1.2.

[22] Or perhaps whether either could, depending on the speaker.

[23] This is essentially the Poverty of the Stimulus argument, of course.

When one assumes that features are innate, then the inevitable follow-up question becomes: how did features arise as part of UG? Since UG is part of the biological world, as Chomsky never tires of stressing, features must have arisen through some evolutionary process. Deacon (1997, ch. 11) talks about this topic at some length and his conclusion is quite damaging:

> The point is that it should be possible to predict which aspects of our language adaptations are more or less susceptible to genetic assimilation, by virtue of the *invariances* they demonstrate. *Universality is not, in itself, a reliable indicator of what evolution has built into human brains.* [emphasis added]
>
> In summary, only certain structural universal features of language could have become internalised as part of a "language instinct", and these turn out *not* to be those that are most often cited as the core of a Universal Grammar. (p. 338)

Deacon argues explicitly that the human species has undergone a range of adaptations for language (cf. also Hauser and Fitch 2003). Crucially, however, he also argues that the fact that something is a universal property of language is not a sufficient reason to assume that said property is innate. Only those (universal) properties that are *invariant* over long enough a period of time can be the basis for adaptation. E-Languages are by no means stable enough to provide a target for adaptation. Languages change from generation to generation and even during a speaker's lifetime. Within the span of a few hundred years, a language may be completely transformed. Specifically, the words, phonemes and grammatical features it employs may change drastically.

Consider in this respect Cinque and Rizzi's (2010) statement: "To judge from Heine and Kuteva's (2002) four hundred or so grammaticalisation targets, the number of functional elements must at least be of that order of magnitude" (p. 57). Although Cinque and Rizzi do not state explicitly that all four hundred or so functional elements must be innate, they do claim that functional heads are universal and that we may assume they are present in any language, even in those that do not have overt evidence for them. The only way in which this is possible is when functional heads are indeed innate.

We may safely assume that there is not and never has been a single language that has overt evidence for all those four hundred or so functional categories.[24] If that assumption is correct, then following Deacon's argument, there is no basis for adaptation: these functional categories could not have become innate through a process of natural selection. The only theoretical option, if features are innate

[24] This assumption is still valid even if the number of features is much smaller, because we have yet to find the definitive set of features that describes *all* languages.

but cannot be the basis for adaptation, is that they have arisen from a single genetic change. This is not a likely occurrence, even if the actual number of features were much smaller than the four hundred that Cinque and Rizzi mention.[25]

Deacon's invariance argument can hardly be ignored. If it is correct, the human brain has specific adaptations for human language, i.e., language is not just a cultural phenomenon, as some would have it (e.g., Tomasello 2003), but there would be no UG in the sense in which it is normally understood in generative linguistics. What this means is that in terms of morphosyntactic structure and features, very little is prescribed by the language faculty,[26] something that has been explicitly assumed in other, non-minimalist frameworks (see, e.g., Müller 2013).[27]

To summarise the discussion, features are too elusive to acquire reliably (i.e., in such a way that each individual ends up with the same set of features) during first-language acquisition, and also too elusive to be the target for (Baldwinian) evolutionary processes (even those features that are clearly visible, because they change too quickly). At the same time, they are too specialised to be the result of a single mutation. It follows that they cannot be innate and therefore cannot be the primitives of syntactic structure.

This does not mean, of course, that language does not employ features. From a symbolic (i.e., high-level) point of view, features are a crucial part of grammar. At a lower level, there are clear suggestions as to how the brain might implement them. For example, Elman (1990) discusses an artificial neural network designed to be able to represent time, or, more specifically, to show how the processing of inputs can be made dependent on prior internal states of the network. Elman uses this to create a model that discovers lexical classes from word order.

Specifically, Elman trained the network with simple intransitive (N-V) and transitive (N-V-N) sentences and tasked it with predicting the next word in a sequence. The sentences were generated by a generator program on the basis of pre-assigned word categories (e.g., VERB-TRANS, VERB-INTRAN, NOUN-HUM, NOUN-ANIM, NOUN-INANIM, etc.) and sentence templates (NOUN-HUM VERB-EAT

[25] Note that Chomsky does seem to be aware of the implausibility of this idea: Hauser, Chomsky, and Fitch (2002) speculate that recursion, not features, may be the only language-specific, innate property of the human species. A single property is already a more likely outcome of a single, lucky mutation than a whole host of features, but Hauser, Chomsky, and Fitch in fact go further: they speculate that recursion may actually be an adaptation, albeit not for language. If that is the case, it raises the question what the trigger was for co-opting this property for language. Hauser, Chomsky, and Fitch do not address this issue.

[26] Note that I use the term "language faculty" metaphorically. I do not wish to imply that there is a specific part of the brain that deals with language: different aspects of language are dealt with by different brain regions, none of which are assigned exclusively to language, presumably. What is meant here is the abstract property of having the ability to speak and understand language.

[27] See also Elman et al. (1997) for a similar view.

NOUN-FOOD, NOUN-HUM VERB-INTRAN, etc.) and they were presented to the network without pauses.

After training, the activations of the hidden units in the network were examined for each word and compared. Each word was associated with a 150-digit vector, a so-called *activation vector*, where each digit in the vector represents the activation of a specific hidden unit. Interestingly, the activation vectors of two words are more similar when they share more features and more dissimilar when they share fewer features. Figure 1.3 shows the cluster diagram produced from the activation vectors of the words used in the sentences.

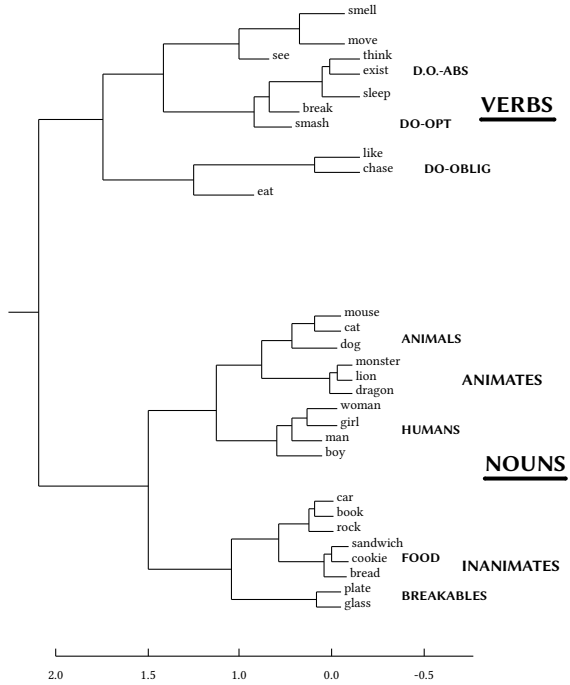


Figure 1.3: Cluster diagram of hidden unit activations (Elman 1990)

Words that are closer together in the cluster diagram have activation vectors that are more similar, whereas the activation vectors of words that are far apart are more dissimilar. The diagram shows that the network has categorised the words in the input sentences according to a number of properties. First and foremost is of course the noun-verb distinction, but within these categories, subcategories have appeared that make more fine-grained distinctions.

What this cluster diagram shows is that the network has developed into a state where, e.g., the verbs *think*, *exist*, and *sleep* have fairly similar activation vectors because of the fact that they occur without objects. Or, more accurately, because of the fact that the network could not predict with any reliability what word would follow after any of these verbs. (Remember that the sentences were presented to the network without any pause, so even intransitive verbs were always followed by a noun, which was the subject of the next sentence.)

Of course, we should not overinterpret the results of Elman's (and much subsequent) work in the field of artificial neural networks. As any researcher in this field will point out, ANNs are not direct models of the brain and the solution that an ANN finds to a given problem does not necessarily reflect the solution that the brain finds. But the experiment does show a few things that are relevant to theories of grammar.

First, even a very simple (artificial) neural network is (obviously) able to extract certain properties from its input. The sentences created by the generator contain implicit category information for the words from which they are constructed by virtue of the fact that each word occurs only in combination with specific other words. The network is able to extract this information[28] and store it as synaptic weights. What this means is that a neural network does not need much in the way of innate disposition in order to extract categorial information from a linguistic input signal, as long as the input signal contains cues that the network can, so to speak, latch onto.[29]

Another important thing that artificial neural networks teach us is that the features they come up with are not categorial in the true sense of the word, nor are they inherent to the objects or events being referred to. Rather, in Elman's network at least, they derive solely from the combinatorial properties of the words.[30] Put differently, the labels that appear in figure 1.3 are completely arbitrary and the dividing lines between the categories are not always clear. Apparently categorial distinctions, such as the one between nouns and verbs, are the result of large differences between the activation vectors for the individual members of both categories, not the result of a bit set in one or the other direction, or of a ±-value assigned to some parameter.

The upshot of all this is that the features in Elman's network are not entities in

[28] Although not in the way a researcher might, or in the way a computer might be programmed to, i.e., by tallying the words and recording which words precede and follow.

[29] In an artificial neural network, the ability to latch onto the right cues is provided by the researcher and the task to be performed. Although Elman's network was not designed to categorise words, it ended up doing so because it was the optimal solution to the problem it was facing: predicting which word would be next in the input string. It is possible, of course, that in human brains, the ability to latch onto the correct cues for language learning are provided by innate dispositions.

[30] In morphologically rich languages, the *formal* properties of words are likely to be a source for categorisation as well.

themselves. They are properties of words but they only exist *in relation to other words*. They are not reified and do not exist on their own. As such, they cannot serve as sole input for computational processes. Features cannot be activated on their own, only words can. When a word is activated, it activates its features by virtue of its activation vector, but these features are not represented separately from the words that contain them.

These remarks hold for the artificial neural network that Elman designed and trained. The question is to what extent they translate to the human brain and thus to human language. Without more detailed knowledge about the actual biological neural structures that support language, this question is impossible to answer definitively. Obviously, for the brain to extract features from the linguistic input without innate knowledge during language acquisition, the input must be rich enough for this to be possible, and furthermore, the brain's architecture must be capable of doing this. Neither requirement is shown conclusively to be true for the brain by Elman's network. The input of Elman's network is not natural language, nor is the way it is trained to learn its task a natural learning setting.

However, if a very simple artificial neural network is able to form word categories from a small corpus of sentences, we can safely assume that the much more complex biological neural network that constitutes our brain is able to do the same, even if the input is more complex. Furthermore, if the simpler network forms categories *implicitly*, it is at least a possibility that the larger network does the same. Since such a system is simpler than one that represents features explicitly, at least in terms of the amount of information that must be represented, this then becomes the null hypothesis for explaining the existence of features in human language. Postulating anything beyond that requires explicit motivation. The common assumption in minimalist theories is that there is good reason to assume more than this null hypothesis. This assumption, however, rests on the Poverty of the Stimulus argument, which, as we have seen in the previous section, is itself problematic.

We are therefore forced to adopt the null hypothesis: features are extracted from the input, and they are stored implicitly as part of lexical entries. They are not reified by the brain (except in the mind of the linguist) and cannot be the objects with which syntactic computations are made. Therefore, a model such as Distributed Morphology, or any model which takes features as primitives and computes syntactic structures with them before relating these to phonological forms, cannot be right. In other words, features cannot be the building blocks of syntax.

The fact that Elman's network learns features that might be considered 'lexical' (category, transitivity, animacy, etc.) and not 'grammatical' (tense, case, definiteness, etc.) does not change this argument. There are certainly word forms that

inherently possess 'grammatical' features (e.g., *was* or *went*; *me*, *him*, *us*, etc.) and there are languages that express 'lexical' features with regular morphology. A distinction between 'lexical' and 'grammatical' features is generally not part of grammatical theories. Current minimalist theories (among which DM) in fact assume that the lexicon consists of roots which are void of *any* features, even of categorial features (a view that is essentially a generalised version of Beard (1988) *Separation Hypothesis*).

There is, then, a major discrepancy between the conclusion reached here on the basis of Elman (1990) and what seems to be the general consensus in current minimalism. I believe, however, that this discrepancy can be resolved, an issue I turn to in the next section.

1.2.3 Representation of features and constraints

Even though I believe that features should not be considered entities in their own right, it does not follow that we should abandon features as a theoretical tool. In fact, without features, it would not be possible to formulate a meaningful theory of language at a high level of analysis. Rules and constraints must be formulated on the basis of features in order to be able to express generalisations. The structure of Elman's network can be described using a set of increasingly more fine-grained categorial features, starting with N vs. V. If we try to identify the categorial feature N in the network, however, all we can point to is the part of the activation vector that all nouns share and that verbs lacks, and vice versa for the feature V.

To make this a bit more concrete, let us consider an (extremely) simplified example. Assume an 8-bit activation vector representing a set of nouns and verbs. Suppose that every word that has the pattern `x1xx1xxx` (with `x` representing either `0` or `1`) is a noun, whereas every word that has the pattern `1xxxxxx1` is a verb. The point is that it is *not* possible is to activate *just* the two bits that constitute the N feature and leave the other bits unspecified. Any input to the relevant units yields an output activation pattern that includes *all* bits and that is tied to that particular input. This is what it means for the feature N/V to be *implicit* in the representation. For any given activation pattern it is possible to say whether it is N or V, but the network has no representation for *just* those features and is therefore unable to perform any sort of computation on it that would require the feature to be an entity in itself, i.e., without involving a particular word that carries the feature.

Obviously, it is in principle possible for a neural network to develop a representation of a feature such as N or V. What is required for this is that the network is faced with a certain kind of task that necessitates such a representation. In a connectionist network, this would simply be the task that the researcher sets for

the network and the error corrections that arise from it. In natural language, it is not clear what kind of input or task might trigger reification of features. The need to classify words categorically does not seem to be a sufficient requirement.

Another possible reason for reifying features might be the need to use them in rules and constraints. After all, rules and constraints should be able to refer to categories of lexical items and the only way to do so is through features that capture those categories. However, there is no direct indication that rules or constraints require explicit encoding of features. In fact, it is possible that rules and constraints can be encoded implicitly as well. For example, Elman's (1990) network embodies the constraint that a verb be preceded by a noun, or, as Elman puts it: "[t]he network is not able to predict the precise order of words, but it recognises that (in this corpus) there is a class of inputs (namely, verbs) that typically follow other inputs (namely, nouns)" (p. 199).[31] Input sequences in which a verb is preceded by a verb would be rejected by the network, in the sense that it would have difficulty processing it and would produce a completely unreliable output. Crucially, this is the *only* way in which the network would 'recognise' that something is wrong. The word *recognise* is in scare quotes because the network itself does not actually register the ungrammaticality of the verb-verb sequence. It just produces gibberish for a moment and then goes on to (try and) interpret the next input. That is, the constraint that a verb be preceded by a noun is not explicitly represented in the network: it is not the case that the input is checked against the constraint. Rather, the connection weights between the units encode the constraint implicitly, with the result that they are not able to correctly parse any input that does not match the constraint.

Nonetheless, despite the implicit nature of constraints and features, it does not follow that we should abandon these abstractions. As discussed in section 1.1.2, in order to fully understand a complex cognitive system such as language, it is necessary to describe and understand it at various levels, from very low-level, subsymbolic descriptions to very high-level, symbolic descriptions. The implicit nature of features and constraints means that they are difficult to see at lower levels of description, in the same way that it is difficult to see what a software package does if one studies the way electric currents flow through logic gates.

Features and constraints are therefore useful abstractions helping us to reason about the system as a whole and about the way it functions. It is important, however, that our abstractions stay as faithful as possible to the way they are implemented at lower levels. When our abstractions deviate too much from the underlying system, we run the risk of turning our model into the object of study.

[31] In actual fact, the network's knowledge is more precise than that. For example, it has learnt that verbs such as *chase, break* or *smash* are always accompanied by a specific class of nouns.

As an example of this, consider the well-known classification of word categories as combinations of [±N,±V]:

(3)

	+N	–N
+V	A	V
–V	N	P

The categorial features N, V, A and P are clearly useful in many languages (N and V probably in all languages). Using two binary features N and V to decompose these categories is an extension of the model that is motivated theoretically: it allows us to reduce four categories to two, yielding what may be considered a "simpler" theory. Furthermore, the analysis has a sense of elegance, because it appears to capture the intuition that adjectives are somehow verb-like and noun-like at the same time, while prepositions are neither.

However, there are serious problems with such an analysis. First, N, V, A and P, and consequently [±N] and [±V] as well, are *morphosyntactic* categories, whereas the observation that adjectives have both verb-like and noun-like properties is a semantic one. Adjectives typically name properties, and although properties in some sense do sit between the typical referents of verbs (i.e., events) and the typical referents of nouns (i.e., objects), it is not clear whether this observation has any relevance semantically. More importantly, however, there is no one-to-one correspondence between semantic types and syntactic categories (abstract nouns do not name objects but rather states or properties, to name just one example), which means that even if there is some semantic relevance to this observation, it does not carry over to syntax. In other words, the idea that adjectives are both verb-like and noun-like is not supported by the facts.

In fact, morphosyntactically, adjectives usually form a distinct category with clear similarities to *either* verbs or nouns, but not to both. In many Indo-European and Semitic languages, for example, adjectival inflection resembles nominal inflection, whereas in languages such as Japanese, adjectives inflect in much the same way as verbs do.

In short, the idea to reduce the categories N, V, A and P to two binary features is not based on any syntactic observation. Instead, it is a theoretical consideration, based on the assumption that having two primitive categories (N and V) is better than having four. This assumption, however, ignores the fact that binary features are different from privative features, at least as high-level analytic devices and that their implementation at a lower level of analysis may therefore be more complex. This would mean that it is not clear whether the decomposition of N, V, A and P into combinations of [±N] and [±V] is really a reduction.[32]

[32] Alternatively, it is possible that binary features must be implemented as privative features at the

The point is that the decomposition of morphosyntactic categories into the features [±N,±V] is motivated by considerations about the structure of the model, without having a real motivation either on the basis of the architecture of the lower-level system in which the model must eventually be implemented, or on the basis of the data to be described. Either of these types of argumentation constitutes a much better argument in favour of a particular analysis than a theoretical one. An exclusively theoretical argumentation turns the model into the object of study, something to be avoided if one wishes to study a real-world phenomenon.

Another way to put this is to say that having postulated features, one might speculate on the properties of a feature system and develop the model along those lines. However, not every conceivable property of (abstract) feature systems is necessarily a property of I-language. Therefore, our models should be informed by possible properties of features systems, properties of the relevant E-language, and properties of the lower level system that implements the relevant I-language.

1.2.4 The past tense debate

Another connectionist model, one that is in several ways more sophisticated than Elman's (1990), provides some hints on the way rules may be implemented in a lower-level network and —more importantly for present purposes— how we can represent rules in a higher-level analysis that should retain compatibility with lower-level analyses. Note that the caveats that hold for Elman's model apply to this one as well: ANNs are not direct models of the brain and therefore cannot tell us anything about the brain beyond showing at a certain level of abstraction how the brain *might* function. They show us what strategies neural networks in general are able to use in order to solve problems.

Incorporating such strategies into a theoretical model of grammar represents an hypothesis about how the brain functions. It is an informed hypothesis, in the sense that we know that neural networks have the ability to function in this manner, but it is really just an hypothesis. Note, however, that *any* theoretical model of grammar constitutes hypotheses about how the brain functions. In that respect, an hypothesis based on what we know about neural networks, however simple, seems better than an hypothesis that explicitly ignores knowledge about the brain.

The network I discuss in this section is developed by Westermann and Ruh (2012). Westermann and Ruh place their model in the broader discussion regard-

lower level, in which case the lower-level implementation would not be more complex, but it would not be simpler, either. I adopt a valued feature system for morphosyntax in section 2.6, on the assumption that it can be reduced to a system of privative features, if the lower-level analysis requires this.n

ing the nature of grammar and language modelling. They defend the position that neural networks do not encode rules but rather "associative processes" that produce a certain output given a certain input and argue that that is a more correct model of how language is instantiated in the brain.

A few remarks are in order here. First, the phrase "associative processes" is of course reminiscent of behaviourism, with which connectionism indeed shares certain aspects and the mentioning of which is often enough to make generativists stop listening. It should be kept in mind, however, that connectionist models are a fair bit more complex than behaviourist models and allow for extensive interaction between inputs. Furthermore, the success of functional programming languages such as Haskell in the last two decades or so suggests that systems that use "associative processes" are not limited in the effects they can produce. Reasoning about operations and processes in associative terms requires a different way of thinking than the more traditional "imperative" model, but it can be done.

Second, even though Westermann and Ruh argue against the idea that rules are instantiated as such in the brain, it does not follow, I believe, that rules (in the broadest sense of the word) are a bad abstraction for thinking about language (cf. also section 1.1.2). We do need to realise, however, that certain types of rules are more likely to be realistic abstractions than others, a point already made above but emphasised and elaborated by the model to be discussed presently.

Westermann and Ruh (2012) start out by discussing the importance of the so-called *past tense debate*:

> Over the past 25 years the English past tense has come to stand at the centre of an intense debate on the nature of language processing and of cognitive processing in general. The main question around which this debate has revolved is whether regular and irregular past tense forms are generated by two qualitatively distinct mechanisms or whether all forms are produced in a single, associative process. This question is interesting from two perspectives. First, the regular past tense is a typical linguistic rule ("to produce a past tense form, add *-ed* to the verb stem"), and debate has focused on whether such rule-like behaviour indicates the mental reality of grammatical rules or whether it can emerge from associative processes without explicit implementation of that rule. Second, regular and irregular verbs dissociate in many aspects of normal and impaired processing in both language learners and adults. This raises the question whether behavioural dissociations between two related tasks imply a modular architecture with qualitatively distinct mechanisms that can be selectively damaged, or whether such dissociations can arise within

> a single-mechanism associative learning system. The English past tense has been called the "drosophila of psycholinguistics" (Pinker 1999) because it can serve to answer these more general questions as a model system for the organisation of the language system and the mind in general. Yet, despite the vigour of the debate, the respective views remain diametrically opposed. (p. 1)

After this introduction, Westermann and Ruh discuss the history and nature of the past tense debate. In this debate, two models oppose each other, as already indicated in the quote.[33] On one side stands the idea of a dual-mechanism theory (also called the Word-and-Rules Paradigm), which assumes two mechanisms: a rule that derives regular past tense forms and a lexicon that stores the irregular forms. On the other side stands the single-mechanism approach, which argues that a single associative mechanism derives both regular and irregular verb forms.

Next, Westermann and Ruh discuss criticisms and problems with both approaches. The main criticism levelled against the dual-mechanism approach is that there are empirical data regarding the acquisition, processing and storage of past tense forms and regarding the loss of past tense after brain damage that do not mesh well with the idea that regular and irregular past tense forms are produced through two separate mechanisms. The single-mechanism approaches proposed in the literature, however, suffer from the problem that they do not get the numbers quite right. As Westermann and Ruh put it: "a full account [of the English past tense data] should provide a comprehensive model encompassing acquisition, adult processing, and selective breakdown after brain damage" (p. 2). The existing single-mechanism models show performance that approaches the empirical data, but still deviates from it.[34]

Westermann and Ruh then develop a model of past tense processing that they call *neuroconstructivist*: a connectionist network that grows new structure during learning as processing difficulties arise. In the simulations conducted by Westermann and Ruh, the model starts out as a simple two-layer model, with an input layer encoding stem forms and an output layer encoding past tense forms. Additionally, there are two hidden units, which are implemented as *receptive fields*: rather than the more common sigmoid activation functions, these units have a Gaussian activation function, which allows a unit to give a graded response to an input that indicates how close the input is to a particular prototypical input.[35]

[33] For references on both types of models, see the discussion in Westermann and Ruh (2012).

[34] Westermann and Ruh discuss a second point of critique, which is the need to postulate a causal role for semantics in the generation of irregular verbs. This point is not relevant for the current discussion.

[35] This prototypical input must not be an actual input. Rather it is the collection of properties most

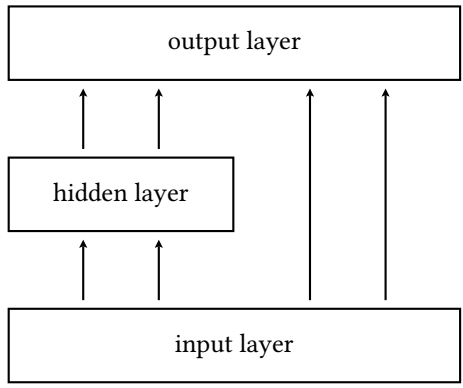


Figure 1.4: Connectionist past tense model (Westermann and Ruh 2012)

During training, the hidden layer is able to grow: as learning stagnates, new units are added to the hidden layer at the location of a hidden unit with a high local error, since such units generally lead to a high output error when activated. "Because a high local error is usually caused by one hidden unit becoming activated for verbs with conflicting past tense forms (such as *read–read* and *need–needed*, or *sing–sang* and *bring–brought*), the insertion of additional resources led to a sampling of the input space that was dense in regions with conflicting forms and less dense in regions where similar verbs shared past tense forms (such as *look–looked*, *cook–cooked*, *book–booked*)" (Westermann and Ruh 2012, p. 4).

To see what this means, consider the highly simplified representation of the end state of Westermann and Ruh's model in figure 1.4.[36] In its end state, the model has a set of direct connections between the input and output layers, and a set of connections between the input and the hidden layer plus a set of connections between the hidden and the output layer. Information flows from the input to the output layer via two pathways: one direct, and one indirect.

At first sight, this representation seems to vindicate the proponents of the dual-mechanism model: it seems that there *are* two mechanisms: one instantiated by the direct connections from input to output layer and one instantiated by the hidden layer. However, there is a crucial difference. Unlike what a dual-mechanism model might lead us to believe, it is not the case that the regular past tense forms are produced solely by the direct input-output connections and the irregular forms by the connections passing through the hidden layer. Wester-

typical for a group of similar inputs. Note that these prototypical inputs are not predefined; they develop as the network learns.

[36] The representation is mine and is a deliberate oversimplification for the sake of argument.

mann and Ruh describe it as follows: "The adult processing system […] is best described as a structured single-mechanism, multiple-representation system. Both representations in this system are based on a verb's phonological structure; one is fully distributed, and the other is quasilocalist and is activated by either a single verb or a range of similar-sounding verbs that share the same past tense form" (p. 15).

In other words, both types of verbs, regular and irregular, depend on both pathways. The distributed representation is the representation encoded on the input/output pathways, where the combined activation pattern of all input/output connections encodes a verb: that is, each verb activates *all* input/output connections. The quasilocalist representation is the representation exploiting the hidden units. It is quasilocalist because a single unit or group of units represents a single verb or a number of similar-sounding verbs, so that a single verb only activates a single hidden unit or a small group of units, but not all hidden units.

Lesion simulations performed on the network do show that the regulars tend to depend more on the direct input-output pathways, while the irregulars depend more on the hidden layer. But this is only a tendency, and crucially, the deciding factor is *not* the distinction between regular and irregular verbs. Rather, the relevant factor is the difficulty of learning the past tense of a specific verb *in the context of all other verbs in the corpus*.[37] For example, a verb such as *blink*, which has a regular past tense, is nonetheless more difficult to learn because of its phonological resemblance to verbs such as *drink, ring, sing, sink*, etc., which all change /ɪ/ to /æ/ to form the past tense and which share with *blink* the phonological chunk /ɪŋ/.

The take-away message for theoretical linguistics is that when an artificial neural network is presented with a corpus of verbs and trained to produce their past tense forms, it does not (need to) generalise a [±regular] feature. Rather, it generalises over phonological forms and creates association groups of different sizes based on those forms. Again, it would be trivial to train an artificial neural network to categorise verbs as [±regular], but it requires a learning task that necessitates such a categorisation.[38] Westermann and Ruh's network suggests that the requirement to form correct past tense forms —really the only conceivable reason why a [±regular] categorisation might be necessary in English— does not in fact necessitate such a categorisation.

The argument here is basically the same as the one made above: Westermann and Ruh's artificial neural network does not require a reification of some [±regular] feature, therefore theoretical models of language should not assume such

[37] In fact, this is not the only relevant factor. Frequency also played a role: highly frequent regular verbs were also represented in the hidden layer (p. 12).

[38] For example, the task of categorising verbs as [±regular].

a feature as a primitive. However, the argument actually goes a step further: the cluster diagram in figure 1.3 shows that even though features are only implicit, the network nonetheless makes a distinction between nouns and verbs that is practically categorial. In Westermann and Ruh's network, however, regular and irregular verbs are not even clearly distinguished. Rather, they are clustered around their phonological properties.

Therefore, unlike with Elman's network, which gave us a good reason to adopt the N/V distinction as part of a symbolic description, the current network does not offer a similar justification for a [±regular] feature. But if we cannot assume a [±regular] feature, how can we describe Westermann and Ruh's network in symbolic terms without losing its essential structure?

First, for the sake of simplicity, the statistical aspects of the network will be ignored here. Not that these are not important, but they complicate the picture and when it comes to core aspects of grammar, the statistical biases are often so strong that they tend toward being absolute. Second, the question we should be asking is not only how we can capture Westermann and Ruh's network in symbolic terms, but also which aspects of it we should try to capture. As already mentioned, we should not mistake a connectionist network for the brain. Such networks at best show us how the brain *might* solve certain problems, not how it actually does.

That being said, there are nonetheless aspects that are worth considering. In the previous section, I argued that the way Elman's (1990) network categorises nouns and verbs allows us to formulate an hypothesis about the way features are represented in the brain. Similarly, Westermann and Ruh's model shows that a simple connectionist network does not need to categorise verbs as [±regular] in order to reliably produce past-tense forms. The fact that the network's performance is statistically indistinguishable from humans in acquisition (learning) and lesion simulations only strengthens the argument that the network's solution is closer to how the brain solves this problem than the traditional analysis with a [±regular] feature.

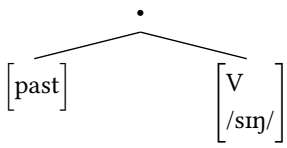
The crucial aspect of Westermann and Ruh's network is that it classifies verbs by their phonological form.[39] In natural language, this is not sufficient, of course: many more aspects about a verb need to be stored. However, the past tense form is obviously something that needs to be stored and since it is generally independent of other properties of verbs, we do not need to take these aspects into consideration for the question at hand. The only thing that we do need to consider is the fact that unlike Westermann and Ruh's network, the human brain

[39] As Westermann and Ruh point out, there have been connectionist models of past tense production that assume an important role for semantics (Joanisse and Seidenberg 1999). However, on the basis of the fact that their model performs better in acquisition and lesion simulations, they argue that semantics does not play a role.

does not uncontrollably produce past tense forms whenever it receives a verb as input. Therefore, in order to activate the part of the network that produces the past tense form, an additional activation must be present, one that corresponds to the grammatical feature [past].

Put differently, the assumption is that the activation of a verb activates (among other things) its phonological representation and, related to this, its past tense form. The past tense form, however, is only weakly activated and requires additional activation from a different source to become strong enough to win out against the network's default verb form (in English presumably the infinitive / non-3sg present tense form). In symbolic terms, we may represent this in the following, not entirely unfamiliar, way:

(4)



In this tree, the feature [past] represents the activation of the relevant property that ensures the past tense form is produced. The feature bundle [V, /sɪŋ/] represents the verb *sing*: the categorial feature V, which is in essence the activation pattern that it shares with all other verbs, and the phonological matrix /sɪŋ/, which represents the activation patterns that represent the phonological form. The question is, how can we represent the top node? Obviously, the phonological form here has to be /sæŋ/. Westermann and Ruh's network produces this form by association of *sing* with a unit (or cluster of units) that detects monosyllabic verbs with a rhyme of /ɪŋ(k)/, turning it into /æŋ(k)/. We can represent this as a schema:

(5)

$$\begin{bmatrix} \textsc{syn|head} & \textit{past-verb} \\ \textsc{phon} & \left\langle \boxed{1} \mid \text{æ, ŋ} \mid \boxed{2} \right\rangle \\ \textsc{h-dtr} & \begin{bmatrix} \textsc{syn|head} & \textit{verb} \\ \textsc{phon} & \left\langle \boxed{1} \mid \text{ɪ, ŋ} \mid \boxed{2} \right\rangle \end{bmatrix} \\ \textsc{n-h-dtrs} & \left\langle [\text{past}] \right\rangle \end{bmatrix}$$

This (partial) schema licenses the structure in (4) and changes the phonological structure from /…ɪŋ…/ to /…æŋ…/. Since this schema explicitly refers to verbs

with a rhyme of /ɪŋ/, it cannot license the combination of a verb such as *work* with [past]. That requires another schema:[40]

(6)
$$\begin{bmatrix} \textsc{syn}|\textsc{head} & \textit{past-verb} \\ \textsc{phon} & \boxed{1} \oplus \langle \text{d} \rangle \\ \textsc{h-dtr} & \begin{bmatrix} \textsc{syn}|\textsc{head} & \textit{verb} \\ \textsc{phon} & \boxed{1} \end{bmatrix} \\ \textsc{n-h-dtrs} & \langle [\text{past}] \rangle \end{bmatrix}$$

Crucially, this schema does not make reference to the phonological form of the head daughter. This allows it to license a node such as {*work*, [past]}. Interestingly, however, it would also be able to license the node {*sing*, [past]} with the form /sɪŋd/. This is of course a classic *Elsewhere-Principle* situation. The schema in (5) is more specific than the one in (6) because it makes explicit mention of a part of the phonological structure. As such, the *Elsewhere Principle* ensures that (5) will win out over (6). In other words, the two schemata are in competition.

In order to properly handle the case of *blink* ~ *blinked* we need another schema:

(7)
$$\begin{bmatrix} \textsc{syn}|\textsc{head} & \textit{past-verb} \\ \textsc{phon} & \boxed{1} \oplus \langle \text{d} \rangle \\ \textsc{h-dtr} & \begin{bmatrix} \textsc{syn}|\textsc{head} & \textit{verb} \\ \textsc{phon} & \boxed{1} \langle \text{b, l, ɪ, ŋ, k} \rangle \end{bmatrix} \\ \textsc{n-h-dtrs} & \langle [\text{past}] \rangle \end{bmatrix}$$

This schema competes with (6) and wins out over it because of the fact that it is even more specific: it specifies the entire word form, not just part of the rhyme.

This symbolic representation directly expresses the fact that there are attractors for certain kinds of verb based on their phonological forms. These attractors are the PHON features of the head daughters in the schemata. It also expresses the idea that there are no rules, just associations. The schemata are simply formal descriptions of structures and the relation between them. This state of affairs may have an interesting consequence: if incorporated into some input/output

[40] I have not indicated the assimilation of the past tense ending /d/ to /t/ that is due to the voiceless stem-final consonant of *work*. That would require a more elaborate phonological representation, e.g., one in which the schema itself contributes a segment that is not specified for [voice] but obtains this specification from the [voice] feature of the final segment of H-DTR|PHON.

system, there is no bias toward what should be the input and what should be the output. The higher node *past-verb* could be the input and have its structure decomposed into its parts, or the *verb* node plus [past] could be the input, being composed into the larger structure *past-verb*.[41]

The symbolic representation sketched here has certain consequences for the notion of lexical item. The traditional view of a lexical item holds that the lexicon stores all lexical items with all relevant morphological, syntactic, semantic and phonological information. However, Westermann and Ruh's network stores the information about past tense formation as part of the production system for past tense forms. In other words, there is no separate lexicon. Instead, the relevant information about lexical items is spread out over the entire system. This fact is mirrored in the symbolic representation by virtue of the schemata that license the various past tense forms, which are not part of the lexicon.

The schemata in (5)–(7) are represented as basic feature-value matrices typical of several representational grammar formalisms, most notably HPSG. This is not a coincidence: I suspect that HPSG's relatively conservative nature makes it a more suitable fit for abstracting the kinds of networks common in connectionist modelling. Grammatical objects in HPSG comprise syntactic, semantic and phonological information, which, apart from the advantages discussed in section 1.2.2, also prevents features such as [past] from being interpreted as heads. The notion of 'head' in HPSG is somewhat different from the notion of 'head' in minimalism, so that a feature such as [past] *could* not be a head, and moreover, [past] is simply not a complete grammatical object, since it lacks the usual semantic, syntactic and phonological descriptions.

Nonetheless, the schemata in (5)–(7) do diverge in some ways from traditional HPSG schemata. Most importantly, the schemata *compete* with each other. They all have the ability to license a structure composed of a verb plus a past tense feature, yet in any given instance of such a combination, only one schema can apply. Such competition is not normally present in HPSG models, the underlying issue being resolved by other means, such as the type system. In neural networks, however, competition is common. In Westermann and Ruh's model, for example, an input such as *blink* will activate the direct input/output pathways, but also the hidden unit(s) that encode the ɪ~æ alternation *and* the hidden unit(s) that encode /blɪŋk/ ~ /blɪŋkt/.

[41] A more complete rendering of Westermann and Ruh's network would of course require additional schemata, in essence one for each irregular verb and each subregularity, and one for each regular verb that might otherwise be captured by a subregularity. Additionally, to capture the multirepresentational and frequency-based nature of the connectionist network, extra schemata could be assumed that capture (high-frequency) regular verbs. This would mirror the actual network structure more accurately than if one were to apply standard economy principles and remove all redundancy from the system, but I will not pursue this here.

In the network, competition determines which of these pathways yields the output. The input/output pathway for *blink* encodes the past tense form /blɪŋkt/ but this alone is not sufficient for suppressing the ɪ~æ alternation: the additional hidden unit encoding of /blɪŋkt/ is required to achieve this. In the symbolic model, the first option that comes to mind to express this competition is the *Elsewhere Condition*, as discussed above. In chapter 5, this idea will be worked out in more detail.

A short remark about the element «[past]» in the schemata above is in order here. At the end of section 1.2.2, I pointed out that the view of features that Elman's (1990) model suggests differs from the view of features that is common in minimalism. From Elman's model, it follows that features are not entities in themselves and only exist as properties of lexical items. In minimalism, features *are* entities in themselves and are generally detached from lexical items ("roots"). The formal (symbolic) rendering of the past tense discussed in the previous section, and especially the tree in (4), may give the impression that [past] should be considered just such a feature. After all, it looks as if it is introduced into the tree as a head without phonological content.

However, the treelet in (4) should not be misinterpreted. The notion of 'head' has different definitions, varying between *projecting element* (e.g., Chomsky 2005, but also HPSG, etc.) and "[a] bundle[s] of features" (Embick and Noyer 2001, p. 559). Crucially, however, heads in minimalist theories are generally thought of as "the basic atoms of phrase structure", as Embick and Noyer state, just before adding that heads are "bundles of features". That is, heads are the building blocks of syntactic structure and can move, be moved to, select other heads, etc., all of which takes place without any links to phonological material, which are established later.

Above, I argued that features are properties of words and cannot be active on their own. For the feature [past] above, we need to extend this idea slightly. What remains is the idea that a feature must be linked to something observable, i.e., something with phonological content. However, this observable element does not have to be a full (prosodic) word, it may also be a partial phonological structure, essentially anything that could be described as a morph. In the case of [past], this morph would be the suffix /-d/, the ɪ~æ alternation and other subregularities, and the irregular past tense forms. That is, it is not necessary that a feature corresponds to exactly one phonological form. Depending on context, it may correspond to several different forms, and occasionally to no form at all, as in verbs such as *set* ~ *set*, whose past tense form is identical to the present tense stem.

This correspondence between [past] and the relevant phonological forms is not directly obvious from the schemata in (5)–(7), but it can be surmised from

the material in the PHON attribute of each structure that is not repeated from the PHON attribute of the head daughter. In the chapters that follow, I adopt a different notation that makes this relation more obvious.

The take-home message for theoretical linguists is that we do not have to avoid features: they are useful and necessary abstractions that help us understand and describe the system. However, we should avoid reifying them and turning them into objects of study. A feature should always be linked to a phonologically overt form, a (prosodic) word or something smaller, such as a syllable, segment or an autosegmental element. What this comes down to is that heads, as bundles of features, cannot be phonologically empty, at least not in most of its occurrences. A head that is phonologically empty in some specific cases is acceptable, as long as it remains the exception rather than the rule.

1.3 Deacon's hierarchy of reference

The conclusion reached in the previous sections, that features must be linked to some overt phonological material, finds support form a very different line of argumentation. In this section, I discuss the theory of reference developed by Deacon (1997, 2003), which points in the same direction.

Adopting terminology and insights from Peirce (1897, 1903), Deacon develops a theory in which three types of reference are ordered in a hierarchy. 'Reference' as Deacon understands it does not mean the same thing that the term means in the field of philosophy of language, as duly noted by Hurford (1988). In philosophy of language (and generally in linguistics), reference is a relation between a word and something extra-linguistic, i.e., something in the real world. For Deacon, the term is more general: it refers to any kind of relation between two entities where mentioning one entity brings to mind the other. In this sense, one object in the world can refer to another object in the world (e.g., when one sees smoke and thinks of fire), and a word can refer to another word (such as when a linguist hears the word *submit* and thinks of the word *abstract*).

These examples highlight an important aspect of Deacon's theory: reference for him is an *interpretative process*. That is, a referential relation is not something static that *is*, it is something dynamic that *happens*. Deacon distinguishes three different kinds of referential processes, which are tied together in a hierarchy of reference. The simplest form of referential process is that induced by *icons*. An iconic referential relation is a process whereby the perception of the icon brings to mind the entity referred to by virtue of the observer not distinguishing between the two.

That is, looking at an icon, say a cartoon of a famous person, brings to mind

the person depicted because in the observer's mind, he *is* looking at that person. Although he is only observing a drawing depicting a famous person, in the observer's mind, the distinction between the cartoon and the person depicted disappears. This does not mean that the observer is not able to distinguish the two, of course. It is just that in this particular instance, he (subconsciously) chooses not to. As Deacon remarks, iconic reference mainly involves *not* doing something, specifically, not distinguishing between icon and referent.

This characterisation of the notion *icon* demonstrates the dynamic nature of reference. The observer looking at the cartoon may choose to ignore the referential relationship and examine the image as an image, considering its drawing style, use of colour, composition, etc. When doing so, he no longer interprets the drawing as a referential entity. Another observer might not know the person depicted and thus never be able to interpret the drawing referentially in the first place. Therefore, there is nothing inherent to the drawing that establishes it as an icon. It is the manner in which it is interpreted that does so.

When an icon and its referent occur together on a regular basis, an indexical relation may be formed in the mind of the observer. An index is an entity that is interpreted as referring to another entity by virtue of the fact that they, as stated, occur together on a regular basis. For example, smoke may be taken as an index for fire, because the observer knows that smoke and fire often co-occur. Crucially, this interpretation is only possible if the individual instances of smoke have been interpreted by the observer as iconic references to fire. Only when this iconic relation has been established often enough, according to Deacon, can the observer establish the indexical relation.[42]

In the indexical relation, there are two entities that are distinguished as such by the observer (unlike in the iconic relation, in which they are not distinguished), whereby one points to the presence of the other. Such a relation can only be acquired if the two entities are seen as belonging together, and it is the iconic relation that establishes this. This may seem contradictory, because iconic reference requires that the observer does not distinguish between the icon and the referent, whereas indexical reference requires the observer to make this distinction. The point is, of course, that one can do both. An observer can interpret an object, image, etc. as an icon referring to something else, i.e., he may treat it as coextensive with the referent for the purpose of the communicative situation at hand, while at the same time being aware that icon and referent are two different things.

Just as indexical reference is built up out of iconic reference relations, symbolic reference is built up out of indexical referential relations, Deacon argues. A child first acquires (content) words as indices pointing to the non-linguistic elements

[42] Deacon does not discuss the question what counts as often enough.

they name. Gradually, however, in the course of language acquisition, the child recognises that certain types of words co-occur with other kinds of words, e.g., that nouns co occur with verbs, adjectives with nouns, etc. Furthermore, the child recognises at a certain point that words have semantic relations to other words. What this means is that words establish indexical relations not only with their non-linguistic referents, but also with other words. These inter-word indexical relations, Deacon argues, end up being more prominent than the indexical relations to the non-linguistic referents of words.[43]

Furthermore, these inter-word indexical relations themselves constitute indices to the corresponding relations between the non-linguistic entities (actions, objects, properties) that the words indexically refer to. It concrete terms: in a phrase such as *read a book*, the words *read* and *book* refer indexically to each other by virtue of the fact that one is a verb and the other a noun that often co-occurs with the verb *read*. At the same time, the indexical relation between these two words is itself an index to the relation between the act of reading and the object category 'book' in the world.[44]

In this way, a complex system of indexical relations arises that allows linguistic symbols (i.e., words) to refer to non-linguistic entities. However, they do so indirectly, through the system of indexical relations that they have with other linguistic symbols. That is, a word's *primary* indexical referential force is toward other words. First and foremost, a word indexically refers to other words; it is only through this system of inter-word indexical relations that a word refers to something non-linguistic. This is what constitutes symbolic reference.

A *symbol*, then, in Deacon's definition, receives its referential force by virtue of being interpreted within a web of indexical relations with other symbols which as a whole has an indexical reference to a web of entities, actions and properties in the world. Symbolic reference is thus built up out of multiple indexical references, similar to how indexical reference is built up out of multiple iconic references.

Because a symbol is embedded in a web of symbols, it can have reference even if there is no actual physical referent that corresponds to the symbol, as is the case with a word such as *unicorn*.[45] In other words, even though the complex of

[43] As a side note, Deacon faces two bootstrapping problems with this account: an acquisition problem (how does the child move from indexical relations to symbolic relations) and an evolutionary problem (how did symbolic referential relations arise in the evolution of the human species). Neither problem needs to concern us here, although Deacon (1997) offers well-argued and thought-provoking answers to both of them.

[44] Or, more precisely, in our mental map of the world.

[45] Here, it is important to keep in mind that for Deacon, *reference* does not mean what it means traditionally in linguistics. In Frege's terms, for example, the word *unicorn* has a sense, but no reference. For Deacon, *unicorn* is a symbol and as such it is characterised by having a multitude of referential relations with other symbols. Stretching the meaning of the term a bit, one might

inter-symbolic indexical references has indexical relations to the non-linguistic world, it is not necessary that *every node* in this web of relations is an index to a real-world entity. The interpretative process for symbols does not depend on such an indexical relation. Rather, it depends on indexical relations with other symbols.

Deacon's analysis is quite a radical departure from more mainstream thoughts on reference, and it would be interesting to see how well it fits with traditional grammar theory. If words are symbols and symbols are characterised by indexical relations to other symbols, how can we implement this in our model of a linguistic sign?

In fact, I believe this question is not too difficult to answer. A verb is a symbol that combines with nouns. For example, the verb *read* is characterised by the fact that it can combine with two nouns, one of which expresses the thing read, the other the reader. What we need, then, is essentially an argument frame. The lexical entry for *read* must specify that it combines with two nouns. In other words, Deacon's idea that words are able to refer as symbols by virtue of the fact that they refer indexically to other words is already part of linguistic theory, although the relevance of this fact goes beyond its usual motivation. We assume that verbs have selectional properties specifying the kinds of elements they combine with because the syntactic system requires this, but if we follow Deacon's view, the selectional properties in a sense *are* the syntax.

The minimalist view on syntax is that it is a combinatorial engine, a module that takes structural objects (lexical items or previously assembled structures) and combines them into (yet) larger structures. Following Deacon's view, however, it is the words themselves that drive syntactic structure building. There is no need for a separate computational system that manipulates words (i.e., symbols). Rather, words combine themselves. It is, simply put, not possible to activate a verb without also activating its indexical references to the types of nouns that can serve as arguments.

At first sight, such an analysis has a much stronger affinity with syntax models such as HPSG, in which lexical items are signs that have a complete argument frame, while being rather far removed from the minimalist view that lexical items are non-categorial roots that obtain their category and argument frame from functional heads. Two points are in order here, though. First, Deacon's analysis needs to be amended in one specific way, as I will discuss below. Second, the difference between HPSG and minimalism is not as large as it seems. In HPSG, lexical items are at the bottom of the type hierarchy, which means that they

perhaps think of the web of indexical relations that tie *unicorn* to other symbols as its *sense*, although that web also includes syntactic relations.

inherit many of their attributes from higher types. For current purposes, the main theoretical difference between the two models is that HPSG is a lexicalist theory while minimalism is not: in HPSG, the lexicon in an active component of the grammar, able to perform complex operations such as deriving new lexical items from existing ones, while in minimalism, the lexicon is merely a list of roots (and possibly formatives, depending on the formalism); 'lexical' operations such as derivation are done in syntax.

HPSG types such as *transitive-verb* and *intransitive-verb* are abstractions over verb types and provide an actual verb stem with the necessary selectional requirements. In minimalism, the relevant information is provided by a syntactic head, usually called *little v* because it also provides the lexical root it embeds with the lexical category V.[46] As such, the two theories share a basic assumption, namely that selectional and categorial information is not part of the core of a lexical item; rather, it is provided through some more general mechanism.

In short, Deacon's inter-word indexical relations can be modelled as rather straightforward selectional restrictions.[47] The fact that syntactic theories generally assume that selectional restrictions are not part of the core of a lexical item is problematic for this view, but note that Deacon largely ignores one important aspect in his discussion: the status of function words and grammatical markers (which I will refer to collectively as functional elements). Deacon explicitly refers to grammatical function words as indices (p. 299), not as symbols, a statement that is presumably inspired by the fact that functional elements do not refer to anything in the world; in Frege's sense, they have no referent, but otherwise, the existence of functional elements is relatively inconsequential to Deacon, because in his view, the truly interesting referential elements are symbols.

Classifying functional elements as indices makes sense: an element such as the determiner *the* points to a following noun, an auxiliary to a following infinitive or participle, etc. But from a linguistic perspective, it should be clear that this is only part of the story. Functional elements are ubiquitous in language, to the extent that it seems impossible for a language to do without them altogether. If the true power of language as a symbol system lies in the way content words establish indexical references to other content words, then there seems to be no need for functional elements. So why then do they exist?

The most likely answer to this question is of course that we underestimate their

[46] Originally, little *v* was an extension of the V projection in a so-called VP-shell analysis (going back to Larson 1988) and as such not the element providing the lexical category, which is a more recent proposal (cf. Borer 2005) commonly adopted in Distributed Morphology.

[47] The status of modifiers and specifiers (in the traditional sense) would need to be clarified. In HPSG, unlike in minimalism, they usually have a selectional feature of sorts that indicates with what sort of expressions they can be combined. Deacon's model suggests that this analysis is correct.

role.[48] Functional elements are indices because they co-occur with specific types of words. The fact that they facilitate word recognition in adults and infants (cf. Hicks 2006) supports this assumption. However, functional elements may do more than that. Take case markers, for example: a case marker not only identifies the word it is attached to as being a noun, it also indicates that the relevant noun is tied to a case assigner. In other words, a case marker carries *two* indexical references: one to the category of nouns and one to the category of case assigners.[49]

As such, a case marker not only introduces an indexical reference to a particular category of words, it also indicates how the word it indexically refers to is tied into the network of indices underlying the utterance it is contained in. For Deacon, the crucial aspect of symbolic reference is the fact that a symbol is a node in a network of indexical references. If this idea is essentially correct, functional elements actually play a central role in the process of establishing symbolic reference, since they establish the nodes in the network.

Other grammatical markers seem to function in much the same way. A tense or aspect marker, for example, is an index to verbs. It differs from a case marker because it does not indicate that its associated verb must also combine with some other element.[50] But they do have another effect: they influence the way in which the clause as a whole establishes indexical reference to our mental map of the world: whether the event expressed by the verb lies in the past, present or future, whether it is hypothetical, possible or fact, etc.

Obviously, this is only a rough description of what tense/aspect markers do, but the general principle seems to conform to the idea that functional markers establish the nodes in the network of indices required to establish symbolic reference. If this assumption is correct, then Deacon's proposal is essentially on the right track, but roots themselves are not the symbols that Deacon takes them to be. Rather, roots are simply indices, referring to whatever they name in the mental map of the world. Symbolic reference is established by functional elements that tie roots into a network of indexical references.

In this respect it is interesting that language acquisition (not just L1, but also unsupervised L2, cf. Jordens 2012) starts out with a predicate-argument system and develops a full-blown functional structure only later. This suggests that the initial steps toward a symbolic system in language acquisition are based on the

[48] Or rather, Deacon does, since theoretical linguistics is generally quite clear on the importance of functional/grammatical elements.

[49] Which entails that case assigners should be subject to two selectional restrictions in a linguistic theory. In some form or other, this seems to be the case in general. In minimalism, for example, a case head selects an NP and requires a case assigner.

[50] Finite tense markers are often different in this respect, since they tend to require the presence of a subject.

collocation properties of content words, establishing a predicate-argument structure, while at a certain stage in the development, functional elements take over, so to speak, to create a more intricate network of indexical references, establishing symbolic reference.

It is no surprise, then, that functional elements are phonologically less conspicuous than content words (i.e., they usually do not constitute prosodic words, they are often unaccented, etc.) Being less conspicuous, a language-learning child will not focus on them until it has first mastered (for some measure of "mastered") the more conspicuous parts of the linguistic input, i.e., the content words and their indexical referential properties. It *needs* to master these first, because the more complex system of symbolic reference is built on top of it. Functional elements, then, conform to the idea that language is especially well adapted to the learning abilities of the developing infant brain, an idea put forward by Elman (1990) that Deacon (1997) explicitly subscribes to.

Seen from Deacon's (amended) point of view, functional elements constitute a kind of skeleton that allows a speaker to construct a sentence, because they establish the network that enables the indices (i.e., the roots) to refer symbolically. This view is not very alien to theoretical syntax, obviously, especially for minimalist theories, which generally assume that functional heads form a skeletal structure on the basis of which a clause or phrase is built.

However, there is of course an important difference with minimalism: in minimalism, the relevant functional heads are assumed to be innate and they may be covert (i.e., phonologically null) without losing their relevance. From Deacon's point of view, functional elements cannot be phonologically null because without a phonological form, there can be no index. There would be nothing to trigger the interpretative process that establishes indexical reference. In essence, functional elements cannot be covert for the same reason that content words cannot be phonologically null. Not saying anything cannot mean "dog" or "house" or "to work". Although remaining silent may have a strong communicative effect, it cannot be used as a symbol or as an index.

There is one imaginable exception to this rule: if the absence of phonological form is conspicuous in some way, it should be possible that this absence takes on an indexical referential force. Specifically, this conspicuity would have to be the result of a paradigm: a contrast with one or more other, overt, forms that can appear in the same position. There are clear examples of this in natural language. In Russian, for instance, the genitive plural form for feminine and neuter nouns is characterised by the absence of an ending, as demonstrated in table 1.8.

Since the genitive plural is the only form in the paradigm that lacks an ending, this absence can function as an index in the same way that the other case/number

	sg	pl
nom	gor-***a***	gor-***y***
gen	gor-***y***	gor-
dat	gor-***e***	gor-***am***
acc	gor-***u***	gor-***y***
instr	gor-***oi***	gor-***ami***
prep	gor-***e***	gor-***ax***

Table 1.8: Russian noun conjugation (feminine)

endings do.[51]

This issue is related to a larger question in linguistic theory: whether or not empty elements can have grammatical force. The answer suggested by the current discussion is that they can, but only to a limited extent. In the form *gor*, the absence of an ending is obviously still able to convey an indexical reference identical to the other endings in the paradigm. A language that completely lacks case endings, however, such as English, cannot have this lack of case function in the same way as a trigger for a set of indexical references.

Instead, English presumably employs a different mechanism for establishing a noun's position in the larger network of indexical relations. It is likely that word order plays this role, since English is a strict SVO language and word order is crucial for decoding the thematic relations in a clause. However we may wish to represent this in a model of grammatical knowledge (minimalist frameworks use the so-called little *v* head for this purpose), we must remember that it is just a representation of a set of indexical relations (modelled as selectional restrictions) triggered by an *observable* property of the linguistic input.

Obviously, an in-depth discussion of this question would go well beyond the purpose of this chapter. What is relevant for us at this point is that functional elements are an important part of linguistic structure (which is hardly new) and that a functional element should in principle have a morphological form that can trigger the relevant indexical relations. If a functional element has no morphological realisation, there should at least be some observable effect from which the existence of the functional element can be surmised, such as an empty ending that is part of a paradigm or a word-order effect.[52]

[51] More specifically, I assume that the absence of an ending can function in this way as long as it adheres to the Clark's (1992) Principle of Contrast as discussed by Carstairs-McCarthy (2004).

[52] The underlying idea here is that functional elements, and features in general, have to be learnable, because they are not innate. When and under which conditions something is learnable is a far from trivial question, of course, but it seems safe to assume that a head that has an overt phonological form, or even a null form as part of an overt paradigm, is learnable. Similarly, it does not seem too contentious to say that a visible word order effect is linked to a feature of some sort.

1.4 A parallel architecture

The conclusion that the primitives of linguistic structure are complex objects bears on another discussion in linguistics: the question what kind of grammar architecture most accurately represents knowledge of language. The two prime candidates are a *sequential* model, in which syntactic structures are generated first and semantic and phonological properties are read off this structure, and a so-called *parallel* model, in which syntactic structures are built in parallel with their corresponding semantic and phonological structures.

The former type of model is of course instantiated in the standard, mainstream generative grammar model of Chomskyff. (1995) *Minimalist Program* and before it Government & Binding Theory. The latter type has been brought to prominence by Jackendoff (1997, 2002) and Culicover and Jackendoff (2005), although they were by no means the first to suggest such a model. Sadock (1992) also develops an explicitly parallel model (see also Sadock 2012) and of course representational grammar models such as GPSG (Gazdar et al. 1985) and HPSG (Pollard and Sag 1994) have much in common with parallel models.

When the building blocks of linguistic structure are complex objects, a parallel model seems unavoidable, for the simple reason that activating (e.g., *merging*, *selecting*, etc.) features without making available the phonological (and semantic) features associated with them is not possible. A purely sequential model would require that those phonological and semantic features remain 'inactive' for some time after the associated morphosyntactic features have been activated. This seems unlikely, since there is no reason why they would remain inactive, and furthermore, it is not clear what 'inactive' means in this context.

What appears to be needed is a model that treats morphosyntactic features as an inherent part of more complex objects that also comprise semantic and phonological features. Representational grammar models generally have this property: descriptions of linguistic structures in e.g., HPSG are complex objects combining (at least) semantic, syntactic and phonological features. Unification in HPSG-style grammars applies to *all* the features in the two structures being unified. Even though HPSG-style analyses tend to focus on the syntactic and semantic part of the analysis, a phonological description is nonetheless part of every structural description.[53]

However, such a strictly parallel / representational model is not absolutely required. Recent developments in minimalism suggest an alternative. A central property of current minimalist models is the idea that the derivation proceeds in *phases* (see Chomsky 2001 for the original proposal and Chomsky 2008 for

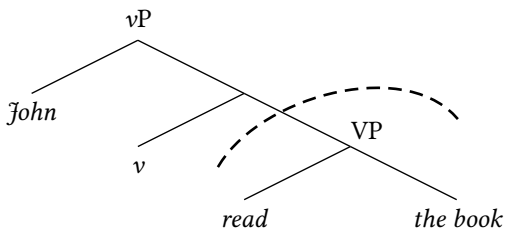
[53]Unification of phonological structures is often assumed to reduce to concatenation, which is a (deliberate) oversimplification. See Bird and Klein (1994), Höhle (1999), Klein (2000), Bildhauer (2007), Tseng (2008), and Skwarski (2009) for work being done on phonology in HPSG.

further discussion). A phase is quite literally a phase in the derivation; each phase ends with the 'transfer' of the structure built up so far to the interfaces with the conceptual-intentional (C-I) and sensorimotor (SM) systems, which convert the structure to something these external systems can handle. On its way toward the interfaces, the syntactic structure is converted to a semantic and a phonological representation, each for its own interface.

Connected with the idea of derivation by phase is the idea of *cyclic derivation*. The core of this idea is that the output of a transfer operation (i.e., the semantic and phonological structure derived from the syntactic structure that was transferred) is fed back to the syntactic system. In practice, cyclic derivation is usually applied to the phonological side of a derivation (although there is nothing in principle that would preclude it from applying to semantics), the assumption being that the (partial) syntactic structure that is transferred is replaced by the phonological representation resulting from the transfer.

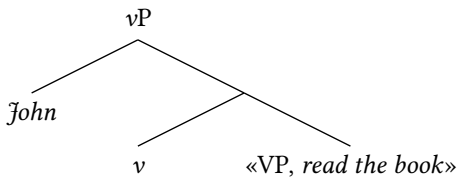
Chomsky argues that C and little *v* are the so-called *phase heads*, i.e., the heads that complete a phase of the derivation and consequently trigger transfer of part of the syntactic structure, usually their complement. That is, once a *v*P has been completed, the VP is transferred:

(8)



The resulting phonological structure is returned and inserted in the syntactic structure at the appropriate place:

(9)



The structure that is returned and reinserted into the syntactic tree is represented here as «VP, *read the book*», which is meant to indicate that the node is *atomic*, i.e., it has no internal syntactic structure. All it has that is visible to the syntactic system is its label, which is VP.

Once a node has been transferred and replaced with the resulting atomic structure, it can only be accessed as a whole. For example, in the structure in (9), the VP *read a book* can only be moved as a whole (e.g., *Read a book, John did*). The individual parts of the VP are no longer visible to syntax.

One point of discussion in the minimalist literature is the question how large the cycle is. According to Chomsky's original assumption, C and *v* are phase heads (an assumption that Chomsky has not abandoned), which means there are at most two cycles in a clause. Other proposals (e.g., Epstein and Seely 2002) suggest that there are more phase heads in a clause and that consequently the cycle is shorter. In the extreme case, when every head triggers transfer, the cycle is so short that the difference with a parallel model, at least as far as the present discussion is concerned, is negligible. A parallel model makes the phonological part of a linguistic head available simultaneously with the morphosyntactic head, while a cyclic derivational model with a short cycle makes it available immediately following the merger of the morphosyntactic head. In both cases, the phonological structure is established before the next morphosyntactic head is merged.

For our purpose, it is not necessary to choose between a derivational or a representational model, or a sequential vs. a parallel one. Our main requirement is that morphosyntactic and phonological information are available at the same time (albeit to different modules), and this requirement can be met in both types of model. Although the rest of this study is written with a clear minimalist background, my aim is to remain as theory-neutral as possible. Minimalist concepts that I appeal to are minimalist versions of general properties of language, which have equivalents in other frameworks.

Two restrictions will be necessary in order to achieve this, which already follow from the discussion in this chapter. The first of these is that I try to minimise the use of phonologically null heads. Functional heads (lexical heads are never phonologically null, even in minimalism) should be discoverable, which means that they should be overt, at least most of the time, in some form or other. Similarly, I will refrain from using movement operations that are not observable, although exactly when movement (or an LD-dependency) is observable is a more difficult question to answer. I will settle on the assumption that if an element of a specific type is placed in a position in which elements of the same type usually do not occur, and if this dislocation correlates with some specific property of that element or of the larger structure in which it appears, it is safe to assume that movement has taken place.

2

Hierarchical and linear structure

2.1 Introduction

Two types of structure are fundamental in syntactic theory: a linear structure and a hierarchical structure. The linear structure is relevant because the object of study, linguistic utterances, present themselves to us as strings of words or signs.[1] The hierarchical structure is relevant because we know that we cannot accurately describe the meaning of an utterance compositionally unless we assume that there is a hierarchical structure underlying a clause. This point is made convincingly clear by examples such as (1), a real-life variation on the traditional textbook example *The boy saw the girl with the binoculars*:

(1) Een agent heeft donderdag een man met een mes in zijn been
a officer has Thursday a man with a knife in his leg
geschoten [...]
shot [...]
'On Thursday, a police officer shot a man with a knife in the leg.'
(From the Dutch newspaper *De Volkskrant*, 8 April 2010.)

This Dutch example has a local ambiguity caused by the PP *met een mes* 'with a knife'. In the intended reading, the PP is construed as part of the noun phrase *een man met een mes* 'a man with a knife'. Syntactically, the PP can be connected to the verb phrase, however, creating an (obviously anomalous) reading in which the police officer shot a knife into the man's leg. The example shows that without reference to the hierarchical structure of a clause, we cannot fully account for its meaning.

[1] Ignoring for the moment the fact that the notion *word* is notoriously difficult to define. See e.g., Haspelmath (2011) and chapter 4 for discussion.

2.2 The Exclusivity and Nontangling Conditions

Having established that a clause has both a linear and a hierarchical structure, we must answer the question how the two are related. Across theories, there are essentially two types of approaches to answer this question: parametrisation and universal ordering. Theories such as GPSG and HPSG assume that there is a set of so-called *Linear Precedence* rules that determine the relative order of sister nodes, depending on various characteristics of those nodes. Such an approach attempts to formulate (language-specific) generalisations and allows for parametrisation of particular heads. Similar approaches exist in minimalist frameworks, but the dominant view on linearisation in minimalism is shaped by Kayne's (1994) influential *Linear Correspondence Axiom* (LCA), which essentially states that if α c-commands β in the hierarchical structure, α precedes β in the linear structure. The LCA is assumed to hold universally and is an attempt to account for certain word order asymmetries that appear to occur cross-linguistically.

Whatever the merits and problems of these two types of approaches, they have two basic assumption in common, which have been formalised by Partee, Meulen, and Wall (1993, p. 442) as the *Exclusivity* and *Nontangling Conditions*:

(2) a. *The Exclusivity Condition*:
In any well-formed constituent structure tree, for any nodes x and y, x and y stand in the precedence relation P, i.e., either $\langle x,y\rangle \in P$ or $\langle y,x\rangle \in P$, if and only if x and y do not stand in the dominance relation D, i.e., neither $\langle x,y\rangle \in D$ nor $\langle y,x\rangle \in D$.

b. *The Nontangling Condition*:
In any well-formed constituent structure tree, for any nodes x and y, if x precedes y, then all nodes dominated by x precede all nodes dominated by y.

The Exclusivity Condition states that for any pair of nodes in a tree structure, one either dominates the other or it precedes the other. This guarantees that a linear order is defined between any two terminal nodes. The Nontangling Condition ensures that if a node dominating a structure «abc» has a sister node dominating a structure «d», the linear structure will be either «abcd» or «dabc», but not *«adbc» or *«abdc».

It should be noted that the constituent structure trees that form the basis for the Exclusivity and Nontangling Conditions are fully specified as to both dominance and precedence. Arguably, such trees are not compatible with more recent linguistic theories, because these place hierarchical and linear structure in different modules. There are alternative principles in the literature, however, that

are formulated as principles governing the relation between hierarchical and linear structure and that have the same effects as the Exclusivity and Nontangling Conditions.

First, Kayne (1994) adopts the principle of *Totality*:

(3) *Totality*:
Given a tree K and the set T of terminals in K:
$\forall x,y\ (x,y \in \mathrm{T} \wedge x \neq y \mid x<y \vee y<x)$

Totality has the same effect as the Exclusivity Condition: it ensures that for any two terminal nodes in a tree a precedence relation is defined between them.[2] Similarly, the Nontangling Condition has an equivalent that links hierarchical and linear structures in a principle proposed by Ackema and Neeleman (2004), *Linear Correspondence*:[3]

(4) *Linear Correspondence*:
If a node X is structurally external to a node Y, then Φ(X) is linearly external to Φ(Y).

What Linear Correspondence says is that if a node X is not contained in a node Y (i.e., not dominated by it), then the phonological component of X is outside the phonological component of Y. It therefore has the same effect as the Nontangling Condition, but just like Totality, Linear Correspondence is formulated on the basis of separate hierarchical and linear structures, so that it avoids the technical problem inherent in the former condition.

Both Totality and Linear Correspondence (and consequently the Exclusivity and Nontangling Conditions) make one fundamental assumption, which is, however, kept implicit. Both principles assume that the terminal nodes are associated with *segmental* material. More specifically, they assume that all terminal (phonologically non-null) elements are all realised on the same phonological tier. To the extent that this assumption is correct for a given combination of phrase structure tree and corresponding linear structure, both principles hold.

However, this assumption does not hold universally. There are cases in which a phonologically non-null head is realised on an autosegmental tier. Take the following example from German Sign Language (*Deutsche Gebärdensprache*, DGS):

[2] Note that *precedence* does not mean *immediate precedence*. Other elements may intervene.

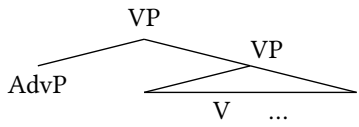
[3] See also Sadock (1992) for a similar principle. The notation Φ(X) (which is a slight adaptation from Ackema and Neeleman's original notation) refers to the phonological component corresponding to the morphosyntactic head X.

(5) **face:** *with effort* *DGS*
hands: STUDENT SIGN-LANGUAGE LEARN
'The student learns sign language with difficulty.'
(Leuninger 2005)

The words in small caps in this example are glosses for manual signs. It is a crucial property of sign language, however, that meaning is not conveyed by manual signs alone. There is a range of so-called *non-manuals*, elements that are not realised with the hands but which are nonetheless relevant for the meaning of an utterance. These elements include facial expressions, body and head position, eye and brow movements, etc. Some have a lexical (adverbial) meaning, some have grammatical functions, others have functions that are comparable to prosody in spoken languages (cf. Sandler and Lillo-Martin 2006; Vermeerbergen, Leeson, and Crasborn 2007).

The example in (5) shows a non-manual adverbial *with effort*, which is realised as a facial expression. The adverb is part of the meaning of the utterance: leaving it out would change the utterance's meaning. Syntactically, then, the adverbial has to be represented in the normal manner, as an AdvP projection adjoining to the VP:[4]

(6)



If this is indeed an accurate representation of the hierarchical structure behind the utterance in (5), we are forced to conclude that the Exclusivity Condition is violated by the adverbial. In the hierarchical structure, AdvP does not dominate V, so the Exclusivity Condition states it should be linearly external to V, i.e., either precede or follow V. Since the adverbial and the verb are realised simultaneously, this is not the case.

Since the non-manual adverb is realised simultaneously with just the verb, one might be tempted to argue that this is a morphological issue, not something that syntax should have to deal with. The same cannot be argued for the example in (7), however, which is an example from American Sign Language (ASL). The crucial property is the negation, which is realised as a head shake accompanying the entire VP, not just the verb:

[4]Or *v*P, depending on the theoretical model. I will mostly abstract away from such details.

(7) **head:** *neg* *ASL*
hands: JOHN BUY HOUSE
'John is not buying a house.'
(Neidle et al. 2000, 44ff)

Again, the negation is an integral part of the clause, the meaning of the utterance changes if it is left out. Therefore, the autosegmental negation must be represented in the syntactic structure as a Neg head adjoining to VP. Again, Linear Correspondence is violated.

One could of course argue that because the negation is not segmental, Linear Correspondence does not really apply and is therefore not violated. But in a way, that is exactly the point: in its formalisation in (4), Linear Correspondence is not meant to deal with cases such as (6) and (7). As such, however, it is insufficient to describe all aspects of the mapping between hierarchical and linear structure. The deeper reason for this problem is the common assumption that as far as syntax is concerned, only a single phonological tier, the one with segmental material, is relevant. The two examples presented here show that this assumption is incorrect: we need to allow for the possibility that autosegmental tiers are relevant for syntax.

2.3 The Separation Hypothesis

In order to do this, we need to clearly distinguish between the morphosyntactic and phonological features of a structure. This distinction goes back at least to the *Separation Hypothesis*, which was proposed by Beard (1988) to distinguish functional elements such as {PL} from the morpho-phonological elements that express them. In English, for example, {PL} is expressed by the suffix /z/ and its allophones /ɪz/ and /s/, but there are a number of other elements and morphological processes that can mark plural (e.g., *-ren* in *child ~ children*, *-en* in *ox ~ oxen*, vowel change as in *foot ~ feet*, etc.) In essence, the Separation Hypothesis expresses the observation that there is no clear one-to-one relation between form and function.

This observation is acknowledged in most syntactic models, either implicitly or explicitly. In HPSG, syntactic and phonological features are clearly separated, being different attributes of a sign; in Distributed Morphology, *Late Insertion* (i.e., the late insertion of morpho-phonological features into the structure) is one of the fundamental principles of the model; in various parallel grammar theories (e.g. Ackema and Neeleman 2004; Culicover and Jackendoff 2005; Sadock 1992, 2012), syntax and phonology are separated in the same way. In minimalist analyses, probably due to minimalism's strong focus on syntactic structure, the dis-

tinction is often not made explicitly, but it is nonetheless a fundamental aspect of the theory. Crucially, and unlike theories such as HPSG, minimalism assumes that morphosyntactic features are initially completely detached from phonological features: the building blocks of syntactic structure are morphosyntactic feature bundles *without* links to phonological structure.

However, as discussed in chapter 1, there are various reasons to assume that the building blocks of syntactic structure are not just bundles of morphosyntactic features without phonological content. Whether we assume true linguistic signs or a cyclic derivational system with a small cycle is not really relevant to the current discussion, but it is crucial that we assume that morphosyntactic features are generally tied to a set of phonological features. Therefore, I will adopt a notation for lexical entries that is borrowed from Jackendoff (1997, 2002) and that combines the semantic, syntactic and phonological parts of the entry in a single representation:

(8) $\lambda x(\mathbf{car}(x)) \leftrightarrow \begin{bmatrix} \text{N, sg} \\ \text{count} \end{bmatrix} \leftrightarrow /\text{ka}\text{ɪ}/$

(8) gives the lexical entry for the word CAR: a semantic predicate, which is linked to a bundle of morphosyntactic features, which is in turn linked to a phonological string. The double arrows are meant to indicate that the different representations are linked to each other.

In what follows, I will use the term *component* to refer to the three parts of a lexical entry. As such, a lexical entry has a semantic, a (morpho)syntactic and a phonological component. I will also sometimes speak of "the phonological component of X", where X refers to a syntactic head. This phrasing is not entirely accurate, as the syntactic head X *corresponds* to, rather than *has*, a phonological component, but it is convenient. I will also use the abbreviations Φ(X) and Σ(X) to indicate the phonological and syntactic component of X, respectively. Here, too, I will be somewhat lax with regard to the X, allowing myself both Σ(*car*) and Σ(/kaɪ/), and similarly Φ(*car*) and Φ(N).[5]

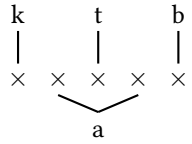
It is important to see that this assumption does not give up the Separation Hypothesis. The essence of the Separation Hypothesis is not that morphosyntactic and phonological features are linked only at the end of or after the derivation. The Separation Hypothesis merely states that morphosyntactic and phonological features are distinct. It does not state anything about the manner in which the two are linked.

[5] I have no corresponding symbol for the semantic component, since semantics is left out of the discussion.

2.4 Autosegmental phonology

It is a common assumption in phonological theory that phonological structure is best described as a collection of independent but interrelated, so-called *autosegmental*, tiers. This idea, introduced by Goldsmith (1976) to describe tone phenomena, has been used very fruitfully by McCarthy and Prince (1990, 1996) to develop their theory of *prosodic morphology*. Consider the well-known Arabic example of *kataba* 'he wrote' (cf. McCarthy 1981, 1986):

(9)



The form *katab* consists of two morphemes, a root /ktb/, which carries the general meaning 'write', and the vowel /a/, which marks a verbal stem with perfective aspect (which, unlike its Slavic namesake, is usually associated with past tense).[6] The segments of these two morphemes are associated with a timing tier, containing slots that are represented by the crosses (×). The segments are associated with the timing slots from left to right, a principle that is known as *Left-to-Right Association*. This association obeys the phonotactic rules of the language in question. In the current example, this means that the /t/ of the root cannot be associated with the second slot, because that would create a syllable with a *t* as nucleus, which Arabic does not allow.[7]

For this reason, the second slot is taken by the morpheme /a/ and the consonant *t* is associated with the third slot. the fourth slot cannot be occupied by the /b/, for same reason as before: it would violate the phonotactics of Arabic. There is no free element that can occupy this slot, however. It needs to be a vowel, but the only available vowel is associated with the second slot. This problem is solved through *spreading*: the vowel /a/ associated with the second slot spreads to the fourth slot. In general, a slot that is not associated with any segmental material is associated with the first segment on its left that it is compatible with.

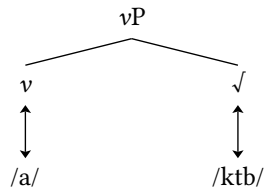
The morpho-phonological structure of the form *katab* is relatively complex (although more complex structures exist in Arabic), but its morpho-syntactic structure is straightforward. The form consists of a root and a verbalising morpheme, roughly along the following lines:[8]

[6] Note that the final *-a* in the form *kataba* is a marker for 3rd person masculine singular. It is irrelevant in the current discussion.

[7] In fact, few languages do, although Berber, a distant cousin of Arabic, is an exception.

[8] I use double arrows in tree structures as well, to indicate that the phonological form is not part of the syntactic structure.

(10)



Note that as far as syntax is concerned, the structure in (10) is complete. There is no need to combine the root and the verbaliser. Combining the two *morphemes* /a/ and /ktb/ (*not* the heads *v* and √) is taken care of by the phonological system: the morphemes are phonological chunks and it is the phonological system that is responsible for constructing a phonologically licit form out of them, using the procedure just described.

This is an important point to stress. If we adhere to a strict separation of syntax and phonology, certain traditional instances of head movement that primarily have the effect of combining a root with its affixes become unnecessary, since their effects can be derived in the phonological component.[9] Moving the root node in (10) to *v* results in a structure that still only consists of √ and *v*, which is not significantly different from the original structure.

2.5 Bare phrase structure

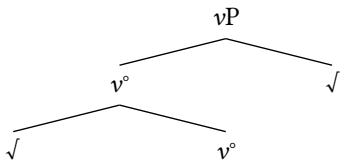
Let us consider this point in a bit more detail. First, it should be kept in mind that the notions *minimal* and *maximal projection* are derivative. A minimal projection is a node that does not branch, while a maximal projection is a node that does not project any further. In this sense, the node √ in (10) above is a minimal and at the same time a maximal projection, while *v* is a minimal projection and *v*P a maximal projection.

In minimalism, this notion of phrase structure, known as *bare phrase structure*, is first discussed by Chomsky (1995a,b), who introduces it as a replacement for X′-theory. Note, however, that this conception of phrase structure is essentially identical to what is assumed in HPSG. In minimalism, a head stops projecting when its features have all been checked, in HPSG a projection is maximal when all of its selectional restrictions have been met, which amounts to the same thing.

Returning to the structure in (10), movement of √ to little *v* yields the following structure:

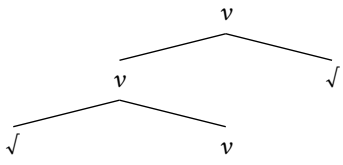
[9] See section 3.2 for discussion.

(11)



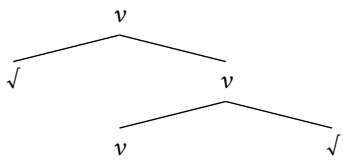
Here, the phonological component has been left out altogether and it is assumed that the movement yields an instance of head adjunction, as indicated by the minimal projection marker ° after *v*.[10] This representation is problematic, however: in bare phrase structure. the intermediate *v* node *cannot* be a *v*° node. It is an intermediate node, and hence by definition not a minimal projection. This means that a more accurate representation of the tree is (12):

(12)



Here, the upper *v*P has also been changed to just *v*, since it is obvious that it is a maximal projection due the fact that it does not project any further. What is interesting about this tree is that it is the mirror image of the tree that would arise if √ had moved to the root of the original tree in (10):

(13)



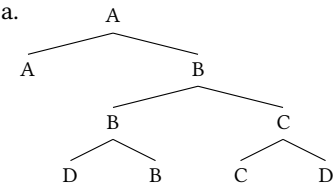
Since syntactic structures are hierarchical and no linear precedence is defined among sister nodes, the trees in (12) and (13) are identical. The only difference between them is their derivational history, but if we adopt Chomsky's (1993) *copy theory of movement*, which is the standard theory of movement in minimalism, the two √ nodes are identical, which means that it is not possible to tell which of the two copies was merged earlier.[11]

[10] Note that head adjunction is actually problematic in minimalism. See section 3.2 for discussion.

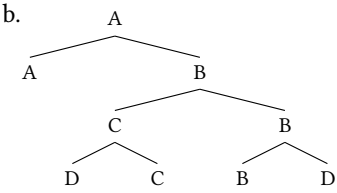
[11] The term *copy theory of movement* is the cause for some confusion, as witnessed by the remark in Chomsky (2008, fn. 17), presumably because it implies that movement (internal merge) creates a *type*-identical copy. Chomsky's original proposal, however, is that an element that is moved is not copied but simply merged a second time, yielding two *token*-identical copies.

In other words, in bare phrase structure, head movement has the problem that it obscures the derivational history of a structure. This is a general problem, as can be seen in the following sample trees:

(14) a.



b.



These trees are identical except for the linear order of sister nodes. That is just a cosmetic difference, of course, because the trees do not define linear order between sister nodes. By convention, the tree in (14a) represents a merge sequence of D-C-B-A (i.e., D and C are merged first, then B is merged, and then A). Additionally, D has been head-moved to B, yielding (in pre-bare phrase structure terms) a complex head. However, if the tree is displayed as in (14b), convention suggests that the merge sequence is D-B-C-A, with D head-moving to C.

The cause of this confusion is the convention to build trees toward the left: in (14b), C is to the left of B, which we interpret as meaning that $[_{B}$ B D] is constructed first and that C is merged with this node. This convention, however, is not the actual problem, it is merely a symptom. The actual problem is the notion of head adjunction within bare phrase structure. In essence, we can say that head adjunction is impossible in bare phrase structure, since it obscures the derivational history of the structure, which is a fundamental problem, at least in a derivational framework, because the derivational history determines the structure itself.[12]

The conclusion is that head adjunction as a means to join two heads morphophonologically is not a feasible syntactic operation. Furthermore, it is important to see that the supposed result of head adjunction, a complex head, is also undefinable in bare phrase structure.[13] It should be noted that this does not mean that head movement is completely impossible in bare phrase structure. There are alternative analyses, such as reprojection (cf. Georgi and Müller 2010; Surányi 2005) or movement to the root followed by morphological merger (Matushansky 2006), which render head movement consistent with Chomsky's proposals regarding phrase structure. (See section 3.2 for discussion.)

[12] I will not go into the question to what extent similar issues arise in representational frameworks, nor will I talk about alternative analyses for head movement here, such as movement to the root followed by morphological merger, parallel workspaces, etc. See section 3.2 for some discussion.

[13] See section 4.3.2 for some discussion.

The point here is that in the case of (10) above, and in a number of other cases, it is not necessary to assume head movement. The relevant cases are those where a *syntactic* trigger for the movement operation is not available and head movement is traditionally assumed to take place to adjoin a base to a suffix. In these cases, I will argue, phonology takes care of combining the two elements (the root and the perfective marker in the current example). The important question is how phonology 'knows' which elements to combine. This question is discussed below in section 2.7.

2.6 Valued features

Ever since "Derivation by Phase" (Chomsky 1999), Chomsky has been moving toward a valued feature system. Instead of uninterpretable features, this system employs *unvalued* features as the triggers for syntactic operations. In this system, a feature is essentially a named attribute that requires a value, much like attributes in frameworks such as HPSG. Once a feature has obtained a value, it can no longer serve as a trigger for syntactic operations.

In HPSG, the use of attribute-value pairs rather than privative features is motivated by the desire to formulate phrase structure rules in a way that makes them as general as possible. Take the basic phrase structure rule for a clause in (15):

(15) S → NP VP

This rule simply states that a sentence consists of an NP and a VP. The NP is the subject of the sentence, which means that in many languages, we need to model agreement between the NP and the VP. One way to do this would be to use specialised symbols for N and V:

(16) a. S → NP_1_sg VP_1_sg
b. S → NP_2_sg VP_2_sg
c. S → NP_3_sg VP_3_sg
d. S → NP_1_pl VP_1_pl
e. S → NP_2_pl VP_2_pl
f. S → NP_3_pl VP_3_pl

Although this works, it becomes unwieldy very fast. For a language that has distinct verb forms for, e.g., three persons and two numbers, we would need six

different N symbols and six different V symbols. If the language has morphological case at the same time, the number of symbols becomes even larger.

The obvious solution to this problem is to use placeholders for the person and number features:

(17) S → NP(Prs,Num) VP(Prs,Num)

The actual values for the placeholders Prs and Num are irrelevant. What is important is that they match throughout the rule. If Prs on the NP has the value 1, then Prs on the VP should have the value 1 as well. These placeholders are the attributes in attribute-value matrices.

Minimalism, which does not rely on phrase structure rules, could in principle do without the placeholders and instead use privative features. Chomsky's reason for moving to a valued feature system has to do with the undesirability of the notion *interpretability*. In early minimalism, features that were uninterpretable at the interface had to be deleted during the derivation, which was thought to be the driving force of the syntactic derivation. However, whether a feature is interpretable or not is something that syntax itself cannot determine. Instead, it would require a form of look-ahead, which is deemed conceptually undesirable. Using unvalued features solves this problem, because it is possible for the syntactic system to determine whether a feature has a value or not without any look-ahead toward the interfaces.

A valued feature system also has a conceptual advantage, because it makes it possible to express certain generalisations that a privative feature system cannot express. For example, if one assumes that nouns have a CASE feature that can have one of the values *nom, gen, dat*, or *acc*, it is immediately obvious that the noun can only have *one* case feature: it is not possible to assign two values to CASE. In a privative system, which does not have CASE but only NOM, GEN, DAT and ACC, it is in principle possible to assign more than one of these to a noun. Furthermore, a binary feature system (which is a valued feature system in which features can have the values '+' or '–') allows the system to refer to the minus-value of a feature, in essence a way to refer to the absence of a feature. This, too, is not possible in a privative feature system. If nominative is expressed as a privative feature NOM, it is not possible to refer to all nouns not having nominative case as a natural class.

Despite these differences, I do not believe there is a big distinction between privative and valued features in terms of lower-level plausibility. A higher-level description of Elman's (1990) network discussed in section 1.2.2 could make use of privative features such as $V_{DO\text{-}ABS}$ (for intransitive verbs), $V_{DO\text{-}OPT}$ or $V_{DO\text{-}OBL}$ (for optionally and obligatorily transitive verbs), or it could make use of valued features that specify a category V and an argument list that is either empty, or

optionally or obligatorily filled. An observation such as the fact that a noun can only have one case assigned to it does not necessarily require an explanation that is part of the grammar. It could be a meta-principle, i.e., a principle that the grammar adheres to without encoding it. Furthermore, if one needs to refer to all non-nominative cases as a natural class, it is always possible to assume a feature NNOM or OBL that captures the relevant nouns.

In other words, a valued feature system may be more elegant for a higher-level analysis, but this does not mean that it poses a problem for a lower-level analysis. At the very least, it can be reduced to a privative feature system, which, as Elman's (1990) network suggests, can be implemented at a lower level.

2.7 Phonological composition†

Any theory of syntax requires the existence of a process that construes syntactic structures. Such a process is either explicitly included in the theory (in derivational frameworks; as the operation *merge* in minimalism) or explicitly excluded (in representational frameworks). In either case, however, the theory must assume that syntactic structures are composed in some way or other. We may call this process *syntactic composition*, because it composes larger syntactic structures out of smaller ones.[14]

In a similar way, phonology needs an underlying process of *phonological composition*, which takes chunks of phonological structure that correspond to syntactic structures (i.e., morphs) and puts them together to form a larger phonological structure and ultimately the entire utterance. Just as in syntactic theory, this process may be explicitly included in phonological theory or explicitly excluded from it.[15] Deciding the merits and feasibility of either option is beyond the scope of the present study, but we can get an impression of what this process should comprise by looking at what it must accomplish.

2.7.1 Arabic verbal nouns

The Arabic verbal noun, or *masdar*,[16] is a good example to use, because its morphological form is fairly complex, whereas its syntax is straightforward. Arabic verbal nouns share a number of properties with the English gerund and they have

†Parts of this section are taken from Kremers (2012b), with adaptations.

[14]Note that thinking of this as a process that actually construes complete syntactic structures may well be a high-level analytic convenience that has no direct equivalent at lower levels of analysis.

[15]In current phonological theory, it seems to be excluded, though generally not explicitly.

[16]In Arabic, the form is called *maṣdar*, pl. *maṣādir*, which literally means 'source'.

often been analysed along the same lines (cf. Fassi Fehri 1993; Malouf 2000). Their main properties are:

- Regular form (in most verb classes).
- Event structure.
- The subject takes genitive case.
- The object takes genitive case when no subject is present.
- Otherwise, the object takes accusative or PP.

Having event structure, masdars are typical *complex event nominals* in Grimshaw's (1990) terms.[17] If a subject is present, the object can be licensed with accusative case or with the preposition *li* 'to, for'. (18) shows an example of the masdar *intiqād* 'criticising', which is derived from the verb *intaqada* 'to criticise' (examples are from Fassi Fehri 1993):

(18) ʔaqlaqa-nī -ntiqād-u -l-rajul-i -l-mašrūʕ-a
annoy.3sg.m-1sg.OBJ criticising-NOM DEF-man-GEN DEF-project-ACC
'The man's criticising the project annoyed me.'

The object of the event is *al-mašrūʕ* 'the project', and has accusative case. This suggests an analysis along the lines of Abney (1987), in which the verbal noun starts out as a verbal root and is converted to a noun by a nominalising affix. The example in (18), which I will refer to as the *accusative masdar* construction, is the equivalent of Abney's *Poss-ing* construction, in which the subject is expressed as a nominal possessor and the object with accusative case. The *prepositional masdar* construction, the equivalent of Abney's *Ing-of* construction, uses the preposition *li*:

(19) ʔaqlaqa-nī -ntiqād-u -l-rajul-i li -l-mašrūʕ-i
annoy.3sg.m-1sg.OBJ criticising-NOM DEF-man-GEN to DEF-project-GEN
'The man's criticising of the project annoyed me.'

I will focus on the accusative masdar construction here, but everything that is said here applies to the prepositional masdar construction as well.

McCarthy and Prince (1990) propose that verbal nouns in Arabic actually contain four different morphemes: a root, a stem marker, a non-finiteness marker and a nominaliser. For *intiqād*, these take the following forms:

[17] A masdar form can also be used as a simplex event or result nominal, although traditionally, the term *maṣdar* does not apply to this use.

(20) Root: /nqd/ Nominaliser: /i.a/
Stem VIII: $(\sigma)\sigma_{\mu}$ Non-finite: $-\sigma_{\mu\mu}$
|
t

Roots in Arabic are category-neutral; that is, the root /nqd/ occurs in both nouns and verbs. The element that turns the root into a verb is the stem marker. Each root in Arabic yields up to 15 possible verb stems, which are numbered I–XV. The verb *intaqada* is a stem VIII form, which is characterised by a *t*-infix after the first root consonant, here /n/.

McCarthy and Prince argue that the stem markers are in fact prosodic morphemes. In the case of stem VIII, the marker consists of an extrametrical syllable (indicated in (20) as a sigma in parentheses: «(σ)») followed by a short syllable (indicated as «σ_{μ}», where the subscript μ stands for a single mora). The extrametrical syllable only has a consonant position, which is syllabified in a post-lexical phonological process with the preceding word or with an epenthetic /i/. The second (full) syllable is associated with a segment /t/, which occupies the syllable's onset position.

The verb stem marker and the root combine to form a verb stem. This verb stem is then combined with a non-finiteness marker, which is also a prosodic morpheme, consisting of a long syllable (as indicated by the two morae in the subscript). This form is then combined with the nominaliser, which serves the same function as the suffix *-ing* in Abney's analysis of English gerunds. The nominaliser consists of two vowels /i.a/, which must end up in different syllables, since Arabic does not allow two vowels in one syllable.[18]

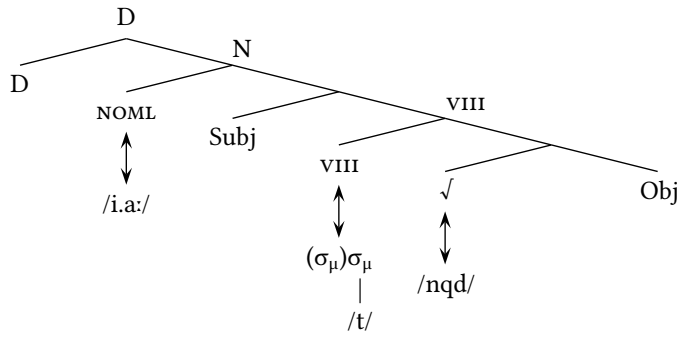
McCarthy and Prince assume that the non-finiteness marker and the nominaliser are two separate morphemes, but in Kremers (2012a) I argue that it makes more sense to treat them as a single morpheme.[19] I therefore assume that the long syllable is part of the nominaliser, which can then be represented as /i.a:/.

The syntactic structure of the masdar construction in (18) is the following:

[18] Arabic does have the diphthongs /aw/ and /aj/, but they are combinations of the vowel /a/ plus an approximant.

[19] Among other reasons, because it does not appear in participles, which are the only other category of non-finite verb forms in Arabic. In fact, not even all masdar forms have a long syllable.

(21)



The important point of this tree is that the three morphemes —the nominaliser, the verbaliser and the root— do not form a distinct subtree. In standard views, this would be problematic, because they must be combined into a distinct (prosodic) word and in order for phonology to be able to do this, they must be linearly adjacent. Under the Nontangling Condition or Linear Correspondence, linear adjacency can only be achieved if the three morphemes form a distinct subtree in syntax. Derivational theories generally achieve this by assuming head movement, which is problematic in its own right, as noted above.[20]

However, on a strong interpretation of the Separation Hypothesis, maintaining a strict separation between syntax and phonology, it is not actually necessary to combine the three morphemes in syntax. We only need to make sure that their phonological forms can be combined successfully in phonology. Phonology 'knows' that the three morphemes are to be combined into a single word form, but there is no need to construct a distinct subtree in order to achieve this. To see why this is the case, we need to consider what an *affix* is under the Separation Hypothesis. Informally, an affix is an element that attaches to a stem and that in general does not constitute a separate prosodic word. Note, however, that this definition of an affix only makes sense if we take the phonological side into consideration: what we see happening is that two phonological chunks combine to form a larger unit.[21]

This is often interpreted as a requirement that the affix attaches to the relevant head *in syntax*, but Ackema and Neeleman (2004) note that the observation is really a *phonological* requirement. In syntax, there is at best a categorial selectional restriction: an affix such as the English gerund suffix *-ing* has a syntactic

[20] Representational theories usually assume that the relevant structure is derived through lexical rules. Such an approach raises the question which structures are lexical and which are syntactic, however.

[21] In most cases, this larger unit is still a prosodic word, although occasionally, it may be something larger: some affixes have PWd status, resulting in a structure that behaves in some ways as a compound (Booij 1995, p. 169; Ackema and Neeleman 2004, p. 149-152, cf.).

and a phonological component. The syntactic component is a nominal head N with a categorial selectional restriction [*u*V]. The point is that in order to satisfy this restriction, the gerund head does not have to attach to a *head* V°. It can also attach to a *projection* of V.

The phonological component of the gerund head *-ing* is the syllable /ɪŋ/. This syllable is subject to a phonological requirement that it attach to a prosodic word, specifically, to the right boundary of a prosodic word, since it is a suffix. Together, the syntactic c-selectional restriction and the phonological restriction create the illusion that *-ing* must attach to a head.

Note that it is not impossible for *-ing* to attach to V° (or any affix to an X° category, for that matter). The syntactic attachment site of an affix does not have to be fixed. Adopting Abney's (1987) analysis of gerunds, Ackema and Neeleman argue that *-ing* can attach to V°, in which case the head is nominalised and the object of the gerund must be licensed within the NP with *of*. Alternatively, *-ing* can attach to VP, yielding a configuration in which the object can be licensed in the VP and consequently takes accusative case.

The two requirements on *-ing* pose a potential problem: if the syntactic component attaches to a projection of V, it is not immediately clear how the phonological requirement can be met, because *-ing* is possibly in the wrong position to take Φ(V) as its phonological host. In order to ensure that an affix is placed properly in syntax and in phonology, Ackema and Neeleman (2004) adopt a second mapping principle, which they call *Input Correspondence*:[22]

(22) *Input Correspondence*:
If A selects (a projection of) B, Φ(A) selects Φ(B).

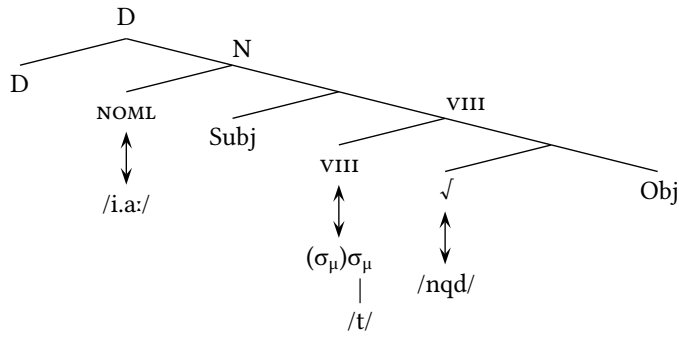
Like Linear Correspondence, Input Correspondence depends on a separation of hierarchical and linear structure. Input Correspondence ensures that even if Σ(*-ing*) attaches to the VP, the phonological component /ɪŋ/ must attach to the verb stem. In Ackema and Neeleman's proposal, this is only possible if the structure is linearised in such a way that the verb stem and the affix are adjacent and in the correct order.[23]

The Arabic masdar shows that this adjacency requirement does not apply when the relevant morphemes are realised on different autosegmental tiers. Consider again the tree in (21), repeated here:

[22]Like Linear Correspondence, Sadock's (1992) proposes a very similar idea, the *Constructional Integrity Constraint*.

[23]In fact, this additional requirement prompts Ackema and Neeleman to argue (following Yoon 1996) that the actual nominalising affix is phonologically empty, with /ɪŋ/ being a non-finiteness affix that attaches to V° — an assumption supported by the fact that the suffix also occurs in the progressive.

(21)



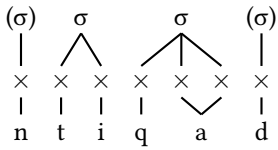
Arabic is a head-first language, which means that if all morphemes in (21) were linearised on the same tier, the order would be D-NOML-Subj-VIII-√-Obj, which obviously violates the adjacency requirement, because Φ(NOML) needs to combine with Φ(VIII) and Φ(√), although it is separated from them by the subject.[24] In Ackema and Neeleman's analysis, this would lead to a violation of Input Correspondence.

However, the various morphemes that make up the masdar are not realised on the same autosegmental tier. Assuming an autosegmental vocalic tier, we can say that since Φ(NOML) consists of the vowels /a.i:/, it is realised on this vocalic tier, which means that it is not linearised with respect to the subject and can appear on either side of it. This does not mean that its placement is arbitrary, of course. The different autosegmental tiers in a phonological representation are connected: each element on an autosegmental tier is associated with an element on the segmental tier or on another autosegmental tier. More importantly, this association is not arbitrary. Elements on an autosegmental tier can only be associated with specific elements on other tiers. Vowels, for example, can generally only associated with nucleus positions in syllables; tones can only be associated with vowels, etc. In the same way, we can think of Input Correspondence as condition on the possible associations of an element subject to it. Since NOML selects a projection of *v*, Input Correspondence restricts the association possibilities of the autosegment Φ(NOML) to Φ(*v*). Obviously,

The figure in (23) visualises the phonological structure of the masdar form *intiqād*:

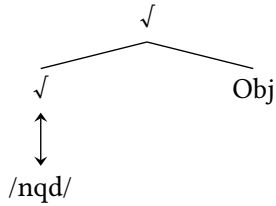
[24] In fact, Φ(D) also needs to combine with the other elements. Furthermore, the masdar form needs to be linearised before the subject. See section 3.2.5 for some details.

(23)



It may be helpful to consider in detail how this structure is composed. The first step in the construction of a masdar is the merger of the object and the root, yielding the following syntactic structure:[25]

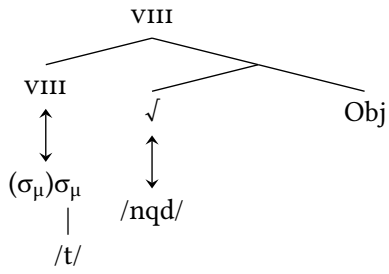
(24)



The object does not factor in the construction of the masdar form, so the first relevant element is the root /nqd/, which is not a phonologically complete unit: consisting of three consonants only, it does not constitute a licit phonological form. The root of course selects its complement, but there is no way in which the root can be phonologically integrated with it: the object does not provide any empty slots in which the consonants could be placed.[26]

The next step is the merger of the class VIII marker. This yields the following syntactic structure:

(25)

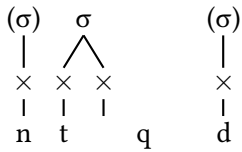


[25]Note that the stepwise presentation given here is not meant to imply a derivational process. It is merely intended to illustrate how the various pieces fit together.

[26]Note that there seems to be a general requirement that predicates and arguments be realised on the segmental tier. In sign languages, for example, certain adverbials can be realised as autosegments, such as the adverbial *with effort, difficultly* in (5), but the corresponding adjective *difficult* is always realised manually.

The class VIII marker is also a phonologically incomplete form. Because it selects the root in syntax, Input Correspondence requires that the two forms be combined. Doing this yields the following phonological structure, which, however, is still incomplete:

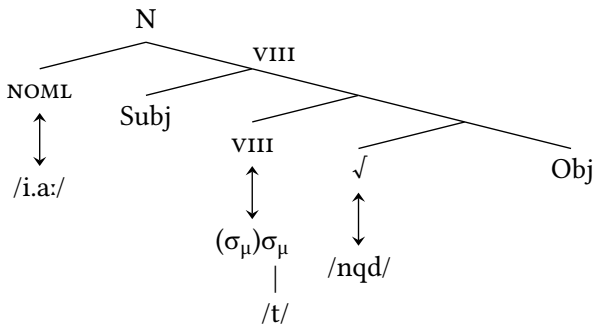
(26)



In this structure, the nucleus of the first full syllable is not filled and the root consonant /q/ is not associated with a slot position. Note that /q/ cannot be associated with the nucleus position, because Arabic phonology does not allow consonants in syllable nuclei. Furthermore, a stem-final extrametrical syllable has been added. This is the result of a general phonotactic rule in Arabic, which states that every stem ends in an extrametrical syllable — essentially a single onset consonant.

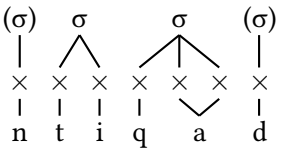
After the class VIII marker, the subject is merged. Phonologically, the subject is treated the same way as the object: phonologically it is complete and therefore it cannot be used to provide a host for the class VIII marker and the root. The next step is of course the merger of the nominaliser. This yields the following syntactic structure:

(27)



In the phonology, the vowels /i.a/ and the long syllable are added to the structure. The phonology can now fill the empty nucleus position of the first full syllable and it can link /q/ to a slot position. The resulting structure is (16):

(28)



At this point, the structure is complete, in the sense that there are no unsatisfied phonological requirements left. Interestingly, no head movement was needed in order to combine the three morphemes into a single word form. Input Correspondence specifies which heads need to combine and phonology has all the information it needs in order to combine them and form the structure in (28).

2.7.2 Mapping and linearisation

The model sketched here differs in certain ways from more standard views on the relation between syntactic and phonological structure. Traditionally, at least in G&B and minimalist theories, the syntax/phonology relation is seen as in some sense sequential: a (partial) syntactic structure is construed, which is then linearised and subsequently interpreted by the phonological component in order to construct a phonological representation. Depending on the framework, linearisation is trivial (e.g., Kayne 1994) or it is a distinct process involving linearisation parameters, as in HPSG but also in some theories within G&B and minimalism, (e.g., Abels and Neeleman 2009; Kremers 2009a; Neeleman and Weerman 1999). Representational formalisms such as HPSG differ from G&B and minimalism in that they do not incorporate actual sequentiality into the model, but it is still the case that the syntactic structure is primary and that linearisation is essentially a (post-) syntactic process, in the sense that it orders *syntactic* constituents.

If the model discussed in this section is on the right track, however, it has implications for the way we think about linearisation in particular. If a morphosyntactic head can be mapped onto an autosegmental tier, and is thus not subject to Linear Correspondence (at least with respect to the segmental tier), then linearisation cannot be a process that is part of syntax or part of the transfer to the interface. The reason for this is simple: in order for a head to be linearised, the linearisation process must have access to its phonological form. It is the phonological form associated with a head that is linearised, not the morphosyntactic head itself, and it is the nature of this phonological form —segmental or (partially) autosegmental— that determines what linearisation options it has.

This does not mean that syntax has no influence on linearisation, but its effects are indirect. The syntactic module does not specify that one element must be linearised before or after another, it only specifies hierarchical relations. The hierarchical structure obviously does influence the phonological structure, but

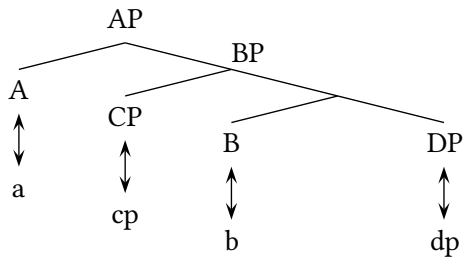
only indirectly. It is not the case that one can read the linear (phonological) form off the syntactic tree. The syntactic tree is mapped onto a complex, multi-tier phonological structure and phonology plays a role in establishing its final linear order.

Below, I discuss the mapping of syntax to phonology in more detail, focusing on the two principles of Linear and Input Correspondence and how the two interact. This section is then followed by a discussion of the way linearisation can be implemented in the current system.

Linear vs. Input Correspondence

It should be clear from the Arabic example in the previous section that Linear and Input Correspondence can place contradictory requirements on a syntactic tree and its mapping to phonology. Linear Correspondence requires two nodes that are sisters to be adjacent in the linear string. Input Correspondence can conflict with this requirement, however. Consider the tree in (29):

(29)



Assume that A in (29) selects category B. Input Correspondence then requires that «a» attaches to «b». Suppose also that BP is linearised in the order «cp b dp». Then Linear Correspondence allows two different orderings for Φ(AP), neither of which is compatible with Input Correspondence:

(30) a. «a cp b dp»

b. «cp b dp a»

These orderings are not compatible with Input Correspondence because «a» and «b» should be combined, which, at least for Ackema and Neeleman (2004), requires adjacency.

As discussed in the previous section, the Arabic masdar provides an example of this problem. The head NOML is not adjacent to the head VIII with which it must combine, at least on the (plausible) assumption that spec-head-comp is the basic order of Arabic. On this assumption, the tree in (46) is linearised as in (31):

(31) NOML subj VIII √ obj

The solution that I proposed is that NOML is exempt from Linear Correspondence since it is an autosegmental morpheme and as such is not realised on the segmental tier. This works for the case at hand, but the matter deserves a more principled discussion. What I will argue is that Linear Correspondence is actually the primary principle of the two, but it may be violated under certain conditions. It is exactly in such cases that Input Correspondence plays a role: an element that violates Linear Correspondence must obey Input Correspondence.

The underlying idea behind the two principles is actually the same. Informally put, both principles act to keep together what belongs together, clearly a necessary precondition for reconstructing a hierarchical structure on the basis of a linear string of segments and autosegments. There are, however, different concepts underlying the notion "belonging together" in the two principles. For Linear Correspondence, "belonging together" translates to sisterhood, while for Input Correspondence it translates to selection.

As an aside, note that for Ackema and Neeleman, Input and Linear Correspondence are *morphological* principles: they govern the mapping of what they call "word-syntactic" structures onto "word-phonological" structures. I explicitly assume, however, that Input and Linear Correspondence also govern the mapping from *syntax* to phonology.[27] There is no real reason to restrict these principles to morphology: like morphology, syntax generates tree structures that must be mapped onto linear phonological structures. Besides, as explained above, Linear Correspondence is closely related to the Nontangling Condition, which is obviously a condition on syntactic structures.

Ackema and Neeleman's solution to the problem of (29) / (30) is inspired by morphology: if $\Phi(A) = \emptyset$, i.e., if «a» is the empty string, then Linear Correspondence does not apply (or applies vacuously), and the configuration is grammatical. In other words, either order in (30) is grammatical if «a» is a null affix. Obviously, however, we cannot resort to this solution for the nominaliser in the Arabic masdar structure, since it is clearly not a null affix.[28]

What this means is that Linear Correspondence is a violable principle. The syntax-phonology mapping adheres to it only to the extent that the phonology

[27] In actual fact, I assume there is no strict distinction between syntax and morphology, contrary to what Ackema and Neeleman (2004) assume. Rather, there is a continuum with clearly morphological and clearly syntactic structures at opposite ends, and with any number of "mixed" structures in between, some more syntactic, others more morphological. The fact that there are principles that appear to apply just to morphology or just to syntax does not contradict this. It may simply be the case that the conditions for a particular "morphological" principle are simply never met at the other end of the continuum. See chapter 4 for discussion.

[28] Besides, as explained in chapter 1, I try to refrain from positing phonologically null elements as much as possible.

allows it. If a phonological form would arise that violates some basic phonological well-formedness principle, Linear Correspondence gives way in order to allow the phonology to construct a licit form. However, when this happens, the mapping adheres to the principle of Input Correspondence. We may see this as simply employing a different method of keeping together what belongs together. If sisterhood does not work, the system resorts to using selection as a criterion for determining what must be kept together.

On this assumption, Input and Linear Correspondence do not have the same status as mapping principles. Linear Correspondence is the primary principle, while Input Correspondence is secondary, only applying when the phonological form of the element(s) involved cannot be mapped onto phonology for some reason. We can in fact make a very specific proposal as to what exactly this reason is: we observed in the case of the Arabic masdar that the solution to the problem caused by configurations such as those in (29) lies in the autosegmental nature of the elements involved. What I propose, therefore, is that elements whose phonological form has some autosegmental requirement are allowed to violate Linear Correspondence.

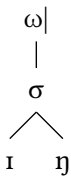
Strictly speaking, however, this violation of Linear Correspondence is only apparent, since the element in question is not realised on the segmental tier and therefore *cannot* be linearised with respect to its sister node if the latter is realised segmentally. Only elements that that are realised on the same tier in phonology can be linearised; elements realised on different tiers must be associated with elements on the segmental tier in order to be integrated into the phonological structure. Input Correspondence is one of the factors that plays a role in this process of association.

It should be noted that what kinds of phonological elements are autosegmental depends on the language in question. There is no universal set of elements or properties that are always autosegmental. That is to say, typical non-segmental elements, such as tone and syllable structure, are (presumably) universally autosegmental, but languages can also use segmental features as autosegments. In languages with vowel harmony, for example, the harmonising feature is often best analysed as an autosegment, and in Arabic, with its consonantal templates and vocalic "melodies", both types of segments are often treated as autosegmental.

Given that for Ackema and Neeleman (2004), Input and Linear Correspondence are morphological principles, it stands to reason that especially the former is relevant for quite a number of cases involving affixes. As indicated in footnote 27, I assume that there is no strict dividing line between syntax and morphology, which is reflected in the masdar analysis presented above: it is clearly syntactic in nature, although the resulting structure, *intiqād* 'criticising', is morphological.

In my analysis, Input Correspondence is also responsible for attaching affixes to their stems, but in a slightly different manner than the way Ackema and Neeleman propose. I assume that affixes in general have a prosodic requirement of some sort. This is obvious for the morphemes that constitute the Arabic verbal noun, but the idea can be easily extended to other affixes if we interpret the fact that a suffix is a suffix follows from a prosodic alignment requirement on the suffix's phonological component. Specifically, I represent the phonological form of a suffix such as the English gerund suffix *-ing* as follows:

(32)



The representation in (32) indicates that Φ(*-ing*) is a syllable consisting of the segments /ɪ/ and /ŋ/. Furthermore, it is associated with a prosodic word boundary, indicated as ω|. The pipe bar symbol indicates the prosodic boundary, the omega indicates the relevant prosodic category (here the prosodic word). This idea is not entirely new. McCarthy and Prince (1993) argue that alignment requirements plays a role in at least some types of morphemes (specifically infixes) and Hayes and Lahiri (1991) treat certain boundary tones in Bengali as morphemic, i.e., as morphemes that are realised as tonal patterns aligned with specific prosodic boundaries. The idea that pre- and suffixes in general have a prosodic alignment requirement is essentially a generalisation of these proposals.

This analysis of pre- and suffixes has a number of consequences, not only for affixes but also for clitics, which are discussed in section 3.3. For the moment, it suffices to note that treating affixes in this way has some precedence in the literature and allows us to deal with interactions of Linear and Input Correspondence in a straightforward manner.[29]

Linearisation

It should be obvious that the implications for linearisation of the current approach are substantial. As mentioned at the beginning of this section, the tra-

[29] It should be pointed out that there are a few cases in which Linear and Input Correspondence seem to interact in a more complex way: although **pass by-er* is ungrammatical, *come-outer* is not. Input Correspondence predicts that the agentive suffix appears on the verb, but in some cases, a realisation on the particle is preferred. Sometimes the suffix *-er* even appears twice: *fixer-upper*. Such cases are rare, however, and do not appear to undermine the general argument. See also the discussion in Ackema and Neeleman (2004, pp. 160–161).

ditional view is that linearisation takes place before phonology proper and that it deals with morphosyntactic heads. However, if we take the Separation Hypothesis seriously, there is no real need to linearise *morphosyntactic* heads, and if the analysis presented here is on the right track, then it would even be pointless to do so, since linearisation is *not* total in the formal definition of Kayne (1994): some heads are not linearised with respect to the rest of the clause. Only when the phonological form itself is taken into account, can it be determined which heads need to be linearised and which do not. It follows that linearisation takes place in phonology, as part of a more general process that composes the entire phonological form, not just the linear order of the segmental elements.

One might be tempted to argue that syntax specifies a partial linear order. However, such a move would raise the question how syntax determines which heads to linearise and which not to. That would seem to require a feature of some sort, but such a feature would obviously be redundant: it specifies information that is implicit in the phonological representation. Furthermore, it does not suffice to assume a feature [±linearise], because two autosegmental elements that are realised on the same tier do need to be linearised with respect to each other. The feature would need to be a multi-valued feature specifying the exact phonological tier on which an element is realised. In other words, it would copy phonological information into the syntax, which makes it obvious that such a feature is redundant.

To avoid such redundancy, we need to assume that linearisation takes place in phonology. More precisely, linear order emerges when the phonology tries to combine various chunks of phonological structure into a larger structure. In a way, linear order is not even the goal that the phonological system is striving for. Rather, it is an emergent property, the inevitable outcome of a phonological system trying to combine the phonological parts of the elements in the structure (see also Kremers 2013 for some discussion).

In other words, linear order is not a syntactic requirement, but a requirement of the medium used to externalise syntactic structures, i.e., sound or sign. Linearisation is required because it is impossible to vocalise or visualise an entire syntactic tree in one go. To use Hauser, Chomsky, and Fitch's (2002) terminology, linearisation is an "external factor", something that should not be part of the language faculty in the narrow sense.

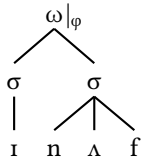
In this view, it makes sense that syntactic structure is actually underspecified for linear order, as is the case in the current approach. The order of the morphemes in the Arabic masdar, or the position of non-manual adverbs and negation in sign languages cannot be determined on the basis of the syntactic tree alone. In order to establish the linear structure, it is necessary to consider the phonological forms of the elements to be ordered.

This point is emphasised by the fact that although linearisation is in many

cases quite regular and predictable, there are cases of idiosyncratic linearisations, as the well-known case of the Degree element *enough* in English (and its equivalent in other Germanic languages) shows:

(33) a. *so/how/very/too* good
b. good *enough*

Adjectival degree heads in English such as *so/how/very/too* are normally linearised before the adjective, which is to be expected given that English is generally a head-complement language. However, in the case of *enough*, the only possible linearisation is Adj-Deg. What is most likely happening here is that there is an idiosyncratic linearisation instruction that places *enough* after its complement. In the current proposal, we can simply incorporate this information in the lexical entry for *enough*:

(34) $\begin{bmatrix} \text{Deg} \\ u\text{Adj} \end{bmatrix} \leftrightarrow$ 

Here, the addition of a φ-boundary ensures that /ɪˈnʌf/ is aligned with the left edge of a phonological phrase, resulting in the correct linearisation. Other Deg heads lack such a φ-boundary in their phonological representations, so that a default linearisation rule applies.

Default linearisation rules can be represented as mapping rules. As a first approximation (to be amended soon), one might envisage a rule such as (35) for head-complement structures:

(35) $[_{XP}\ X_a\ YP_b\ ] \longrightarrow \{\ \omega_a\ \varphi_b\ \}_\varphi$

This rule essentially states that a head X and its complement YP must be linearised in the order «Φ(X) Φ(YP)». However, it is not very sophisticated: it requires indices (the subscript *a* and *b*) to keep track of which syntactic element corresponds to which phonological element. A more sophisticated treatment is possible based on the observation that there is a correlation between stress placement and linearisation of head-complement structures:

(36) a. head-initial VP ↔ φ-final stress (English: *read **the book***)
b. head-final VP ↔ φ-initial stress (German: ***das Buch** lesen*)

The observation goes back to Truckenbrodt (although the examples he provides are not from English and German). Similarly, Nespor, Shukla, et al. (2008) also find that stress-initial phonological phrases are head-final, and vice versa.[30]

In order to use these facts for a theory of linearisation, it is necessary to adopt two fairly standard assumptions: **(a)** the assignment of stress within a phonological phrase depends on the hierarchical structure of the VP; and **(b)** in a structure $[_{VP}$ V Obj] phrasal stress is assigned to the object.[31] To implement this, we can start out with some standard mapping rules between syntax and phonology (cf. Nespor and Vogel 1986; Selkirk 1984; Truckenbrodt 2007). First, syntactic heads generally map onto a prosodic word. Truckenbrodt (1995) formulates it as in (37a), I adopt a slightly different formulation (37b):

(37) a. Wrap-X: A syntactic head is contained in a prosodic word.
b. X ↔ ω

Note that the syntactic head in these rules is a lexical head, i.e., a content word. Functional items generally do not constitute prosodic words (cf. Selkirk 1996).

Similarly, Truckenbrodt adopts a mapping rule for phrases:

(38) a. Wrap-XP: A syntactic phrase is contained in a phonological phrase.
b. XP ↔ φ

In a complement-head structure ZP consisting of the head Z and a complement YP, Z is subject to Wrap-X, while YP, which is a maximal projection, is subject to Wrap-XP. Note, however, that ZP itself is also subject to Wrap-XP. For Truckenbrodt, this is not problematic, because his constraint says that an XP must be *wrapped* in a p-phrase. That is, Wrap-XP does not say that each XP must correspond to a p-phrase, only that it must be contained in one.[32]

[30] The data Nespor, Shukla, et al. (2008) discuss even shows that the phonetic realisation of stress differs in both cases: φ-initial stress is mainly realised through pitch and intensity, while φ-final stress is realised differently. This, Nespor, Shukla, et al. claim, may allow infants to establish whether a language has head-complement or complement-head structures at a very early moment, even before they are able to parse words, let alone phrases.

[31] Stress assignment is of course a much more complicated topic, but the facts suggest that these two basic assumptions are essentially correct (cf., e.g., Truckenbrodt 2006b).

[32] This 'wrapping' aspect is not obvious in the representation of Wrap-XP in (38), which literally states that each XP must correspond to a p-phrase. There are technical solutions to this problem, e.g., by allowing the constraint to be violable, but I will not go into this issue here.

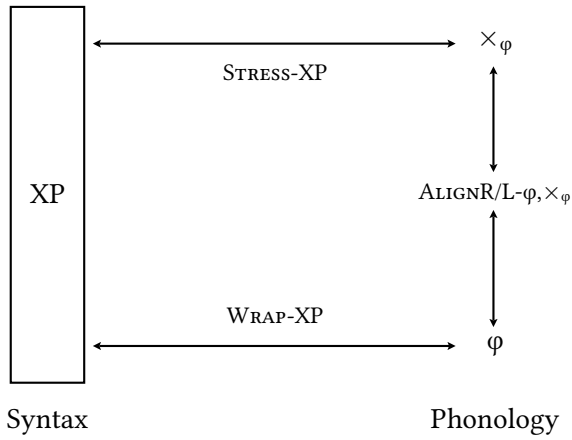


Figure 2.1: Stress-XP, Wrap-XP and stress alignment

Truckenbrodt assumes two further constraints: Stress-XP and AlignR/L-φ,×φ. Stress-XP states that every syntactic phrase has φ-stress. In the current framework, this can be captured with a mapping rule of the form in (39b):

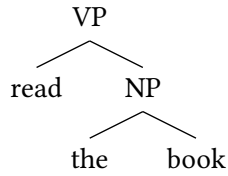
(39) a. Stress-XP: Every syntactic phrase has a φ-stress.

b. XP ↔ ×
|
φ

AlignR/L-φ,×φ is a purely phonological constraint in Truckenbrodt's model, and states that stress must be aligned to the left or right edge of a p-phrase, depending on the language in question. Schematically, this can be represented as in figure 2.1. What this schema represents is the manner in which a syntactic XP (to the left) is mapped onto a phonological structure. The constraint Wrap-XP ensures that the XP corresponds to a p-phrase, while Stress-XP ensures that it has phrasal stress. AlignR/L-φ,×φ then ensures that this stress is aligned to the left or right edge of the phrase, depending on the language.

For Truckenbrodt, linearisation is syntactic operation, which means that the mapping principles Wrap-XP and Stress-XP apply to linearised structures. Take, for example, the simple VP *read the book*, represented in (40):

(40)



The table in (41) shows how Wrap-XP and Stress-XP together make the right prediction with regard to stress placement. Since the NP is also subject to Stress-XP, the only grammatical structure is the first one, with phrasal stress on the object, as summarised in table 2.1.

	Wrap-XP		Stress-XP	
/riːd ðəˈbʊk/$_{\varphi}$	✓ VP	✓ NP	✓ VP	✓ NP
/ˈriːd ðəbʊk/$_{\varphi}$	✓ VP	✓ NP	✓ VP	*NP

Table 2.1: Wrap-XP and Stress-XP

The point is that it is not necessary to assume that the VP in (40) is linearised before phonology. If we include Align-R/L-φ,×$_{\varphi}$ in the overview, we see that the structures in which the object is linearised before the verb are ungrammatical as well, as shown in table 2.2.

	Wrap-XP		Stress-XP		AlignR-φ,×$_{\varphi}$
/riːd ðəˈbʊk/	✓ VP	✓ NP	✓ VP	✓ NP	✓
/ˈriːd ðəbʊk/	✓ VP	✓ NP	✓ VP	*NP	*
/ðəˈbʊk riːd/	✓ VP	✓ NP	✓ VP	✓ NP	*
/ðəbʊk ˈriːd/	✓ VP	✓ NP	✓ VP	*NP	✓

Table 2.2: Wrap-XP, Stress-XP and stress alignment

On the assumption that English has an alignment requirement for phrasal stress to be placed on the right, only the first of the four phonological structures in (40) is grammatical. The second and fourth structures, with phrasal stress on /riːd/, violate Stress-XP for the NP. Additionally, the second structure violates the alignment requirement. The third structure, with stress on the NP but

linearised in the order NP-V, also violates the alignment requirement. In other words, using the provided mapping principles, we can linearise a VP in phonology, without having to specify a linear order on (the basis of) the syntactic structure. Instead, we can assume two stress alignment rules:

(41) a. head-initial: XP $\leftrightarrow \times|_{\varphi}$
b. head-final: XP $\leftrightarrow {}_{\varphi}|\times$

In a language like German, which has different linearisation rules for different domains, it is necessary to associate the stress alignment rules to syntactic structure. For the VP, this can be done in the following way:

(42) VP $\leftrightarrow {}_{\varphi}|\times$

This rule states that a VP in syntax is associated with a phonological structure in which the stressed PWd is right-aligned in the phonological phrase. Essentially, (42) states that a verbal category is associated with a phonological phrase in which stress is left-aligned. It does not state that the V head is associated with the stress mentioned in the rule.

Obviously, the approach sketched here raises many questions. For example, it is not clear what happens when stress does fall on the verb, e.g., when the verb is (contrastively) focused. This would require being able to distinguish between nuclear stress and other kinds of stress.[33] However, in spite of these questions, the approach sketched here seems worth pursuing further, especially because it would be compatible with theories of prosodic bootstrapping (cf. Christophe et al. 2008), which argue that children pay attention to prosodic and stress cues very early in the process of language acquisition and are in that way able to establish the branching direction of the language in which they grow up.

It also remains to be seen whether all linearisation requirements can be reduced to stress alignment and at first sight, it seems unlikely that they can. This would mean that we would lose the in itself attractive notion that all linearisation statements are derived in the same way, but I do not see this as a problem. If linearisation is indeed a requirement of the modality (i.e., of the sensorimotor system), then any method for deriving linear order from a tree structure will do and different methods may be used in different contexts.[34]

[33] It may be possible to do this by adding an appropriate semantic component to the rule in (42).

[34] This is not really a new idea, by the way. For example, Abels and Neeleman (2009) argue that moved elements are typically linearised to the left. This is a linearisation principle that is not universally applicable for the simple reason that not every element is necessarily moved. See also Kremers (2009b), where I reach the same conclusion on different grounds.

2.8 Concluding remarks

It is a well-established fact about human language that we need both a linear and a hierarchical structure in order to describe utterances satisfyingly. Given this observation, it becomes necessary to specify the relation between the two. In this chapter, I have established the main assumptions that I make. I have also sketched the main ideas of the model that I propose. Essentially, the underlying idea of the current chapter is that this relation is essentially the relation between syntactic and phonological structure. That is, there is no syntactic or quasi-syntactic linear structure (e.g., PF or something similar) that defines a linear order of syntactic terminals. The main argument for this view is in order to determine the exact linear order, we need to know the phonological form with which a syntactic head corresponds. Since this information is by definition not part of the syntactic structure, it is not possible to define a linear order on the syntactic elements in the hierarchical structure.

3
Phonology-syntax interaction

The model sketched in the previous chapter allows us to deal with a number of phenomena in a new way. Instead of analysing them syntactically, we can treat them as the result of interaction between syntax and phonology. This is most clearly the case with simultaneity, briefly discussed in the previous section and the topic of the first section of this chapter. There are, however, other phenomena that I believe benefit from an analysis in terms of interaction between syntax and phonology. One of these is head movement, the topic of the second section of this chapter. Finally, I take a look at clitics, which are generally considered to be syntactic, although analysing them as such often raises problems.

3.1 Simultaneity

Simultaneity has already been discussed briefly in section 2.1. The term refers to phenomena in which two meaningful (syntactically represented) elements are realised simultaneously. The example in (5) in chapter 2 is repeated here as (1):

(1) **face**: *with effort* *DGS*
hands: STUDENT SIGN-LANGUAGE LEARN
'The student learns sign language with difficulty.'
(Leuninger 2005)

In this example, the adverbial *with effort* and the verb LEARN are realised simultaneously. The verb is a manual sign, the adverbial is realised through a facial expression and is commonly called *non-manual*. In this section, I discuss a number of such examples, from different syntactic categories, and show how they can be accounted for in the model described in the previous chapter.

3.1.1 Negation

Different sign languages realise negation in different ways, but most share the ability to realise negation through a simultaneous construction. Here, I will limit myself to negation in DGS.

As Pfau and Pfau and Quer discuss in several papers,[1] sentential negation in DGS is expressed through the combination of a non-manual and a manual marker, the manual marker being optional:

(2) **head**: *neg* *DGS*
hands: MOTHER FLOWER BUY (NOT)
'Mother does not buy a flower'
(Pfau and Quer 2002)

The non-manual marking, here indicated as *neg*, consists of a head shake and is realised together with the verb and the manual negation. If the manual negation is omitted, the head shake accompanies the verb alone. Note that the non-manual negation must always be supported by manual material, it cannot appear on its own.

Pfau and Quer (2002) analyse the DGS data as follows: DGS is a language with split negation, i.e., it has two negative elements that, when realised together, form a single negation, similar to French *ne...pas*. The manual negation NOT is one element, the head shake associated with the verb[2] is the other. NOT is in Spec,NegP and is lexically associated with a head shake. The head shake associated with the verb is analysed by Pfau and Quer as a [+neg] feature sitting in the head position Neg°. Because the [+neg] feature is affixal, the verb moves to Neg° in order to support the head shake phonologically. Note that DGS is an SOV language; Pfau and Quer assume that V and T (and Neg) are head-final, which means that this verb movement is essentially string-vacuous.

As Pfau (2002, 2008) shows, it is also possible for the head shake to spread over the entire VP, even though this is less common:

(3) **head**: *neg* *DGS*
hands: MOTHER FLOWER BUY
'Mother does not buy a flower'
(Pfau and Quer 2002)

[1] See Pfau (2001, 2002, 2008) and Pfau and Quer (2002, 2007) for details on negation in DGS, ASL and Catalan Sign Language (LSC); see Pfau (2008) for some discussion of negation in a number of other sign languages; see also Zeshan (2006) for a typological overview of negation in sign languages.

[2] More accurately, with the predicate, as the non-manual negation can also be used to negate copular sentences, which do not contain a verb.

Note that both forms of negation in DGS are sentential, there is no difference in meaning between (2) and (3). The crucial point for the current analysis is that the spreading of the head shake over VP is phonological in nature, as argued for by Pfau (2008). That is, the head shake spreads not over the VP but over the phonological phrase that corresponds to the VP. Evidence for this analysis comes from the observation that the head shake may spread as far as the subject if this subject is a pronoun and therefore part of the same phonological phrase as the VP:

(4) **head:** *neg* *DGS*
hands: IX:3 FLOWER BUY
'She does not buy a flower'
(Pfau 2008)

In this example, the subject of the clause is an indexical, i.e., a pronominal element. Unstressed pronouns are cross-linguistically phonologically weak: they usually do not constitute a prosodic word, instead cliticising onto an adjacent prosodic word. Assuming that this is true of indexicals in sign languages, we predict that spreading of the head shake is possible with indexical subjects. Pfau (2008, p. 26) argues that this is indeed common and provides the example in (4) as illustration.

For Pfau (2008), the manual negation is "syntactic", while the head shake is "prosodic", which for him entails that it is not syntactic. It should be clear that I do not make the same distinction: the phonological form of an element may be subject to prosodic restrictions, but that does not entail that it is not syntactic. I propose the following analysis for the non-manual part of the negation:[3]

(5) **not** ↔ Neg° ↔ *hs*
|
φ

The semantic component in (5) is obviously oversimplified but is not at issue. The syntactic component is a straightforward syntactic head Neg°. The phonological component consists of an autosegment, which I represent here as *hs*.[4] This autosegment has an additional prosodic requirement, which states that it must

[3] I ignore the manual sign NOT, because it is not relevant to the present discussion.

[4] Note that the phonological representations of sign language data here and below are simplified: *hs* is just an abbreviation for *head shake*, not an accurate phonetic representation. Although technically not really correct, the simplified representations are easier to understand than a system such as HamNoSys (basically an IPA for sign languages, cf. Hanke 2004).

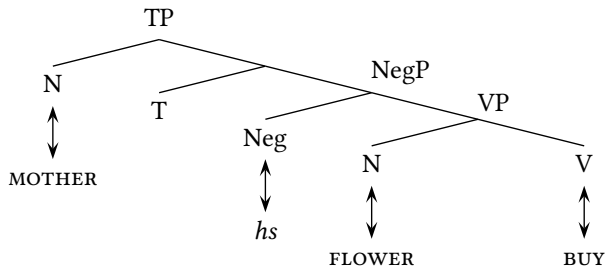
align with a phonological phrase (φ). It is this requirement that causes the head shake to spread over the phonological phrase corresponding to the VP in DGS.

As discussed above, the negative head shake in DGS can also be limited to just the verb, as in (2). The phonological component of the negation's lexical entry therefore must be assumed to be variable. Apart from the form in (5), there is a variant in which the head shake is associated with the prosodic word (ω):

(6) NOT ↔ Neg° ↔ *hs*
|
ω

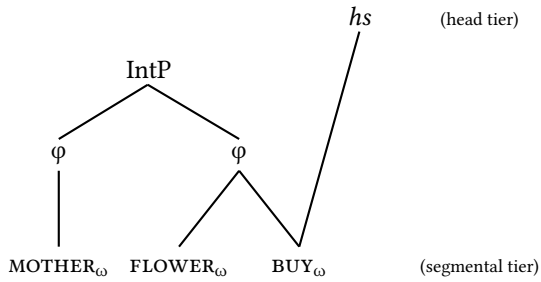
Let us see how this representation of the negation yields the correct structure. Since I assume that the difference between (2) and (3) lies in the phonological component of the negation, the syntactic structure for both clauses is the same:

(7)



This structure is mapped onto the following phonological representation:

(8)

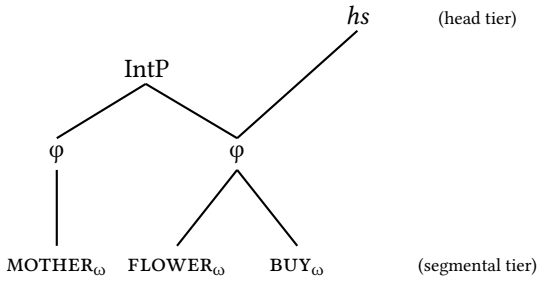


This representation contains the manual signs on the segmental tier and the prosodic structure in which they are contained. The head tier is indicated as well and contains the head shake. Assuming that *hs* is the variant in (6), it must be associated with a prosodic word. Since there are three prosodic words in the structure, we need to ensure that *hs* is associated with BUY. In fact, the principle

of Input Correspondence already ensures this: Neg° selects a verbal category, so that Input Correspondence requires that Φ(Neg) be associated with Φ(V), the head of the projection that Neg selects.

This analysis is not significantly different when instead of the phonological form in (6), the form in (5) is chosen. Input Correspondence still requires that the head shake be associated with the verb. The difference is that Φ(Neg) must be associated with a phonological phrase. Both requirements can be met by associating the head shake with the phonological phrase containing the verb:

(9)



The model provides a straightforward analysis of the simultaneous negation. Other simultaneity phenomena function essentially in the same way.

3.1.2 Adverbials

The example above in (1) shows a non-manual adverbial in DGS. It should be no surprise that non-manual adverbs also occur in other sign languages. The following examples are from ASL:

(10) a. **face**: /*mm*/ *ASL*
hands: HER HUSBAND COOK-[dur] DINNER
'Her husband has been cooking the dinner with pleasure'

b. **face**: /*th*/
hands: HER HUSBAND COOK-[dur] DINNER
'Her husband has been cooking the dinner inattentively'
(Corina, Bellugi, and Reilly 1999, p. 310)

/mm/ and */th/* represent different facial expressions, associated with the adverbial meanings *with pleasure* and *inattentively*, respectively, as indicated in the translations. Anderson and Reilly (1998) identify eleven different facial adverbials for ASL. Similar examples can be found in other sign languages: Kyle and Woll

(1985, p. 86) discuss a number of facial expressions that contribute adverbial (or in some cases adjectival) meaning in British Sign Language (BSL), Meir and Sandler (2008, pp. 173–176) report non-manual adjectives and adverbs for Israeli Sign Language (ISL) and Johnston and Schembri (2007, p. 150) discuss facial adverbials in Australian Sign Language (Auslan).

The phenomenon is therefore widespread in sign languages. Importantly, as most authors note, the forms and meanings of these facial expressions are conventionalised (i.e., not iconic), they differ from one sign language to the next, and they contribute to the meaning of the utterance: leaving them out changes the meaning of the phrase. That is, as already argued for by Liddell (1980), these facial expressions are not mere gestures, they are linguistic. We therefore expect them to be represented in the syntactic structure underlying the utterance.[5]

One might perhaps argue that these adverbials are morphological, not syntactic (in fact, Sandler and Lillo-Martin 2006, p. 61–63 suggest as much), especially considering the fact that they are realised simultaneously with the verb, not the entire VP. However, although this may be a possible analysis for the non-manual adverbs demonstrated so far, there are non-manual adverbials that cannot be analysed morphologically. Happ and Vorköper (2006, p. 363) give the following examples from DGS:

(11) a. **f/b:** *presumably*
h: (POSSIBLE) SVEN WORK:3 GO:Perf:3
'Presumably, Sven has gone to work already.'

b. **f/b:** *possible*
h: (POSSIBLE) SVEN WORK:3 GO:Perf:3
'Sven has possibly gone to work already.'

The examples in (11) contain sentential adverbs. Specifically, there are two components to the adverbs: a manual one, POSSIBLE, which is optional, and a non-manual one, a combined facial expression and body position.[6] The two components are independent from each other: the manual component is optional and

[5] Another interesting cross-linguistic characteristic of these facial adverbials is that they all involve the lower part of the face, in particular the mouth and cheeks. The upper parts of the face tend to mark sentence types and properties such as topic and focus (cf. Meir and Sandler 2008; Wilbur 2000, p. 175).

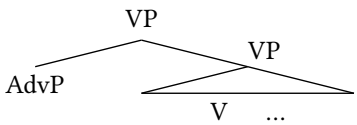
[6] In fact, it is possible that the facial expression and body position contribute independently to the meaning. Such combinatoriality is often assumed for different facial and eye brow configurations (e.g., Dachkovsky and Sandler 2009; Wilbur 2000). Note, however, that even if there is a certain combinatoriality in non-manual components, this system lacks the so-called *duality of patterning* of the segmental system, as the independent components all have a specific meaning.

the two do not need to share the same meaning.[7]

The relevant aspect of these examples is the fact that the non-manual component spreads over the entire clause. As with the VP adverbials above, these non-manual sentential adverbials contribute crucial information to the utterance and their forms are conventionalised. We therefore expect them to be represented syntactically. Since morphological analyses are by definition limited to the word/sign, these non-manuals cannot be morphological, as they involve more than one manual sign.[8]

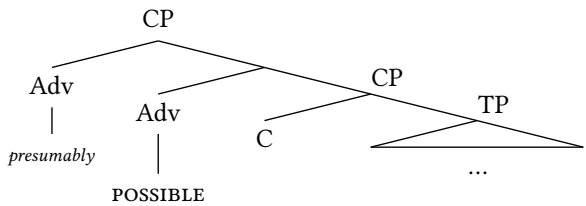
Let us see how adverbials can be handled in the system proposed here. For the sake of simplicity, I assume that adverbials are adjuncts. The VP adverbs in examples such as (1) and (10) can then be represented as adjuncts to VP:

(12)



Similarly, sentential adverbs such as those in (11) are adjoined to the CP. If we separate the manual and non-manual components, we may be dealing with two adjuncts:

(13)



What is relevant here is the fact that the adverbials are part of the syntactic structure in a position c-commanding the element(s) with which they are simultaneously realised. The LCA would therefore predict that they are realised sentence-initially (in the case of sentential adverbs) and preverbally (in the case of VP adverbs) but obviously they are not.

It should be clear at this point how these adverbials can be analysed: the non-manual (components of) the adverbials are autosegments that need to be associated with segmental material. This association takes place through prosodic

[7]Possibly, the manual component expresses a general modality, which is then specified further by the non-manual component.

[8]In a way, this argument sounds almost too simplistic. The argument itself is correct, however, and in my opinion shows the arbitrariness of the syntax/morphology distinction. See also chapter 4.

structure: the VP adverbs, like the negation, require association with prosodic words, the sentential adverbs in (11) with (presumably) intonational phrases.

One element is missing here, however. In most minimalist analyses, adverbials are optional elements: there is no c-selectional relation between the adverbial and the element it modifies. Therefore, Input Correspondence does not apply and the association between the non-manual adverbial and the segmental string cannot be established. There are in principle two ways to deal with this: assume that there is some sort of (c-)selectional relation between the adverbial and the modified element, or extend Input Correspondence in some way.

The latter option would be fairly *ad hoc*, since there is no reason to do so other than the problem under discussion. There is precedent for the first idea, however: in frameworks such as HPSG, it is a common assumption that there is a certain selectional relation between adverbials and the elements they modify. In such analyses, it is the adverbial that selects the modified element, however, not the other way around. The reason for this is that adjuncts can generally occur only with particular kinds of elements, while the types of adjuncts that can occur with a particular kind of element (e.g., noun phrase, verb phrase) is usually not restricted: nouns can be combined with adjectives, prepositional phrases or relative clauses, while an adjective must be combined with a noun, it cannot be combined with a verb or a preposition.

Interestingly, this kind of selection is exactly what is needed in the current analysis: Input Correspondence says that if X selects Y in syntax, Φ(X) associates with Φ(Y) in phonology. Since the adverbial is the autosegment, it is the element that requires anchoring, i.e., association with some segmental element. The verb, on the other hand, is segmental and therefore lacks a prosodic requirement, which means that Input Correspondence does not even apply. In order for the analysis to work, it is therefore the adverbial that must have the selectional requirement. In the HPSG analysis of adverbials, this is the case.

We can conclude, then, that a VP adverbial such as *with effort* in (1) has a lexical entry along the following lines:

(14) WITH EFFORT $\leftrightarrow \begin{bmatrix} \text{Adv} \\ u\text{V} \end{bmatrix} \leftrightarrow \begin{matrix} w.e. \\ | \\ \omega \end{matrix}$

Obviously, the *w.e.* in the phonological representation is a simplification. What is important here is that the syntactic component of this lexical entry has a selectional restriction requiring it to combine with an element of category V, and the phonological component has a prosodic restriction requiring it to associate

with a prosodic word. Input Correspondence then assures that the phonological component is associated with the verb.

The adverbials discussed so far have been facial adverbs: they are realised on (the lower part of) the face. There are other kinds of simultaneous adverbials, however, and furthermore, multiple non-manual adverbials may be combined. Consider the following DGS example:

(15) **body**: *exclusive*
face: *reluctant*
hands: APPROACH-EACH-OTHER.$\text{Cl}_{\text{uprightG}}$.slow
'two people approach each other slowly, reluctantly, and with hostility
(Leuninger, Hohenberger, and Menges 2005, p. 331)

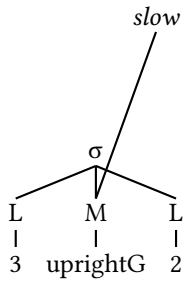
This example contains two or possibly three simultaneous adverbs. The manual sign is a two-handed sign, with both hands consisting of an upright G-handshape (stretched index finger). This handshape functions as a classifier for human beings. The two hands move toward each other, expressing the meaning *approach each other*. The first non-manual adverbial is expressed through body posture: the upper body of the signer is leaned backwards, expressing exclusivity, which is understood as hostility in the current case (cf. Wilbur and Patschke 1998). The second non-manual adverbial is the signer's facial expression, conveying the meaning *reluctant*.

Additionally, the movement of the hands is realised more slowly than usual, which expresses another adverbial, which can be translated as *slowly*. Because of the obvious iconic nature of this part of the utterance, it is difficult to tell whether it is merely gestural or whether it is truly morphemic. On the assumption that *slowly* is morphological, however, it can easily be analysed in the same manner as facial adverbs. In sign language phonology, several kinds of *movement* are recognised. One of these is so-called *path movement*, where the signing hand changes its location.[9] In the sign in (15), the two hands undergo a path movement: they move from outward right and left positions toward each other. The adverb *slowly* modifies this path movement: it reduces the speed at which it normally occurs.

We can therefore treat the adverb *slowly* as an autosegment realised on a tier for path movement. A partial phonological structure of (15) is then (16):

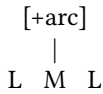
[9]Other types are sometimes collectively called *internal movement*, a term that describes movement in which the hand does not change location, but its orientation or configuration changes. See Sandler and Lillo-Martin (2006) for details.

(16)



Here, I adopt a representation of the structure of the syllable in sign languages proposed by Sandler (1986, 1989): a sign consists of a path movement (M) and two locations (L), the starting and end points of the path movement.[10] In the sign in (16), the component *slow* is associated with the movement (M) part of the syllable.

The analysis is in fact similar to the rule of *arc linking* proposed by Sandler (1990, p. 25). [+arc] is an ASL morpheme expressing duration. It is realised as a change in the path movement of a sign, which is changed from a straight movement to an arc-shaped movement. Sandler represents it as follows:

(17)


The representation in (17) associates a feature [+arc] to the movement (M) part of the sign syllable, indicating that the sign is made with an arc movement. [+arc] is only associated with the movement, not with the locations, because the feature is not compatible with them.

We may similarly assume that the adverb *slowly* is realised by something like a duration feature [+dur] that is associated with the M in (16). Phonetically, these features are comparable to distinctive features in spoken languages, in that they determine the phonetic shape of the segment (sign). This contrasts with the non-manual adverbials in earlier examples, such as *anxious* in (`[ex:04-27]`), which are not distinctive features. Rather, they bear strong similarities to tone in spoken languages.

In the current model, however, the phonetic difference between the two kinds of elements are not relevant. Both behave as autosegments and are associated with segmental material through standard phonological principles (e.g., Left-to-Right Association).

[10] An earlier model of sign language syllables posits holds (H) instead of locations, cf. Liddell and Johnson (1989). See Sandler and Lillo-Martin (2006, pp. 128–138) for a discussion of both models.

3.1.3 Interrogative marking

Another typical phenomenon that is usually expressed non-manually (and hence simultaneously) in sign language is question marking. Interrogatives are normally marked by certain non-manual markers (which can take various forms, depending on the type of interrogative and other factors; see Sandler and Lillo-Martin 2006, 459ff for details and further references). In the case of polar questions, this is usually the only way the utterance is marked as interrogative:

(18) **face:** *Q*
hands: HE COME
'Is he coming?'

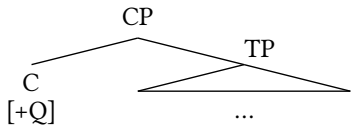
The interrogative marking extends over the scope of the interrogative but excludes elements that are not in its scope:

(19) **eyebrows:** *Q* *BSL*
hands: K.I.L.B.Y BEFORE GOOD NOW GOOD
'Kilby, who was good before, was he good now?'
(Kyle and Woll 1985, p. 156)

As has often been noted, such facial interrogative markers function in much the same way as interrogative intonation in spoken languages. Such intonational patterns are normally considered non-syntactic, so that by analogy, a non-syntactic analysis suggests itself for interrogative marking in sign languages as well.

However, interrogative marking in sign language is amenable to the same analysis that we applied to negation and non-manual adverbials. For example, assuming that interrogatives are headed by a [+Q]-marked C head, the basic structure of clauses such as (18) is (20):

(20)



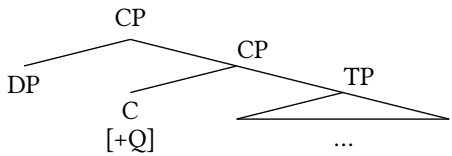
Now we simply need to assume that this interrogative C head can be represented with a lexical entry such as in (21):

(21) INT ↔ [C,+Q] ↔ *r.b.*
|
IntP

The *r.b.* in the phonological component stands for *raised brows.*[11] Prosodically, it must be associated with an intonational phrase. Input Correspondence requires that *r.b.* be associated with the head of the category that C selects, which is T, or more precisely, Φ(T). The prosodic requirement ensures that *r.b.* spreads over the entire phrase.

In (19), the interrogative marking does not extend over the first part of the utterance, K.I.L.B.Y BEFORE GOOD 'Kilby, who was good before'. This is expected, however, given the fact that this part of the utterance is a constituent that functions as a topic, and that topics tend to form independent intonational phrases (cf. Nespor, Shukla, et al. 2008; Selkirk 2005). Assuming that the left-dislocated phrase is adjoined to CP,[12] the syntactic structure for (19) and (27b) is the following:

(22)



This structure is of course identical to the one in (20), with the addition of the topic in Spec,CP. Φ(C[+Q]) spreads over the prosodic domain associated with the CP, i.e., the IntP.

3.1.4 Topics

Like negation and interrogatives, topics are often marked non-manually in sign language (cf. Sandler and Lillo-Martin 2006, pp. 406–413 and references cited there). The phonological form of topic marking varies depending on the type of topic (see Aarons 1994, p. 156 and Wilbur 1994 for ASL) and (presumably) also varies form one sign language to the next. A typical example from Danish Sign Language is (23):

[11] A brow raise is an essential part of non-manual marking for yes/no questions (constituent questions are usually marked with furrowed brows), but the representation in (21) is not intended to be exhaustive.

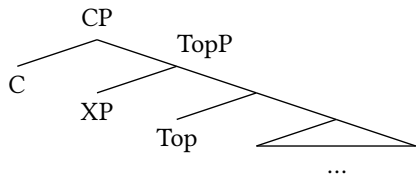
[12] Nothing really hinges on this assumption, one may also argue that the left-dislocated phrase is in the specifier of some high functional head in the C-domain.

(23) **face:** *topic* *DSL*
hands: NURSERY *index:3* 1:SEND:3
'I took (her) to the nursery'
(Engberg-Pedersen 1994, p. 75)

In (23), the signer first introduces the goal argument of the clause and uses an indexical to locate it in signing space. Simultaneously, NURSERY and the indexical are marked non-manually as topic. Next, the signer signs the verb SEND, which is a so-called *agreeing* verb[13] that agrees with the subject (indicated with '1' for first person) and the goal of the action, which is the nursery. This locative agreement is indicated with the index '3', which refers to the location in signing space that is associated with the nursery.

Views on the syntactic position of topic markers differ. In syntactic approaches that are based on work by Rizzi (1997) and Cinque (1999), topics appear in the specifier position of a dedicated Top° head:

(24)



In the structure in (24), the XP would be topic. In languages that have an overt topic marker, the argument goes, this marker is an overt realisation of the Top° head. However, if this is indeed the correct structure, the phonological marking of the topic in sign languages (and in spoken languages that use intonation for this purpose) functions in a way that is very different from interrogative marking and negation. Both interrogative marking and negation extend over their complements (or, more precisely, over the phonological domain the complement is embedded in). The same is essentially true for non-manual adverbs: VP adverbs are realised simultaneously with the verb, which is in the c-command domain of the adverb. In the case of topics, however, the phonological marking associated with the Top° head would extend over its specifier.

However, there is data to suggest that this cannot be the only source of topic markers. If the topic marker is a Top-head in the verbal projection line, we would expect that a topic marker can co-occur with a case marker on the DP in Spec,TopP. However, this is not always the case. In Japanese, for example,

[13]Not all verbs in sign languages agree. Those that do can be divided into several different categories. See e.g. Sandler and Lillo-Martin (2006) and references cited there for details.

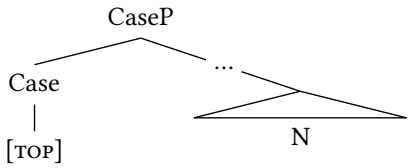
the topic marker *wa* is in complementary distribution with the case markers *ga* (nominative) and *o* (accusative):

(25) a. Taiyō ga noboru. *Japanese*
sun NOM rise
'The sun rises.'

b. Taiyō (*ga) wa noboru.
sun NOM TOP rise
'The sun rises.'

The most straightforward analysis for these data is one in which the topic marker and the case marker (and presumably also adpositions) are reflexes of the same functional head in the nominal projection line, as schematised in (26):

(26)



I will assume that the sign language examples under discussion have essentially this structure.[14] On this analysis, the non-manual topic marking would fall in line with the interrogative and negation markers: the topic marker is phonologically realised simultaneously with the element in its scope, i.e., with its c-command domain. As above, I assume that the actual domain is defined phonologically: the topic marking spreads over the phonological domain (presumably an intonational phrase) that contains the phonological component of the N° head.[15]

3.1.5 Simultaneity in spoken languages

One possibly interesting consequence of the sign-language data and the analysis presented here is that it can perhaps be extended to deal with intonational phenomena in spoken language. For example, if we look at the English translations of (18) and (19), we see that they are identical in relevant respects:

[14]The actual category of the head carrying the [TOP]-feature may not be Case°. The label is inspired by the Japanese example, but since sign languages generally do not have case marking, it is probably not the most appropriate label here.

[15]he analysis sketched here obviously raises the question whether there are indeed two kinds of topic markers and if so, whether they co-occur or not. I leave this issue open for future research, however.

(27) a. **int:** *Q* *English*
segm: Is he coming?

b. **int:** *Q*
segm: Kilby, who was so good before, was he good now?

Just as in the BSL examples above, question intonation in English extends over the entire clause in (27a) but only over the second part in (27b). We can analyse this in essentially the same way:

(28) INT $\leftrightarrow$ $\left[\text{C,+Q}\right] \leftrightarrow \text{H}|_{\text{IntP}}$

In (28), the semantic component is again highly simplified. The phonological component consists of a high tone that is right-aligned with an intonational phrase. This is just an approximation, but suffices for the point at hand. What (28) then says is that the interrogative intonation in a yes/no question is essentially the phonological component of an interrogative C head. In some languages, this phonological component is segmental, i.e., Russian *li*, Standard Arabic *hal*, Japanese *ka*, etc., but the current proposal allows us to say that languages such as English are not much different. The only difference is the prosodic autosegmental tier onto which the phonological component of the interrogative C head is mapped.

The idea of treating intonational patterns as reflexes of syntactic heads may be counter-intuitive at first. However, I believe this impression stems from our intuition that syntax deals with the arrangement of words and that words consist of segments, an intuition that is formalised as the Exclusivity Condition. For example, given a syntactic structure such as in (20), the Exclusivity Condition (or Totality, cf. section 2.1 of chapter 2) leads us to expect C to be linearised before or after TP. Since intonational patterns are by definition not linearised as part of the segmental string, they are not commonly thought of as being 'syntactic'. However, the phrase "being syntactic" has a very particular meaning here. An element is syntactic if it is represented as a feature on a head in the syntactic structure. Therefore, saying that the interrogative intonational pattern is syntactic means that this pattern, which in itself is a phonological phenomenon, is linked to a feature on a syntactic head. Put this way, the idea of treating prosodic patterns as part of syntax should not appear so counter-intuitive.

3.1.6 Concluding remarks

The discussion shows that the simultaneity data are amenable to an analysis in terms of prosodic syntax. The primary advantage of the analysis is that it allows

simultaneity to be analysed in syntactic terms, something that is not immediately obvious. Furthermore, the analysis does not require new assumptions. We can take an existing model, that of prosodic morphology, and extend it in a natural way to syntactic structures.

One important aspect of this analysis is the assumption that adverbial facial expressions and negation are aligned with prosodic categories and not with (morpho)syntactic categories. It has been argued that non-manuals in sign language are aligned with syntactic categories (e.g., Neidle et al. 2000), but it is by now quite clear that certain types of non-manuals, especially interrogative and topic marking, are aligned with prosodic domains (see Sandler 2011, who summarises the relevant discussion).

For non-manual negation and adverbials, matters are not as clear. One datum suggesting that at least the non-manual negation in DGS aligns with prosodic categories and not with syntactic ones is the example in (4), where negation spreads over the subject pronoun. For VP-adverbs, however, the data are compatible with both types of analyses. They align with the verb; whether they align with the syntactic head or with the prosodic word is not so easy to determine.

One fact that tentatively supports the assumption made here is that the current analysis makes one particular prediction. Because phonological structure is organised along the prosodic hierarchy, and because the prosodic hierarchy obeys the Strict Layer Hypothesis (Nespor and Vogel 1986; Selkirk 1984), we expect that two autosegmental morphemes are either coextensive or that one completely contains the other. That is, we do not expect to find structures of the following type:

(29)

tier 1:		*x*			
tier 2:	*y*				
segm:	pword	pword	pword	pword	

In this hypothetical example, an autosegmental marker on tier 1 overlaps with a marker on tier 2, but only partially. Given the assumption made here that autosegmental markings coincide with prosodic constituents, the situation in (29) could only obtain if the relevant prosodic constituents violated the Strict Layer Hypothesis.

Presumably, the same prediction would hold if all autosegmental elements would align with syntactic categories. Crucially, the prediction would not hold if some autosegmental elements aligned with prosodic categories, while others aligned with syntactic categories. Because of the fact that syntactic and phonological constituents do not map onto each other one-to-one, cases such as in (29) might be possible. A survey of the relevant literature (cf. Kremers 2012b) sug-

gests that such cases indeed do not occur.[16] We may tentatively conclude from this that because at least some autosegmental elements align with prosodic categories, they all do so. More research is needed, however, to see if the prediction really holds.

Note, however, that under the Separation Hypothesis (cf. section 2.3), alignment with a syntactic category is strictly speaking not even an option. A syntactic head does not carry phonological information, it consists of morphosyntactic features only. Therefore, it has no temporal extent, no left and right edges with which an autosegmental phonological element could be aligned.

If we phrase the question in such a way that it is compatible with the current proposal, we should ask whether non-manuals align with a prosodic category that is lexically associated with a specific head or whether they can also spread over a prosodic domain that is not introduced by a lexical head. Although this question in itself is an interesting one, it is difficult to answer and, moreover, of little consequence for the proposal at hand unless it can be shown that a non-manual spreads over a domain that is clearly not (also) phonological.

Take, for example, the VP adverb in (30), repeated here:

(30) **face:** *with effort* *DGS*
hands: STUDENT SIGN-LANGUAGE LEARN
'The student learns sign language with difficulty.'

The lexical entry for the verb LEARN is (31):

(31) $\lambda e\, \lambda x\, \lambda y$ (**learn**(x,y,e)) $\leftrightarrow$ [V, *u*N] $\leftrightarrow$ LEARN

The phonological component, which I have indicated simply as LEARN, is a prosodic word. The non-manual adverb *with effort* is aligned with this PWd by virtue of its lexical specification. The relevant question is now what happens with the adverb when the verb's PWd is modified in some way. The problem is that there are no relevant modifications that one can think of that may provide a clue. In larger prosodic domains, e.g., IntPs or Utterances, it is not uncommon for additional boundaries to be introduced (e.g., when a parenthetical is added to the structure). At the level of the PWd, there is no equivalent process (at least, not that I am aware of) and therefore no way to determine what the non-manual does when the PWd is broken up.

At best, one may glean something from processes that extend the PWd, for example, by incorporating a clitic. But in such cases, it would first need to be established on independent grounds that verb + clitic really to form a (single)

[16] Although the adage "absence of evidence is not evidence of absence" of course applies here.

PWd and not a larger prosodic category such as a clitic group or a recursive PWd. Only when it can be established independently that the verb by itself does not constitute a PWd anymore after the addition of the clitic,[17] do we have reason to argue that at least some non-manuals align with syntactic categories. However, to the best of my knowledge, such evidence does not exist.

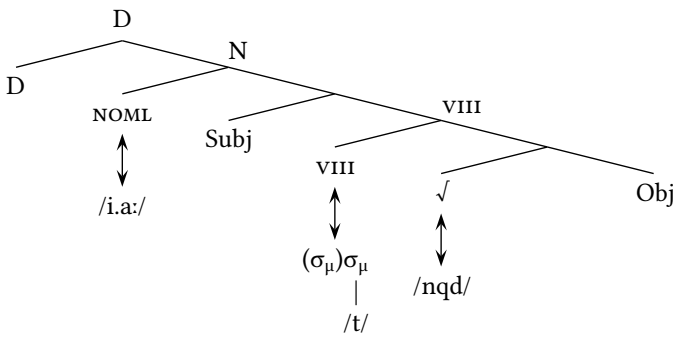
Summarising, the grammar architecture basically forces the conclusion upon us that non-manuals align with prosodic categories: the component of the non-manual that must be aligned is the phonological one, not the syntactic or semantic one, and the only type of element it can align with is also phonological. The question that needs to be answered in future research is whether non-manual adverbs align with the prosodic constituent that the verb introduces or whether they can also align with prosodic constituents that are not introduced lexically.

3.2 Head movement

3.2.1 Introduction

Some of the examples discussed so far, in particular the Arabic verbal nouns discussed in section 2.7.1, suggest an analysis in terms of head movement, at least, for those working in a minimalist framework. Consider again the tree in example (21) in chapter 2, repeated here as (32):

(32)



In chapter 2, I argued that the word form *intiqād* is created in phonology and that there is no need in the syntactic structure to combine the three heads √, VIII and NOML into a distinct subtree. On first sight, however, an analysis in terms of head movement (a syntactic operation) seems capable of accounting for the structure.

[17] Which, according to Scheer (2008), means that there must be some phonetically observable and phonologically relevant boundary between verb and clitic.

If we were to adopt such an analysis, however, we would still need an account of the way in which the three morphs are combined. Most head movement analyses assume that head movement results in head adjunction, with the moved head adjoining to the left of the target of movement. In this view, the target head appears as a suffix on the moved element. The Arabic verbal noun cannot be analysed in this way, however. Although there is a root, /nqd/ in the example above, the two elements that it combines with, Φ(VIII) and Φ(NOML), are not suffixes. Therefore, a phonological explanation is still required to account for the word form.[18]

Head movement in minimalism (and in G&B before it), in fact, is not a unitary phenomenon. There are at least three different types of motivations for head movement analyses, resulting in different concepts of head movement. Firstly, there are head movement analyses that are motivated syntactically, in the sense that an element which is taken to be a head appears in a position that is assumed not to be its base position. V-to-C movement in verb-second languages like Dutch and German or T-to-C movement in English questions are two typical examples. In both cases, the finite verb in the clause appears in a position that is thought to be too high to make the correct (lexical) semantic contribution and to govern its dependents (arguments). Furthermore, there are sentence structures in which the finite verb is not in the same high position: subclauses in Dutch and German, and declaratives in English.

A second kind of motivation for head movement is morphological: a word form consists of multiple morphemes which are assumed to have different structural positions. V-to-T movement is a typical case: the idea that the verb stem picks up its inflectional affix in T. In his original proposal for this type of head movement, Pollock (1989) argued that head movement is also visible syntactically, by showing that the verb appears in different positions in the clause (with respect to negation and adverbials) that correlate with the suffixes on the stem. This type of analysis has become so widely accepted that it is often considered unnecessary to provide actual syntactic evidence of head movement.

In fact, head movement has become such a widely-used analytic tool in minimalism that it is often invoked in cases where even morphological evidence is lacking. For example, Longobardi assumes that N-to-D movement takes place in English noun phrases even though this movement is not "visible", i.e., it is assumed to take place at LF. Other examples are analyses of clause structure that assume a split CP, with the (finite) verb moving to different positions depending on the semantics of the phrase (Rizzi 1997). This kind of head movement can be seen as theoretically motivated, since there is no direct evidence for its existence.

Syntactically motivated head movement differs in an important way from mor-

[18]Which is, of course, the whole reason why prosodic morphology was developed in the first place.

phologically motivated head movement.[19] The former constitutes *head substitution*, while the latter constitutes *head adjunction*. When V moves to C, for example, it is generally assumed that the C position is empty before movement, in the sense that there is no phonological material associated with it.[20] Morphologically motivated head movement combines two morphs, i.e., there is morphological material associated with the target position.

It has often been noted in the minimalist literature that head movement is problematic on theoretical grounds: it violates the Extension Condition, and the moved head no longer c-commands its trace, under the standard, simplest definition of c-command. Various proposals have been made to solve these issues: one proposal argues that head movement actually extends the tree (sometimes called *reprojection*), cf. Bury (2003), Georgi and Müller (2010), and Surányi (2005), possibly followed by morphological merger (Matushansky 2006). Alternatively, Nunes (2001, 2004) argues that head movement takes place via a "separate workspace", or head movement is reanalysed as remnant XP movement (Koopman and Szabolcsi 2000). There have also been suggestions that head movement takes place at PF (e.g., Boeckx and Stjepanović 2001; Chomsky 2001), although no worked-out proposal has been offered and the idea has been criticised by e.g., Zwart (2001) and Matushansky (2006).

I will not go into the merits and deficiencies of theories of head movement here. What I will do is argue that *morphologically* motivated head movement is more properly handled in phonology. The analysis differs qualitatively from the type of analysis suggested by Chomsky (2001) and Boeckx and Stjepanović (2001), and is therefore not subject to the same criticism. Note that I explicitly do not want to claim that *all* kinds of head movement should be handled in phonology. Especially syntactically motivated head movement cannot be handled in the manner described here.

3.2.2 Motivating movement

Before we turn to the question of head movement, let us take a step back and consider what sort of observations motivate the assumption of movement. Foremost is the observation that a phrase can be interpreted in two positions, while it appears in only one:

(33) a. **Diesen Film** wollte ich schon immer mal *t* sehen.
this.ACC film wanted I yet always once see
'I've always wanted to see this film.'

[19] The same is true of theoretically motivated head movement.

[20] It is usually assumed that there are features present at the target position, e.g., in order to trigger movement, but no phonological material.

b. **What** are you doing *t*?

In (33a), the noun phrase *diesen Film* 'this film' acts as topic: it occupies the first position in the clause and the clause makes a statement about it. Crucially, however, *diesen Film* is also the accusative object of the verb *sehen* 'to see', as demonstrated by the accusative case on the demonstrative *diesen*. Objects, however, normally appear lower in the structure, linearised (in German) to the (immediate) left of the main verb, the position indicated with *t* in (33a).[21] The clause in (33b) is very similar, the main difference being that the element *what* is not the topic of the clause but a *wh* word, which is fronted in English. At the same time, however, it is also the object of the verb *doing*, which would require placement after the verb, English being a VO language.

Crucial for the argument that the boldface elements in (33) are interpreted in two positions is the observation that objects generally occur in a different position from the one the boldface elements are in. The boldface elements can be said to be in a 'marked' position, in the sense that they can only occupy this position under specific circumstances, which we can formalise as a morphosyntactic feature on the dislocated element, [+top] and [+wh] in the cases at hand. Furthermore, the dislocation must be *observable*, that is, it cannot be string-vacuous. It has to be visible in a substantial number of instances for it to be detectable.

If we apply these criteria to head movement, it is immediately obvious that head movement is often problematic. Specifically, morphologically motivated, morpheme-collecting movement does not have the property that the moved element is in a marked position, or, equivalently, that it can appear in two different positions depending on context. Usually, the position in which the relevant element appears in is the *only* position it can appear in.

For example, V-to-T movement in German does not alter the position of the verb relative to other constituents:

(34) a. … dass er das Buch liest.
… that he the book reads
'… that he is reading the book.'

b. … dass er das Buch lesen wird.
… that he the book read will
'… that he will read the book.'

[21] In fact, in German, a non-topicalised object, especially a definite object, may scramble and thus appear anywhere between the finite verb and the main verb (*wollte* 'wanted' and *sehen* 'to see' in the example). This complication is not relevant to the issue at hand, however.

In the German verb complex, the finite verb appears at the right edge, while infinitives precede it. The position of *liest* 'reads' is not observably different from the position of *lesen* 'to read', however: it follows the object *das Buch* 'the book'. Therefore, there is no direct empirical evidence for the assumption that *liest* has been moved, which is why V-to-T movement is motivated on theoretical grounds.

One obvious exception to this is V-to-C movement in verb-second languages such as German. In German and Dutch, the finite verb can appear in the first or second position of the clause or, alternatively, in clause-final position as part of the verbal complex.[22] As has been argued by, e.g., Zwart (2001) and Truckenbrodt (2006a), V-to-C movement has a semantic effect, as well. As such, it is virtually inevitable to treat V-to-C movement as a real case of movement (or LD-dependency, if one prefers the term): the V2 position is marked and is clearly distinguishable from the clause-final position of the verb in embedded clauses.

3.2.3 Morphologically motivated head movement

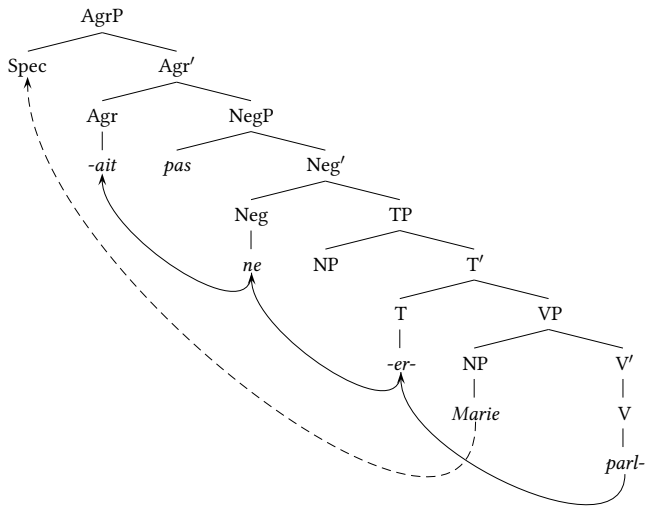
Although head movement often lacks a clear empirical basis, it is nonetheless a staple of (mainstream) generative theories. Many instances of head movement that lack such an empirical basis are *morphologically* motivated: head movement is assumed to take place in order to account for the observation that two or more elements that theoretical considerations place at different positions in the tree, nonetheless appear as a single word form. Take the following example, adapted from Pollock:

(35) Marie **ne** **parl-er-ait** pas.
Marie NEG speak-COND-3SG NEG
'Marie would not speak.'

Pollock analyses the form *ne parlerait* as the result of a series of movement operations in which the V head (*parl*) moves to T to combine with *-er*, after which the complex T-V form moves to Neg to pick up (part of) the negation. This is followed by a movement of the Neg-T-V complex to Agr, which establishes agreement with the subject and adds the suffix *ait*:

[22] With its position in the verbal complex depending on various factors, which are not relevant to the present discussion.

(36)



A more recent analysis would probably not include the negative element *ne* as part of the verb form, because it appears before the stem, not after it (cf. Julien 2002, ch. 4). The general principle of the analysis has not changed, however: a stem moves up the tree from one head position to the next, picking up bound morphemes as it goes.

It should be noted that one of Pollock's arguments in favour of head movement of V is the position of the (finite) verb relative to adverbs. However, this argument only holds if adverbs are immobile, something that Pollock explicitly assumes (cf. Pollock 1989, p. 372-373). Although there is little reason to assume that adverbs can move (except into a topic or focus position), it is not impossible that their placement may be flexible.

Reviewing Pollock's data in detail is beyond the scope of the present discussion, but there is some discussion in the literature (e.g., Ayoun 1999, 2000; Iatridou 1990) that suggests that the data are not as clear-cut as Pollock suggests, in the sense that the exact placement of adverbs and negation with respect to finite an non-finite verbs (also) depends on the type of adverb/negation. If this is indeed correct, the relative position of verbs and adverbs/negation cannot tell us much about the movement of verbs.

3.2.4 An alternative analysis

The model outlined in the previous chapter provides us with a general solution to the problem that morphologically motivated head movement operations pose. One argument in favour of this position is the observation that the movement

operations are often not necessary in syntax, because the syntactic tree before the (assumed) head movement operations can be interpreted properly by the semantic system. I illustrate the analysis here with the French example in (36). The analysis assumes that the morphemes making up the verb form *ne parlerait* are combined in phonology rather than in syntax. In order to make this possible, we need to assume that for example the T head is represented by the following lexical entry:

(37) T ↔ əʀ|$_{\omega}$

In this lexical entry, Φ(T) consists of the syllable /əʀ/ combined with a phonological alignment requirement. As discussed in section 2.7.2, I assume that a suffix is characterised by the fact that it has an alignment requirement with a prosodic word. As a result, suffixes are subject to Input Correspondence, which requires that the element /əʀ/ attach to Φ(V) in phonology.

In this way, we can attach /əʀ/ to /paʀl/ without having to resort to head movement. The suffix /ɛ/ (*-ait*) can be attached in the same way. Similarly, the negative particle *ne* can be attached to the verb stem as a prefix if we assume that it has the following lexical entry:

(38) Neg ↔ $_{\omega}$|nə

Here, the prosodic alignment requirement is on the left rather than on the right. In this way, *ne* is marked as a prefix. Whether this is the best way to treat the negative particle *ne* in French is a question that I will not go into here. The important point is that it is not necessary to assume that there is such a thing as morpheme-collecting head movement in syntax. Head movement operations that only apply to attach an affix to a stem are superfluous, because there is no need to form a distinct subtree in syntax containing just the stem and the affixes.

Instead, we can apply the mechanism described in chapter 2 to ensure that affixes are attached to their stems. By encoding the affixal nature of elements such as *-er*, *ne* and *-ait* in their phonological forms as a prosodic alignment requirement, we create elements that are subject to Input Correspondence, which allows us to attach the affixes to the verb stem without V having to move.

3.2.5 Head placement

Treating affixes as prosodic elements works well to ensure that they are combined with the stems they attach to. The analysis does not account for the placement of the stem+affix combination in the clause, however. Consider the following pair of sentences:

(39) a. Jean embrasse souvent *t* Marie.
Jean kisses often Marie

b. John often kisses Mary.

This pair of sentences, which goes back to Pollock (1989), is used to argue for the assumption that French has V-to-T movement, while in English, the verb remains in its base position.[23] The position of the verb relative to the adverb *souvent* / *often* is argued to show this: if the adverb adjoins to VP universally, then obviously the verb *embrasse* in French has moved out of the VP, while in English it remains in its base position.

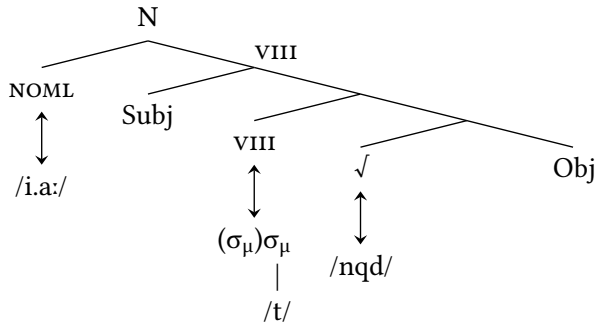
The problem here is not how the verb stem and the ending are combined: the mechanism for this is the one described above. Rather, the problem is how we can account for the fact that in French, the verb appears to be in a higher position than in English. Note that this problem is not unique to the analysis proposed here. The standard analysis, which assumes head movement, has the same problem in reverse: it must account for the fact that *kisses* in English combines with the suffix *-s*, although it does not actually move to T, where *-s* is assumed to be base-generated. The traditional analysis assumes *lowering* of the affix, which is a problematic notion because of the Extension Condition (Chomsky 1995b), which states that all operations must take place at the root. For this reason, lowering has sometimes been reanalysed as an operation that applies after syntax (e.g., Embick and Noyer 2001).

In the head movement analysis, the default, so to speak, is for the stem+affix combination to appear in the position of the affix (which is structurally higher). In the analysis proposed here, the default is the exact opposite: the stem+affix combination appears in the position of the stem. The sign language examples are quite clear in this regard: the non-manual negation is structurally higher than the verb, but Φ(V+Neg) appears in the position that Φ(V) occupies.

This, however, is problematic for French finite verbs and even more so for Arabic verbal nouns. Consider again the structure of the verbal noun construction in example (27) from chapter 2, repeated here as (40):

[23] Or, alternatively, that V-to-T movement is covert in English, i.e., takes place after spell-out.

(40)



If Φ(√)+Φ(VIII)+Φ(NOML) would be placed at the position of Φ(√), the structure would yield the wrong order, since the verbal noun precedes the subject, as shown in example (18) of chapter 2, repeated here as (41):

(41) ʔaqlaqa-nī -ntiqād-u -l-rajul-i -l-mašrūʕ-a
annoy.3sg.m-1sg.OBJ criticising-NOM DEF-man-GEN DEF-project-ACC
'The man's criticising the project annoyed me.'

Consequently, we need some way to ensure that the word form *intiqād* is placed in the right position. The first idea that comes to mind, based on the tree in (40), might be to say that the word form is placed in the position of the highest head involved, which would be the position of Φ(NOML) in the current example. However, when we examine the *masdar+li* construction, we immediately see that the former position is not tenable. Consider again the data:

(42) ʔaqlaqa-nī -ntiqād-u -l-rajul-i li -l-mašrūʕ-i
annoy.3sg.m-1sg.OBJ criticising-NOM DEF-man-GEN to DEF-project-GEN
'The man's criticising of the project annoyed me.'

Because the object in this construction is licensed with a preposition, we must assume that there is no verbal projection. Instead, the verb root is nominalised immediately, by merging NOML after the stem VIII marker, but before the object and the subject. If the word form *intiqād* were placed in the position of Φ(NOML), the expected order would be SON, contrary to reality.

The only other alternative is that there is a higher head position in the structure, one that c-commands the subject, and that NOML occupies this position. There is an obvious candidate for this head: D. Its presence in the constructions in (41) and (42) is a common assumption and arguments in favour of it are fairly straightforward. As has been amply discussed in the literature (cf. Borer 1999;

Fassi Fehri 1993; Kremers 2003; Siloni 2001, and related work), the combination of the masdar and the subject forms a so-called *genitive construction*. The genitive construction is a construction in which a head noun and a genitive-marked dependent noun form a tight syntactic unit. The head noun and the dependent are strictly adjacent, not even adjectives can break up the sequence:

(43) a. bayt-u -l-rajul-i -l-kabīr-u
house-NOM DEF-man-GEN DEF-large-NOM
'the man's large house'

b. *bayt-u -l-kabīr-u -l-rajul-i
house-NOM DEF-large-NOM DEF-man-GEN

In the examples in (43), the adjective *kabīr* 'large' modifies *bayt* 'house', not *rajul* 'man', as indicated by the case ending. However, it must follow the genitive noun. Placing it in the position immediately adjacent to the head noun leads to ungrammaticality.

A further typical property of the construction is the fact that the head noun, here *bayt*, has a special morphological form called *construct state*,[24] which is marked by the total absence of definiteness marking. A noun in (Standard) Arabic can have one of three states: apart from the construct state, there is the *definite state*, marked by the definite article *al-*, and the *indefinite state*, marked by one of a small number of indefiniteness markings (the most common one being the suffix *-n*).[25]

In Kremers (2003), I argue that the D head is crucial in establishing the genitive construction. The D head is phonologically null, but it carries a [+poss] feature which enables it to assign genitive case. Specifically, I argue that genitive case in Arabic is a structural case that is assigned through a standard Agree process between the D head and the highest nominal in its c-command domain. In the case of masdars, this is the subject if one is present and the object otherwise.

Longobardi (1994) argues that an empty D head cannot survive in syntax. The noun must move to it in order to provide it with semantic content. I adopt this assumption, although in a somewhat different form. First, the notion of an "empty

[24] In the generative literature, the term *construct state* is often applied to the genitive construction, which is also, at least informally, common practice in (Western) philological traditions of Arabic. Technically, however, this use is incorrect: as explained in the text, the term *construct state* actually refers to a particular morphological form of the noun. For this reason, I prefer to use the term *genitive construction*.

[25] It should be noted that the spoken Arabic vernaculars, which are the actual native languages of speakers of Arabic, do not have case endings and lack all forms of indefiniteness marking. The construct state is still marked morphologically, however, by the absence of the definite article (usually *il-*). Furthermore, feminine nouns, which end in the feminine suffix *-a*, change this suffix to *-it* in the construct state.

D head" is ill-defined in the current model, and in the sense it was intended by Longobardi, it is useless here. For Longobardi, an empty D head is a head with no phonological content. However, in this sense, all morphosyntactic heads are empty in the current model, because the phonological features are not part of the syntactic structure. Morphosyntactically, the construct state D head is definitely not empty, because it contains at the very least a [+poss] feature.[26]

To implement Longobardi's idea, I assume that a D head requires an index, a syntactic feature modelled on the equivalent feature in HPSG: the index feature allows the nominal to be identified syntactically; it consists of a set of φ-features that enable this identification and that are available for agreement. A D head that does not possess an index must obtain one, normally from the noun that it embeds. The D head in a construct state lacks this index feature but obtains it through movement of N to D.

Introducing the notion of head movement at this point raises a few questions. First and foremost, above I argued that movement should be detectable. Detectability is best achieved by cases in which the relevant element is not required to appear in its target position but can, depending on context, also appear in its base position. Additionally, the two positions must be distinguishable on the basis of the linear string. V-to-C movement is a typical example, because the position of V with respect to its argument(s) changes: a verb that has moved to C will precede its subject and object, while the unmoved verb follows its subject and, depending on language, also its object.[27]

Similar, though more subtle, variability is available for Arabic nominals. In non-construct state constructions, the head noun of a noun phrase can be preceded by the demonstratives *hāḏā* 'this' and *ḏālika* 'that' (which additionally require the presence of the definite determiner *al-*). In a genitive construction, however, the head noun is always the very first element in the phrase: demonstratives may not appear before the noun. Instead, they are placed after the genitive dependent, much like adjectives:

(44) a. hāḏā -l-bayt-u
this the-house-NOM
'this house'

[26] In my analysis in Kremers (2003), it also contains an unvalued [def] feature, which accounts for the so-called *definiteness effect*: the definiteness of the entire genitive construction depends on the definiteness of the dependent (genitive- marked) noun. The head noun itself has no definiteness marking.

[27] I woud argue that relative placement with respect to arguments is the strongest kind of evidence, because arguments tend to have fixed positions. As hinted at above, relative placement with respect to adverbials, for example, is much more problematic, because it is possible that the placement of adverbs is less constrained. In other words, if the order V-Adv co-occurs with the order Adv-V, it is not clear whether it is V or Adv that occupies different positions.

b. bayt-u -l-rajul-i hāḏā
house-NOM DEF-man-GEN this
'this house of the man's'

This strictly phrase-initial position suggests that the noun is not in the same position that it occupies in non-genitive constructions. Furthermore, the so-called *definiteness effect* is a further indication that the noun moves to D in the genitive construction. The head noun of a genitive construction, itself not marked for definiteness, inherits a formal definiteness feature from its genitive dependent:

(45) a. riwāy-at-u mu?allif-i-n jadīd-at-u-n / *al-jadīd-at-u
novel-f-NOM author-GEN-INDEF new-f-NOM-INDEF / *the-new-f-NOM
'an author's new novel'

b. riwāy-at-u -l-mu?allif-i *jadīd-at-u-n / -l-jadīd-at-u
novel-f-NOM the-author-GEN new-f-NOM-INDEF / the-new-f-NOM
'the author's new novel'

Attributive adjectives in Arabic agree with the noun they modify in definiteness: if the noun is definite, the adjective takes a definite determiner *al-*, if it is indefinite, it takes the indefiniteness suffix *-n.* This fact can be used to determine the definiteness of the head noun. The adjective in both examples in (45) modifies the head noun, as witnessed by the agreement in both case and gender.

The example shows that the adjective appears to agree in definiteness with the genitive modifier: both are either definite or indefinite, they cannot have differing values for definiteness. There is, however, no formal mechanism that could establish an agreement relation between the genitive dependent and the adjective. Rather, we are forced to assume that there are two agreement relations: one between the genitive dependent and the head noun (often called *definiteness inheritance*) and one between the head noun and its attributive adjective.

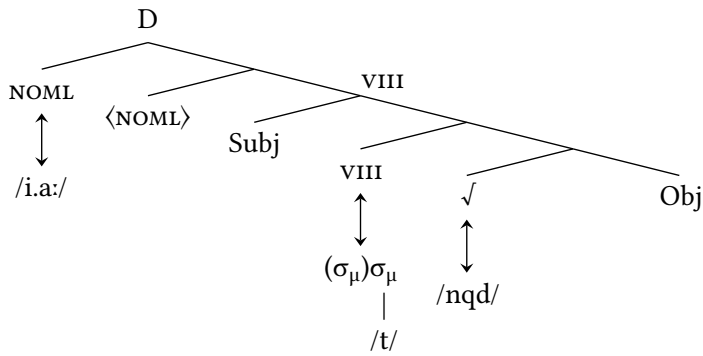
The details of these agreement processes are not relevant to the present discussion.[28] What is relevant here is the fact that the head noun obviously exhibits a definiteness feature and that this feature may be taken as evidence for the assumption that in the genitive construction, the noun occupies the position of D°. In other words, despite the fact that the effects are subtle, there seems to be

[28]See Kremers (2003, ch. 3.1) for an analysis of the definiteness inheritance between head noun and genitive dependent, and Kremers (2003, ch. 4.3) and Struckmeier and Kremers (2014) for definiteness agreement between noun and adjective.

sufficient reason to assume that N-to-D movement is detectable in Arabic noun phrases.[29]

The only remaining question is which head is moving to D in the verbal noun construction. Since head movement to D takes place in order to provide D with an index, it seems unlikely that it is the root √ or the stem marker *v*, since neither is equipped with the necessary index. The only head that can be plausibly assumed to provide such an index is the nominalising head NOML. Let us therefore assume that it is indeed NOML that moves to D. The resulting structure is (46), with angle brackets indicating the trace:

(46)



It is important to note here that although in syntax only NOML moves, phonology has to respond by positioning the entire word form containing Φ(NOML) in the corresponding position. The reason for this is obvious: Φ(NOML) consists of the vowels /i.a:/, which do not constitute a phonological constituent. As a result, the word form *intiqād* ends up phrase-initially.

In light of the discussion in chapter 1, it may seem surprising to adopt the notion of a D head that can be the target for movement, given that I argue in section 1.2.3 that features should always be connected to phonologically overt material and that they do not constitute independent elements, which means that they cannot move or be moved to independently of some phonological form. Nonetheless, this seems to be exactly what happens in the analysis proposed here: the D head in a genitive construction is assumed to be phonologically empty, yet N targets it for movement.

It should be kept in mind, however, that *movement* is really just a theoretical abstraction. It expresses the observation that under certain circumstances, an

[29]Note that the cues described here correspond to the traditional arguments for assuming N-to-D movement in the genitive construction (see the references above).

element can occur in what might be called a marked position. In this position, it has additional features that it does not have in its unmarked position. In the case at hand, the noun is in initial position in the noun phrase and has a [±def] and a [+poss] feature that it normally does not have. How exactly the noun ends up in its noun-phrase initial position and how it obtains these additional features are question that a high-level theory of grammar does not necessarily have to answer (representative theories do not), but head movement and some form of feature amalgamation are available as theoretical options if one prefers to have an account.

Put differently, the [±def] feature in the structures under discussion is a property of the construction, one that must be visible. The principle element that makes [+def] visible in the Arabic noun phrase is the definite determiner *al-*. Since in the genitive construction, the head noun lacks *al-* but still displays a definiteness feature, and because its position in this construction is compatible with the position of *al-* in nominals without a genitive dependent, the hypothesis that a construct state noun occupies the position of *al-*, i.e., D°, is a reasonable one.

What this means is that positing a head movement operation does not conflict with the notion of features discussed in chapter 1, as long as movement is observable. The question when exactly movement is observable in the sense intended here will have to be left open for future research, but N-to-D movement in Arabic seems to qualify.

That being said, one may wonder why we cannot simply assume that in the Arabic verbal noun construction, √ and VIII (i.e., little *v*) move to NOML, after which the entire complex head moves to D. After all, Φ(√) and Φ(VIII) are visible in the position of D as well. Though this is true, there is no *syntactic* trigger for these movements. N-to-D movement (NOML-to-D in the case of verbal nouns) is motivated by the D head's need to have an index, but similar motivations do not exist for √ moving to VIII or VIII moving to NOML.

3.2.6 Concluding remarks

Head movement is a staple of G&B and minimalist frameworks, though it is not without its problems. In a prosodic syntax model, certain instances of head movement, specifically, those that are motivated by the need for a stem to "pick up" its affixes, can be analysed in a different manner, by placing the actual formation of the affixed form in phonology, with the mapping from syntax being controlled by Input Correspondence. In this way, head movement is no longer necessary in syntax, which is a desirable result, because there is usually no true syntactic motivation for the proposed movement beyond the need for the stem to combine with its affixes.

It does not follow from the analysis, however, that all kinds of head movement should be treated in this way. In cases where head movement is visible in syntax, i.e., where it can be shown that a certain type of element occupies a marked position and carries additional features under specific circumstances, syntactic head movement is still an option. Feasibility in terms of a lower-level analysis is probably maintained if the relevant marked position can be shown independently to exist and to be associated with the relevant features.

3.3 Affixes and clitics

3.3.1 Introduction

As briefly mentioned in section 2.7.2, the prosodic syntax model proposed here offers a new way of looking at affixes and clitics, suggesting that these two kinds of elements are in fact closely related. More specifically, I argue that there is a range of elements going from bound morphemes on one side to free forms on the other, with a number of elements in between. In other words, what I will argue is that there is no discrete distinction between affixes, clitics and words. Rather, there are a number of properties that we need to distinguish, both morphosyntactic and phonological, and the combination of these determine the behaviour of individual elements.

The important point is that terms such as *affix, clitic* or *word* are not ontological primitives. They are convenience labels that sometimes obscure rather than clarify matters and which should therefore be used with care. The goal of the current section is to discuss and clarify the properties that are relevant in the description of affixes and clitics and which can help us establish a basic overview of affix and clitic types.

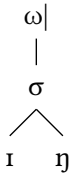
3.3.2 Affixes

First, let us look at affixes. As already mentioned above, I propose that an affix is a syntactic head with no special morphosyntactic properties. That means that the syntactic system does not "know" that an element is an affix: it just sees a head. Rather, it is the phonological form of an element that tells us that an element is an affix. Let us consider a few cases to see how this idea works out.

English '-ing'

The first example to be discussed here is the English gerund affix *-ing*. As mentioned earlier, I adopt the following representation of *-ing*, repeated here from example (32) in chapter 2:

(47)



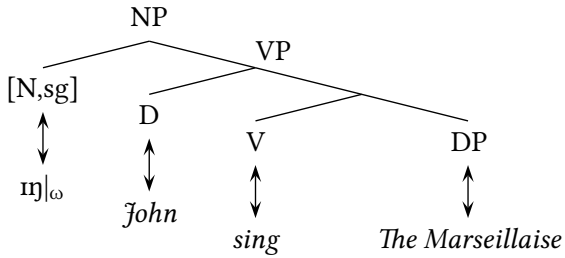
As is well-known, English allows three types of gerund constructions, for which Abney (1987) adopts the terms *Acc-ing*, *Poss-ing* and *Ing-of*, respectively:

(48) a. John singing the Marseillaise
b. John's singing the Marseillaise
c. John's singing of the Marseillaise

Abney attributes the different constructions to different attachment sites of the suffix *-ing*: in (48a), *-ing* attaches to the IP, turning it into a DP. Since the IP is present in the structure, a clausal licensing mechanism is available for the subject, which accordingly takes accusative case. In (48b), *-ing* attaches to the VP, turning it into an NP. Since the object is VP-internal, it can be licensed by assigning it accusative case, but the subject must be licensed through a nominal mechanism and consequently receives genitive case. Lastly, in (48c), *-ing* attaches to V, turning it into N. As a result, no verbal licensing mechanisms are available, so that even the object must be licensed by a nominal licensing mechanism.

The core idea of Abney's can be easily transferred to the current model: *-ing* is a nominalising head (i.e., it has category N) that attaches at different levels in the structure. Since it attaches to a verbal category (or its extended projection), it must have a syntactic selectional feature [*u*V]. Consider, for example, the case of *-ing* attaching to VP:

(49)



In the mapping to phonology, Input Correspondence applies to *-ing* because of its prosodic requirement and requires it to attach to Φ(V), the phonological

constituent of the head of the structure selected by *-ing*. The prosodic alignment requirement on Φ(*-ing*) attaches it to the right of Φ(V), yielding the form /sɪŋɪŋ/.

Tagalog '-um-'

The analysis works well for suffixes such as *-ing*, and it is easy to see how it can be adapted to prefixes. Infixes require a more specific treatment. As an example, I discuss one of the classics of Optimality Theory: Tagalog *-um*-affixation (Prince and Smolensky 2004, pp. 40–43). As is well-known, the Tagalog affix *-um-* (often called the *agent trigger*, i.e., it signals that the noun marked with so-called "direct" case is the agent of the verb) is realised as a prefix on stems that begin with a vowel but as an infix on stems that begin with one or more consonants:

(50) a. *aral – um·aral*

b. *sulat – s·um·ulat*

c. *gradwet – gr·um·adwet*

In Optimality Theory, these data are explained by the interaction of two constraints: NoCoda, which states that syllables should not have codas, and AlignL, which aligns the affix with the left edge of the stem. The following tableau shows the derivation of *um·aral*:

(51)

		NoCoda	AlignL
☞	u.ma.ral	*	
	a.um.ral	**!	a
	a.ru.mal	*	a!r
	a.ra.uml	*	a!ra
	a.ra.lum	*	a!ral

The constraint NoCoda ensures that the affix precedes the stem. If the stem starts with a consonant, however, this same constraint has the effect of moving the affix into the stem:

(52)

		NoCoda	AlignL
	um.su.lat	**!	
☞	su.mu.lat	*	s
	su.um.lat	**!	su
	su.lu.mat	*	su!l
	su.la.umt	*	su!la
	su.la.tum	*	su!lat

Since *-um-* changes the argument licensing properties of the verb, we may assume it is an instantiation of the *v* head, even though nothing really hinges on this assumption. The model described by Prince and Smolensky uses two constraints. Of these, AlignL does not really say anything more than that *-um-* is a prefix of sorts. The question we need to answer here is which of the following two options is better: to formulate a general but violable principle that prefers left alignment, or to simply incorporate this information in the lexical entry of *-um-*? Since Tagalog also has suffixes, AlignL is at best a tendency, not a constraint that applies to all affixes. I therefore assume that it is better to incorporate the alignment property into the lexical entry, in the same way that the suffix status of *-ing* is encoded in its lexical entry.

The NoCoda constraint expresses a more general property of language: there is a general tendency to avoid codas, which expresses itself in many ways, perhaps most prominently in the fact that CV is the most basic syllable structure of human language. What we should ask ourselves, however, is whether we should incorporate such a constraint into the grammar or whether we are dealing with a third-factor effect. For example, there are strong indications that this general preference for onsets vs. codas has to do with the way the brain times phonetic events (cf. Goldstein, Byrd, and Saltzman 2006; Nam 2007).

I therefore assume that the coda dispreference is not a rule of grammar but the result of architectural constraints of the brain that are essentially independent of language. In other words, NoCoda is not part of UG, nor is it part of the I-language of a speaker. Rather, it is a factor that influences the development of an I-language in an individual. As a result, a language may incorporate rules into its grammar that in some form or other disprefer codas, such as phonotactic restrictions, lexical or postlexical phonological rules, etc. However, the extent to which a language incorporates such rules will differ from language to language, because there may be other factors in the development of I-languages or in the historic development of a particular language that push the language toward having codas. For example, if in a language that has CV syllables word-final vowels are reduced to schwa and then dropped, CVC syllables result.

Therefore, rather than adopting a NoCoda constraint that is a violable rule of grammar, I propose to encode the coda-avoiding property of *-um-* directly into its phonological component: despite appearances, *-um-* is not a single syllable, it is bisyllabic. There is a syllable boundary between the two segments, or put differently, /m/ must occupy an onset position.[30] We can express this by representing the morpheme as /u.m/, or, more elaborately, as in (53):

[30] In fact, McCarthy and Prince (1993, p. 79) observe this themselves when they state that *-um-* "falls as near as possible to the left edge of the stem, so long as it obeys the phonological requirement that its final consonant *m* not be syllabified as a coda".

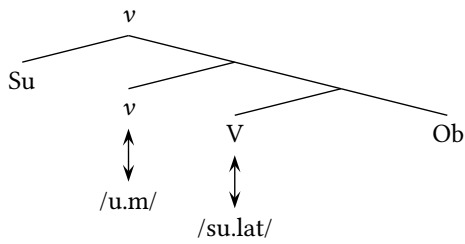
(53) $\left[v,+\text{ag}\right] \leftrightarrow {}_{\omega}|\sigma \quad \sigma$

u m

Here, the feature [+ag] is simply meant to represent the fact that *-um-* is an agent trigger. For convenience' sake, I have left out the semantic part of the lexical entry. The important part in (53) is the phonological part: a bisyllabic structure with two segments occupying slots in different syllables. Additionally, there is a prosodic word boundary at the left edge. This ensures that /u.m/ is left-aligned and is positioned as closely to the left edge of the stem as possible.

Analysing *-um-* in this manner turns it into a prosodic morpheme, with the additional effect that it is subject to Input Correspondence. On the assumption that *-um-* is syntactically a *v*-head, a clause containing *-um-* has a syntactic structure along the lines of (54):

(54)



By now it should be obvious that we do not need to assume much more than this structure: Input Correspondence ensures that Φ(*v*) selects Φ(V) and the phonological requirement that /u.m/ is left-aligned and bisyllabic can only be accommodated by creating the form *sumulat*.

In the prosodic syntax model, affixes are characterised by two properties: they have a c-selectional feature in syntax, and a prosodic requirement in phonology. The combination of these two properties together with the mapping principle Input Correspondence ensures that affixes are attached to the phonological component of the head of the category that they select. There is no particular reason, however, why both properties should always occur together. When they do not, we obtain clitics.

3.3.3 Clitics

As a descriptive term, 'clitic' refers to elements that sit somewhere between affixes and words. Elements referred to as clitics are usually elements that cannot

stand on their own, just like affixes, but unlike affixes, they often attach to different classes of words, or, if not that, then at least to different subclasses of the same class (e.g., Romance object clitics). Apart from this rather vague classification, there does not seem to be much that clitics have in common. Consider, for example, some of the properties of a few elements commonly classified as clitics.[31]

- Romance object clitics attach either to the left or to the right of their host, depending on the host's morphosyntactic form.[32]
- English reduced auxiliaries occupy the position of full auxiliaries in the clause and attach to whatever word precedes.
- English possessive *'s* attaches to the last word of the phrase it is syntactically associated with, regardless of the word's category.
- Second-place clitics occupy the second position of the phrase they belong to, without there having to be any direct syntactic relation between the clitic and its host.

Interestingly, though, a similar confusion surrounds the term 'word'. Words are equally notoriously difficult to define, as illustrated by Haspelmath (2011), who argues that there is no general, cross-linguistic property that distinguishes words from phrases and our intuitions in this regard are all too often shaped by our writing system. The solution that is usually chosen in order to deal with this fact is to separate out the various properties of words, yielding three separate notions: prosodic word, morphosyntactic word, and lexeme.

Having established these separate notions, we can, for example, explain why compounds[33] sometimes behave as if they are two words: phonologically, they *are* two (p-)words. Similarly, it is no longer problematic to have lexemes that consist of more than a single morphosyntactic word, such as particle verbs (*to shut up*, *to write off*, etc.). That is, by separating out the semantic, morphosyntactic and phonological properties of elements that we traditionally call 'words', we can actually distinguish different types of words, each with its own particular behaviour.

The purpose of this section is to do something similar for clitics. That is, I assume that the term 'clitic' is essentially a pre-theoretical notion, just like the

[31] Note that I use the term *host* to refer to the element to which the clitic attaches in phonology. This of course does not have to be an element with which the clitic has any *syntactic* relation.

[32] I do not discuss Romance object clitics, mainly because of their placement properties, which would require a more in-depth look into Romance sentence structure than I can provide here.

[33] And some stem+affix combinations, such as Dutch *roodachtig* 'red-like', which displays final devoicing of the 'd' in *rood* 'red'.

term 'word'. It has no theoretical status in linguistics, it is merely an imprecise but occasionally convenient shorthand. The discussion will focus on the morpho-syntactic and the phonological properties of clitics. The semantic properties are not relevant to the current discussion and will be ignored.

Klavans (1985) proposes three parameters along which clitics can vary, which are summarised by Anderson (1992, p. 203) as follows:

(55) a. The clitic is located within some syntactic constituent (S vs. VS vs. NP, etc.) which constitutes its domain.

b. The clitic is located by reference to the {first vs. last vs. head} element of a specified sort within the constituent in which it appears.[34]

c. The clitic {precedes vs. follows} this reference point.

Although these parameters are descriptively correct, it is not necessary to specify them in the lexicon for each clitic that a language has. In many cases, the placement of a clitic follows from the syntactic structure of the language in question. In the sections that follow, I discuss several clitics that show what features are necessary for a description of most clitic types. The analysis shares several insights with the analysis of Anderson (1992), although the framework in which it is couched is quite different.

Latin '-que'

The Latin conjunction *-que* is a typical second-place clitic, appearing after the first element of the second conjunct:

(56) bon+ī puer+ī bon+ae-que puell+ae
good+.pl.f boy+pl.m good+pl.f-and girl+pl.f
'good boys and good girls'

In (56), the second conjunct is *bonae puellae* 'good girls', the element *-que* breaks up this sequence. Embick and Noyer (2001) analyse *-que* in terms of their Distributed Morphology model and argue that it is positioned after the first morphological word of the second conjunct. They appeal to the concept of *Local Dislocation*, a type of movement operation that is sensitive to the morphemes involved and can only apply to string-adjacent elements.

However, there is a major problem for Embick and Noyer's analysis, something they even note themselves. There are data that the analysis cannot account for:

[34] The 'head' option is not part of Klavans's (1985) original definition and seems to have been added by Anderson in order to account for Romance object clitics.

(57) a. circum-**que** ea loca
around-and those places
'and around those places'

b. contrā-**que** lēgem
against-and law
'and against the law'

(58) a. in rēbus-**que**
in things-and
'and in things'

b. dē prōvinciā-**que**
from province-and
'and from the province'

When *-que* conjoins prepositional phrases, its placement depends on the size of the preposition. If the preposition is disyllabic, *-que* attaches to the preposition. If it is monosyllabic, however, *-que* is placed after the first word of the preposition's complement.

For Embick and Noyer, this fact is difficult to explain. The preposition is a morphological word, no matter whether it is mono- or disyllabic. They argue that in the case of monosyllabic prepositions, a string-vacuous instance of Local Dislocation turns the preposition and the following word into a single morphological word. This, however, is merely an *ad hoc* technical solution that does not provide an explanation of the facts.

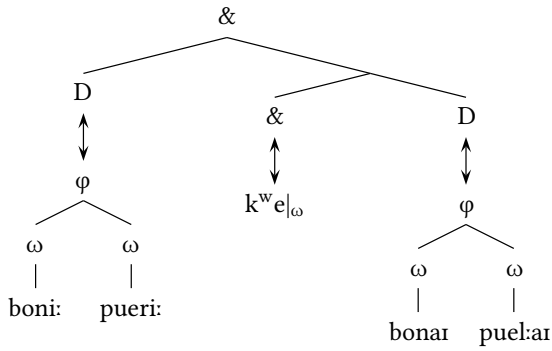
A more insightful account presents itself when one considers the prosodic structure of the examples. Although Latin is not a living language and we can only speculate about its prosodic structure, it seems safe to assume that monosyllabic prepositions did not constitute prosodic words, while disyllabic prepositions did (cf. Agbayani and Golston 2010 for an extensive discussion of this point). Assuming this to be correct, we can give the following lexical entry for *-que*:

(59) AND $\leftrightarrow$ &° $\leftrightarrow$ $k^we|_\omega$

The phonological form of *-que* has the same kind of prosodic requirement that the gerund suffix *-ing* has: it must be right-aligned to a prosodic word. Unlike *-ing*, the syntactic head of *-que* does not have any selectional restrictions, however. Being a conjunction, it can conjoin phrases of any type.

Let us see what happens when *-que* conjoins two noun phrases. Assuming an &° head for the conjunction, the syntax tree looks like (60):

(60)

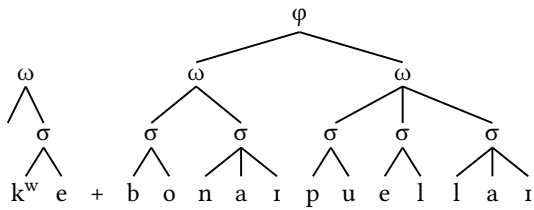


Embick and Noyer's account relies on a movement operation (called *local dislocation*) that moves *-que* to the position following *bonae*. This is a so-called *morphological* operation, which means that it applies after syntax but before Vocabulary Insertion (the operation that attaches phonological forms to syntactic nodes), i.e., just before the syntactic structure is transferred to phonology.

There is no need, however, to move the &° head to a position following *bonae* in syntax (or, as in the DM model, before Vocabulary Insertion). If we assume that the conjunction forms a prosodic domain with the second conjunct,[35] we can account for the placement of *-que* by appealing to Left-to-Right Association, a standard principle of autosegmental phonology.

Let us discuss this process in more detail. The two elements that need to be combined by the phonological system are the conjunction /kʷe/ and the second conjunct /bonaɪ puel:aɪ/:

(61)

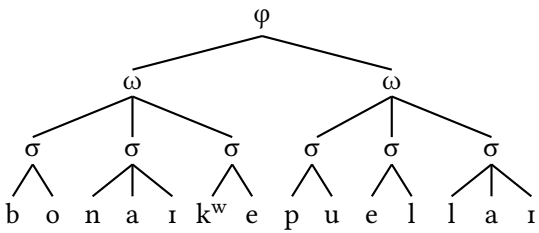


As discussed, /kʷe/ must align with the right edge of a prosodic word. The principle of Left-to-Right Association says that the first prosodic word that is

[35] There are some indications that this is quite generally true: for example, it is usually easier to insert a pause between the first conjunct and the conjunction than between the conjunction and the second conjunct, cf. *a man — and a woman* with *a man and — a woman*: the latter seems more marked. In some languages the conjunction actually cliticises onto the second conjunct, making separation almost impossible, cf. Arabic *al-nahār wa-l-layl*, lit. 'the-day and-the-night', i.e., 'day and night'.

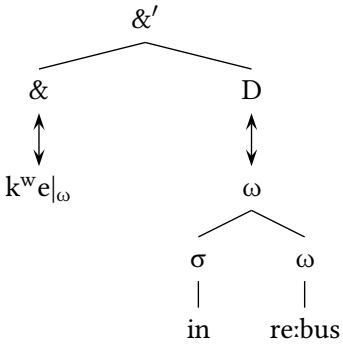
available should be targeted, which in this case is /bonaɪ/. Note that the first available prosodic word is not /boni:/ of the first conjunct, because we assumed that the conjunction forms a prosodic domain with its second conjunct and the prosodic word to which Φ(*-que*) attaches must be found in this domain. The resulting structure is depicted in (62):

(62)



The analysis can also account for the behaviour of *-que* in the examples in (58). The syntactic structure of the (a) example (with the first conjunct left out) is (63):

(63)



It should be clear that the structure in (63) yields a phonological form in which /kʷe/ follows /re:bus/, not /in/: the first prosodic word in the structure is /in-re:bus/, since /in/ by itself does not constitute a prosodic word.

The phonological component of *-que* has the same structure as the phonological component of *-ing*: both consist of a single syllable and both have a prosodic alignment requirement. The crucial difference between the two elements is the fact that the syntactic component of *-ing* c-selects a verbal category, while the syntactic component of *-que* has no c-selectional restrictions. For this reason, *-que* is not subject to Input Correspondence. It is still a prosodic morpheme, of course, being realised on an autosegmental tier, which is the mechanism that enables Φ(*-que*) to appear in a position that does not correspond to its position

in the syntactic tree. The difference with *-ing* is that Input Correspondence does not dictate to which prosodic word *-que* must attach.

English reduced auxiliaries

English reduced auxiliaries represent a different kind of clitic. They do not occur in second position, but rather occupy the same position that their phonologically non-reduced counterparts occupy.[36] Let us look at a concrete example, the auxiliary *will* and its reduced form *'ll*. I assume the following lexical entry for the reduced form (which I refer to as *'ll*):

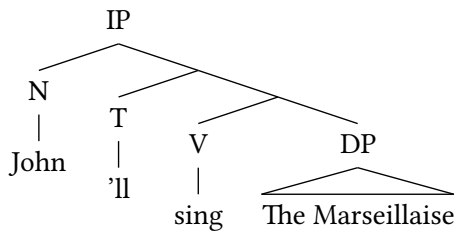
(64) FUT ↔ [T,*u*V] ↔ l

Again, the semantics is highly simplified. In syntax, *'ll* is a T head that selects a projection of V, the standard feature make-up of auxiliaries. In this respect, *'ll* differs from *-que*, which does not have a syntactic selectional restriction, but it is similar to *-ing*, which does have such a restriction.

Phonologically, however, the auxiliary differs from both *-ing* and *-que*. First, rather than being a syllable, it consists of only one segment, /l/, but more importantly, it does *not* have a prosodic requirement: the reduced auxiliary does not specify that it must right-align in a p-phrase. This means that it is not a prosodic morpheme and therefore it is not subject to Input Correspondence. That is, even though the syntactic head selects a (projection of) V, it is not the case that the phonological form must be associated with the head of the VP that the auxiliary selects. Instead, the reduced auxiliary is subject to Linear Correspondence, which means that its linear position is derived in the usual manner:

(65) a. John'll sing the Marseillaise.

b.



[36] That is to say, there are contexts in which auxiliaries cannot be reduced. These are generally contexts in which the auxiliary does not govern an overt verb, which implies that the reason for this impossibility is phonological: the auxiliary must carry the stress that would otherwise fall on the verb; cf. *I know he will* vs. **I know he'll* (but: *I know he'll do it*).

In this tree, the T head is structurally outside the VP, so that Φ(T) will be linearly outside Φ(VP). Given the value of the head parameter in English, Φ(T) appears before Φ(VP), leading to the word order in (65a).

Nothing so far indicates that the reduced auxiliary is a clitic. In fact, what has been said so far also applies to the non-reduced form *will*. The property that turns *'ll* into a clitic is the fact that it consists phonologically of a single segment /l/. Since a segment cannot constitute a prosodic word on its own, it must be incorporated into one.

English reduced auxiliaries are enclitics: they take the preceding word as phonological host (cf. Lakoff 1972). Klavans (1985) argues that the direction of phonological cliticisation is essentially arbitrary and must be specified for each clitic independently. Anderson (1992, p. 203), on the other hand, argues that a particular language's rules of *Stray Adjunction* determine the direction of cliticisation. Stray Adjunction refers to the process of incorporating phonologically deficient material (e.g., extrametrical syllables) into word-level phonological units for reasons of stress assignment. Stray Adjunction rules usually involve a preferred direction of incorporation, which, Anderson argues, is the direction of cliticisation for clitics such as *'ll*.

Booij (1996, p. 233) argues for a more nuanced position: he suggests that Stray Adjunction provides the preferred direction of cliticisation in a language, but specific contexts may trigger cliticisation in the opposite direction. Booij specifically looks at Dutch, which generally has enclitics, but clause-initially, (some) weak pronouns act as proclitics, because in this particular context, no other option is available. On the assumption that English, like Dutch, favours encliticisation, we can account for the fact that *'ll* takes the preceding (prosodic) word as its phonological host, regardless of its category.

At this point we have seen three types of elements: affixes, which have a c-selectional restriction and a prosodic alignment requirement, second place clitics, which have a prosodic alignment requirement but no c-selectional restriction, and "normal" clitics (for want of a better term), that have no prosodic alignment requirement but which are prosodically deficient, in the sense that they do not constitute a PWd. They may or may not have a c-selectional restriction, but this is irrelevant for their placement, because they are not prosodic morphemes and are therefore not subject to Input Correspondence.

Dutch 'het'

The effect of Stray Adjunction can be demonstrated very well with clitics whose syntactic and prosodic alignment properties do not match. An example of this category is the Dutch neuter definite determiner *het*, phonologically /ət/. *Het* and elements like it, such as the other Dutch determiners, *de* (non-neuter defin-

ite; phonologically /də/) and *een* (indefinite; phonologically /ən/) or the English determiners *the* and *a*, have the same phonological deficiency that other clitics have; that is to say, they are not PWds and therefore have to attach to some PWd in phonology.

Syntactically, determiners select the noun they combine with,[37] but phonologically, they do not (necessarily) attach to this noun. This is demonstrated in the following Dutch example from Booij (1996, p. 219):

(66) Jan kocht het boek.
$(jɑn)_ω$ $(kɔx_σ$ $tət_σ)_ω$ $(buk)_ω$
John bought the book
'John bought the book.'

The definite determiner *het* /ət/ is syntactically part of the noun phrase *het boek* 'the book', but phonologically it attaches onto the prosodic word /kɔxt/ 'bought', in line with the preferred direction of Stray Adjunction in Dutch, which is leftward. If there is no element preceding /ət/, however, it can encliticise onto the following PWd:

(67) Het boek ligt op tafel.
$(ət$ $buk)_ω$ $(lɪxt$ $ɔp)_ω$ $(tafəl)_ω$
the book lies on table.
'The book is on the table.'

In this example, the determiner *het* is sentence-initial and consequently has no preceding material it can take as a host. In such cases, the clitic can attach to the following word, in this case *boek* 'book', in violation of the preferred direction of Stray Adjunction.

Like *'ll*, Dutch *het* and elements like it are not prosodic morphemes and they are not subject to Input Correspondence. Their placement is therefore determined by syntax and Linear Correspondence. Their clitic behaviour follows from the fact that they are phonologically deficient, i.e., they do not constitute PWds, and must be incorporated into a PWd.

Dutch '-ie'

Booij (1996) argues that the principle that clitics obey Stray Adjunction has another exception, apart from the already observed fact that clitics may deviate

[37] Depending on the theoretical framework, they are also selected by the noun or, in early G&B terms, a spec-head agreement relation holds between the two.

from the preferred direction if they appear in a context in which they could otherwise not be attached to a PWd. According to Booij, individual clitics may be lexically specified as to their direction of cliticisation. A typical example is the Dutch element *-ie*, which is the clitic form of the third person singular masculine pronoun. Booij claims that *-ie* is explicitly marked as enclitic and therefore cannot occur clause-initially, unlike other clitic pronominals. Compare (68) with (69):[38]

(68) a. Daar was 'k nog nooit geweest.
There was I yet never been
'I had never been there yet.'

b. 'k Was er nog nooit geweest.
I was there yet never been
'I had never been there yet.'

(69) a. Daar was-ie nog nooit geweest.
There was-he yet never been
'He had never been there yet.'

b. *ie Was er nog nooit geweest.
he was there yet never been
'He had never been there yet.'

The clitic pronoun *'k* (pronounced /ək/) can precede or follow the finite verb and cliticises onto it, acting as an enclitic in (68a) and as a proclitic in (68b). (69) shows that *-ie* (pronounced /i/) can be enclitic but not proclitic. Booij argues that this fact is simply recorded in the lexical entry for the clitic. The first idea that comes to mind, then, is to describe this requirement as a prosodic alignment requirement, which we can specify in the clitic's lexical entry:

(70) $x_{3\text{sgm}} \leftrightarrow \begin{bmatrix} \text{D, m} \\ 3\text{sg} \end{bmatrix} \leftrightarrow \text{i}|_{\omega}$

This, however, is probably the wrong way to go, as it suggests that Dutch *-ie* is similar to Latin *-que*. The prosodic requirement would cause *-ie* to be mapped onto an autosegmental tier and its placement to be determined by Left-to-Right

[38]Note that rather than attaching clitics with an equal sign, I follow Dutch (informal) writing conventions in using an apostrophe to indicate a schwa (except at the end of the orthographic word) and attaching *-ie* 'he' to the preceding word with a hyphen.

Association, which is a phonological principle. The placement of *-ie* is clearly not determined by prosodic/phonological factors, however. Rather, it is syntactic.

In order to see this, we need to look at the placement of Dutch unstressed pronouns in general. Although scrambling in Dutch is much more restricted than in German, unstressed pronouns behave similarly (though not identically) to their German counterparts, in that they are scrambled closer to the front of the clause. They are placed in a position often referred to as the *Wackernagel position*, i.e., a position immediately following the complementiser or finite verb. Furthermore, the order of the pronouns in this position is S-DO-IO, whereas the order of non-pronominal arguments is S-IO-DO:[39]

(71) a. … dat Jan zijn vrouw het boek heeft gegeven.
that Jan his wife the book has given
'… that Jan gave his wife the book.'

b. … dat-ie 't 'r heeft gegeven.
that-he it her has given
'… that he gave it to her.'

In (71a), the indirect object *zijn vrouw* 'his wife' precedes the indirect object *het boek* 'the book', while in (71b) the order is reversed: the direct object *'t* 'it' precedes the indirect object *'r* 'her'. In both cases, the reverse orders are not possible.

These facts would be compatible with a prosodic account of clitic pronoun placement on the assumption that clitic pronouns in Dutch (and German) are second place clitics of some sort, in combination with a morphological template, in a way similar to Schütze's (1994) account of Serbo-Croatian second place clitics, assuming that we would be able to define the domain of their attachment appropriately. There is one fact, however, that is not compatible with a prosodic analysis. In Dutch (unlike in German), object pronouns cannot scramble before a subject:

(72) *… dat 't 'r Jan heeft gegeven.
that it her Jan has given
'… that Jan has given it to her.'

The only possible placement of the pronoun cluster in Dutch is after the subject. In spite of this, however, the order of the pronouns is still DO-IO:

[39] In German, S-IO-DO is the base order for non-pronominal arguments, but depending on information structure, any other permutation is possible. Dutch has the same base order, but scrambling is limited: S-DO-IO is possible if the indirect object is realised as a PP, but other orders are generally not allowed, although non-agentive verbs show more flexibility.

(73) a. … dat Jan 't 'r heeft gegeven.
that Jan it her has given
'… that Jan has given it to her.'

b. *… dat Jan 'r 't heeft gegeven.
that Jan her it has given
'… that Jan has given it to her.'

There is no possible way to target the position following the subject on the basis of prosodic criteria: the rule applies regardless of whether the subject is a pronoun (74a) or a noun phrase modified by a relative clause (74b):

(74) a. … dat-ie 't 'r heeft gegeven. =(71b)
that-he it her has given
'… that he has given it to her.'

b. … dat de man die ik gisteren het boek heb gebracht 't 'r
that the man REL I yesterday the book have brought her it
heeft gegeven.
has given
'… that the man that I brought the book yesterday has given it to her.'

The indicated positions are the only positions in which the clitic pronouns can appear and the order DO-IO is the only possible order. In this respect, Dutch differs from German, which does allow scrambling of object pronouns before a (non-pronominal) subject:

(75) … dass es ihr Hans gegeben hat.
that it her Hans given has
'… that Hans has given it to her.'

Within a minimalist framework, a straightforward explanation of this difference between Dutch and German is available with Struckmeier's (2011) analysis of the German clause. Struckmeier argues that in German it is not the subject but the entire *v*P that checks the EPP feature on T. The order of arguments inside the *v*P is S-IO-DO, but constituents can move out of the *v*P before it is moved to Spec,TP, yielding scrambled orders. Adopting this analysis, we can account for the Dutch and German pronoun facts and the differences between them if we assume that movement into the Wackernagel position is actually adjunction to *v*P. On the assumption that in German, the entire *v*P moves to Spec,TP, it follows that the pronoun cluster is placed immediately adjacent to the element in C°.

For Dutch, we must assume that checking of EPP is done in the same way as in English, i.e., by moving the subject to Spec,TP, and that non-pronominal constituents cannot move out of the *v*P. This ensures that the subject is always the first argument following the element occupying C°, that the pronoun complex in the Wackernagel position follows the subject, and that the arguments in the Dutch clause are always in their canonical order.

As stated, this analysis relies heavily on minimalist assumptions and is not readily adaptable to other frameworks. For this reason, I will not go into the obvious questions that it raises.[40] The important point is that placement of unstressed object pronouns in Dutch is clearly governed by syntax. Although *-ie* is a subject pronoun, there is no reason to assume that it has properties that differ from other unstressed pronouns, so that we must assume its placement is dictated by syntax as well.

A prosodic account of placement of *-ie* would not work for another reason. A prosodic requirement of the type that Latin *-que* has the consequence that the relevant element is placed after the first prosodic word in its alignment domain. For *-que*, this alignment domain is (by assumption) the phonological phrase constituted by the second conjunct, as discussed above. For subject pronouns, this alignment domain would most likely be the intonational phrase corresponding to the containing CP. The subject (in Dutch) is in Spec,TP, but TP is not known to correspond to any prosodic constituent, which makes the CP the first syntactic domain corresponding to a prosodic domain.

If this assumption were correct, we would expect *-ie* to appear after the first prosodic word in the clause, which is arguably correct in subclauses (since *-ie* appears directly after the complementiser), but not in main clauses: in main clauses, *-ie* appears after the finite verb, but this verb may be preceded by a constituent. As a result, the finite verb is usually not the first prosodic word in the alignment domain of *-ie*. A prosodic account would therefore predict a different position for the pronoun.

A prosodic account in which *-ie* is provided with a prosodic alignment requirement would therefore need to distinguish between *-que* and *-ie*: the former is an autosegment, the latter is not. The only way to do so appears to be to add a diacritic to (one of) these elements in the lexicon marking the appropriate autosegmental tier, which is obviously an undesirable solution. The conclusion that we are forced to draw, then, is that Booij's suggestion that the direction of

[40] For example, the variability in the argument order of psych verbs in Dutch would have to be accounted for by assuming variable base generation, which means that we actually have two sources of scrambling: base generation and movement out of *v*P. This may not be a very desirable option at first sight, but it is not unlikely that the first option is restricted to psych verbs, which are cross-linguistically exceptional when it comes to argument realisation (cf. Belletti and Rizzi 1988).

attachment of *-ie* is fixed in the lexicon is not tenable in the current model.

There is, however, another way to ensure that *-ie* cannot appear in proclitic positions. All unstressed (clitic) pronouns in Dutch have a schwa vowel, except for *-ie*, which consists of a tense vowel. It is conceivable that this fact is sufficient to account for the impossibility of *-ie* to appear as a proclitic. Phonologically, Booij (1996) argues that proclitics are syllables that are attached directly under a prosodic word:

(76) Proclitic prosodic structure:

That is, according to Booij, proclitics are not contained in a foot, which is a violation of the Strict Layer Hypothesis, but Booij argues that proclitics, being weak elements, cannot constitute a foot on their own. And since they cannot be incorporated into the following foot, given that feet in Dutch are left-headed, the only option is to attach them directly under the prosodic word.

Booij also argues that clitics are phonologically identical to affixes, so that the prosodic structure in (76) also holds for prefixes.[41] Looking at prefixes in Dutch, one can observe that they never contain a tense vowel. They either have a schwa (e.g., the verbal prefixes *be-* /bə/, *ge-* /ɣə/, *ver-* /və/, the verbal marker *te* /tə/ 'to') or a lax vowel (the verbal prefixes *ont-* /ɔnt/ and *er-* /ɛʀ/, and *ver-* /vɛʀ/ in a variant pronunciation), never a tense vowel.[42] We may assume that these facts reflect a rule of Dutch phonology, namely that a (ω-initial) syllable not embedded in a foot does not allow a tense vowel. Although admittedly a tentative proposal, it would allow us to account for the fact that *-ie* can only occur as an enclitic without having to take recourse to a diacritic in the lexicon.

The conclusion, then, is that the Dutch determiners and the clitic pronouns including *-ie* have the same properties as *'ll*: they have no syntactic c-selectional restriction and no prosodic alignment requirement, which means that their placement is determined by syntax. The fact that they behave as clitics can be attributed to the fact that they are phonologically deficient. The difference between *-ie* and other reduced pronouns in Dutch stems from the fact that its phonological form is not compatible with a proclitic position.

[41] Anderson (1992, ch. 8) makes the same claim. He points out that the rules for clitic placement and for affix placement are essentially the same, the difference in his account being that clitics attach to phrases while affixes attach to words.

[42] The prefixes in separable particle verbs appear to be exceptions, e.g., *aan-komen* 'on-come' (i.e., *arrive*), but these are distinctly different elements. Most importantly, they have a different prosodic structure, as reflected by the fact that they carry word stress.

3.3.4 Affix and clitic types

The discussion in the previous sections suggests that we can distinguish between two types of clitics on the basis of their properties. Latin *-que* has a prosodic alignment requirement, while *'ll* does not. These two types correspond to Anderson's (1992) *special* and *simple* clitics, respectively. For Anderson, simple clitics are elements that behave normally from a syntactic perspective. The only property that sets them apart from words is their phonology: they are prosodically deficient. Special clitics are clitics that have a "special syntax", in the sense that their placement is not determined by normal syntactic rules. Both clitic types have subcategories, based on the direction of alignment and the prosodic constituent to which an element must be aligned, and on the syntactic category that is selected, but simple and special clitics are the two basic categories.

The claim that I make here is that the placement of special clitics can be handled in terms of phonological categories, ideally by alignment to some prosodic constituent.[43] In addition, I claim that affixes function in essentially the same way. The main difference between affixes and clitics is the fact that the former have an additional c-selectional restriction, which makes them subject to Input Correspondence.

	syntactic selection	prosodic alignment	PWd status	further restrictions
-ing	✓	✓	–	
-um-	✓	✓	–	bisyllabic
-que	–	✓	–	
'll	✓	–	–	
'k	–	–	–	
-ie	–	–	–	not proclitic
het	✓	–	–	
MAN	–	–	✓	
HIT	✓	–	✓	

Table 3.1: Affix and clitic types

Table 3.1 lists the various affix and clitics types discussed here. For completeness' sake, it also lists the lexical items MAN and HIT. The table shows quite clearly that affixes and clitics are in fact very similar elements. They are not categorically distinct, but rather share fundamental properties. This is a claim that

[43] This appears to be true in at least a majority of cases. There are some types of special clitics, such as Serbo-Croatian second place clitics (cf. Schütze 1994), for which this claim is problematic, because they are organised in a template. I return to such cases in section 4.6.4 of chapter 4.

Anderson (1992) also makes, and it is confirmed by Booij's (1996) observation mentioned above that clitics and affixes are phonologically identical. The main difference, based on the table in table 3.1, is that affixes are subject to both syntactic selection and prosodic alignment, while clitics are at most subject to one of these.

Also listed in the table are additional restrictions that apply to individual elements. The Tagalog affix *-um-* is bisyllabic, as discussed, which causes it to appear as an infix in certain circumstances, and Dutch *-ie* cannot be proclitic, due to its phonological form and general restrictions on proclitics in Dutch. The point is that even though syntactic selection, prosodic alignment and PWd status are the main features that define affixes and clitics, individual elements may be subject to additional restrictions, giving rise to more specific behaviour.

Klavans (1985) three clitic parameters cannot be carried over directly into the current model. Klavans's first parameter defines the domain of a clitic as a syntactic unit. In the current model, clitics do have what I call an *association domain*, but it is defined prosodically. For example, I argued that the association domain for Latin *-que* is the phonological phrase corresponding to its second conjunct. In general, however, the association domain of a clitic is the prosodic domain corresponding to the syntactic phrase that contains the clitic, which should not create serious discrepancies compared to the domain defined by Klavans's theory.

The second and third parameters in (55) together define the location of a clitic. The second parameter defines whether the clitic is placed with respect to the first or last element in the domain.[44] This parameter has no direct equivalent in the current system. For a clitic such as Latin *-que*, its effect is obtained by Left-to-Right Association, which suggests that left-association (i.e., in Klavans's terms, association with the first element of the domain) is not really a matter of some lexical parameter of the clitic. Association with the right edge of the domain cannot be derived in a similar way. Left-to-Right Association is a general principle of autosegmental phonology, but there is no equivalent Right-to-Left Association. Right-association can be a language-specific principle (e.g., the Arabic rule that all stems end in an extrametrical syllable, which results in the association of the last root consonant with the right edge of the PWd), or a lexical property of the relevant element. There is also a third option: the clitic could be placed by syntax, i.e., it could be a simple clitic in Anderson's terms, that ends up at the right edge of its containing phrase for syntactic reasons. Looking at Klavans's examples suggests that this is even the prevalent option.

[44] In Anderson's (1992) overview, they can also be placed with respect to its head, an option that Klavans (1985) does not include and I ignore, as mentioned in footnote 34.

Klavans's third parameter, the placement with respect to the element selected by the second parameter, has a more direct equivalent, In the current model, a clitic can have an alignment property with respect to some prosodic category, usually its host. However, not every clitic has such a requirement: some clitics are subject to Stray Adjunction or to language-specific prosodic requirements. Furthermore, the alignment property that some clitics have is the same that aligns affixes with respect to their hosts and is therefore not a property exclusive to clitics.

Summarising, the placement of clitics is handled rather differently in the current proposal. Following Anderson (1992), I assume that simple clitics do not need any placement parameters, their placement being determined by the syntactic structure and Linear Correspondence. Special clitics do have a placement parameter, implemented as a prosodic alignment requirement, but such a requirement is not unique to clitics. The term 'clitic' is therefore not a primitive notion of any linguistic theory. It is a descriptive term that does not pick out a single category of elements. In this sense, it is not very different from the term 'word', which is also imprecise and not suited as a linguistic term.

3.4 Concluding remarks

The goal of the current chapter has been to show how the idea that part of what we traditionally call "morphosyntax" takes place in phonology can be implemented. The chapter title refers to prosody-syntax interaction, but it should be clear that what is meant is meta-level interaction. Syntax and prosody both contribute to construct linguistic forms. Specifically, syntactic structure alone is not sufficient to determine linear order.

The basic idea itself is not new: it has been developed and employed to model non-concatenative morphology in many languages. The current proposal extends this idea, however, by arguing that the same principles underlie concatenative morphology and even syntax. More specifically, the proposal is based on the assumption that there is no clear distinction between syntax and morphology (as discussed in the next chapter). Note, however, that this does not entail that every aspect of syntax is governed by prosodic principles. At lower levels of the prosodic hierarchy, the effects of prosody are stronger, while at higher levels, other principles start playing a role as well.

The examples in this chapter focus on phenomena at lower levels of the prosodic hierarchy (head movement and the nature of clitics), but the sign language data show that the principles behind these examples apply to syntax as well, because syntactic heads can have phonological components that are associated with prosodic words, phonological phrases and even larger prosodic constituents.

4
Syntax vs. morphology

4.1 Introduction[†]

The claim of the present proposal is that we need to extend the principles behind prosodic morphology to syntax, yet many of the examples discussed in the previous chapters traditionally fall within the realm of morphology. They have been included in the discussion because within the model developed here, they can be analysed as syntactic structures that only start to look like morphological ones when they are mapped onto phonology. It is an interesting additional effect of the model of prosodic syntax that it offers the opportunity to unify morphology and syntax into a single morphosyntax module.

The distinction between syntax and morphology is well established in both descriptive and theoretical linguistic thinking. Even though there are various proposals in the linguistic literature that argue for eliminating the distinction (e.g., Ackema 1995; Halle and Marantz 1993; Lieber 1992; Selkirk 1982; Siebert 1999, etc.), none of these can really escape the distinction between word structure and phrasal structure and each must make special arrangements in its syntactic model to accommodate word structure and the properties that set it apart from phrasal structure.

In this chapter, I argue that the idea behind proposals to unify syntax and morphology is essentially correct: there is indeed only a single computational system combining morphosyntactic elements into larger structures. The empirical differences between word structure and phrasal structure that we observe in linguistic forms, and which are the reason why the authors mentioned in the previous paragraph all need to assume special syntactic operations that cater

[†]Parts of this chapter appeared previously as Kremers (2015).

specifically to word structures, should, I believe, be explained through the interaction of syntax with phonology.

Like the authors just cited, the current proposal cannot escape the fact that there are superficial differences between syntax and morphology, and any theory that aims to unify them must accommodate the properties that appear to make morphology different from syntax. Unlike previous theories, however, the current model does *not* need to adapt syntactic theory, e.g., by assuming special syntactic principles for "complex heads". Rather, I will argue that the "morphological" properties of an element follow from the interplay of its syntactic and phonological properties. In other words, it is the combination of particular properties occurring in the lexical entry of an element (which contains both the syntactic and the phonological properties of that element) that cause it to appear to us as morphological.

Put differently, I claim that the distinction between morphology and syntax is essentially an optical —or rather, acoustic— illusion: it is not the structure-building mechanism that determines whether a given structure is morphological or syntactic, rather it is the phonological structure of the form onto which the output of the structure-building computation is mapped that lets us *perceive* a given structure as morphological or syntactic. Correspondence effects between syntax and semantics strengthen this effect. There is a strong tendency in language to express concepts (i.e., basic units of semantic computation) as prosodic words. As a result, it appears as if the grammar treats words differently from phrases, but in fact, the same principles underlie both domains. More precisely, there is only one domain.

One consequence of the proposal is that we must assume that certain generalisations and constraints that we observe in language are imposed by external systems. These include restrictions placed on the grammar by the modality in which it is expressed (e.g., the requirement of linearity), restrictions imposed by the parser (e.g., recoverability of the base position of a displaced element), and generalisations and constraints resulting from the correspondence between semantic and syntactic structure and between syntactic and phonological structure.

While the latter type of restriction can still be considered grammatical, since the mapping between semantics, syntax and phonology is part of grammar in the broad sense, the former type of restriction is not grammatical. Rather, restrictions of this kind follow from what Chomsky has called *interface* or *third-factor effects* (i.e., from "principles of structural architecture and developmental constraints that are not specific to the organ under investigation, and may be organism-independent" Chomsky 2008, p. 133) or the Faculty of Language in the Broad sense (FLB, Hauser, Chomsky, and Fitch 2002). They do not constrain the grammar of an I-language directly, rather they constrain the *development* of an

I-language.

4.2 General aims

Before I go into the discussion, I would like to provide an overview of what it is that I am aiming for, and in particular also what I am *not* aiming for.

In generative approaches to syntax, the main aim is to discover the general principles governing the construction of syntactic structures. A consequence of this aim is that an attempt is made to eliminate any and all irregularities from the theory. Even properties that appear to be irregular on first sight must be accounted for by the interplay of specific rules. This interplay may become so complex that it creates the appearance of irregularity, but in actual fact, we are looking at a complex set of rules.

Whenever it is not possible to reduce some fact to the application and interplay of certain rules, the phenomenon is relegated to the lexicon, or, if there appears to be some idiosyncratic (i.e., language-specific) generalisation at work, to morphology. Phonology is similar to syntax: the general aim of phonologists is to find the general rules of the system. Anything that cannot be described as the result of some general rule is relegated, again, to the lexicon or to morphology. In this way, morphology and the lexicon form the place in the grammar model where idiosyncrasies, both facts and generalisations, are stored.

The fact that there are not just language-specific facts, but also language-specific generalisations serves as motivation for the assumption of a separate morphology module, a computational subsystem that is part of the larger grammar system. Any theory that argues for a unified morphology-syntax module assumes that this separate morphology module does not exist and that morphological structures are syntactic. Every theory to this effect that I am aware of (as cited above) needs to make accommodations in the syntactic component for morphological structures. Ackema (1995), for example, argues that morphological structures should be analysed within X-bar theory as projections of negative X-bar levels, i.e., projections from X^{-2} to X^{0}. Similar proposals are found in Selkirk (1982), who proposes a distinction between X° and X^{Af} elements, and Lieber (1992), who argues that X-bar theory should be extended with a recursive projection rule at the X°-level of the form $X^{\circ} \longrightarrow X^{\circ}$.

The general problem with such proposals is that even though they claim to do away with the distinction between syntax and morphology, they reintroduce the same distinction by adopting special constraints and principles in the syntactic component that only apply to word-level structures. Negative X-bar levels, or a recursive X° level are nothing more than morphological structures in a "syn-

tactic" jacket.[1] Moreover, they constitute a set of operations that are located at a specific point in the derivation, namely below the head level. These operations apply first, before the original syntactic operations, i.e., those operations that raise the X-bar level beyond X°, take place.

Distributed Morphology (Halle and Marantz 1993 and much subsequent work), the leading morphology theory in minimalist frameworks, is no exception, despite the fact that it assumes "syntactic structure all the way down". However, even though DM does not assume a separate morphology module, there are still a number of so-called "morphological operations" such as fission, fusion, impoverishment and merger. Crucially, these also apply at one specific point in the derivation, albeit at a different point than negative X-bar levels. They apply at the end of the syntactic derivation, right before Vocabulary Insertion. Such a model is not very different from a model that constructs a syntactic structure and then feeds this structure into a separate morphological module. In other words, DM suffers from the same problem: the syntax has to be augmented with a special set of operations that deal with word-level structures, and these operations are located at a specific point in the grammar model.

It should be noted that the idea that syntax needs to be augmented with specific operations hat apply only to word-level structures is in itself not immediately problematic. It becomes problematic under (one of) two conditions: (a) they all apply at the same point in the derivation; and (b) they apply to a pre-defined set of elements. The second condition refers to elements that are marked in the lexicon as X^{Af} or $X^{\leq 0}$. Since such elements are marked in the lexicon as being subject to morphological rules, they negate the idea of a unified morphosyntax module.

Having separate rules for words in the syntax module is not problematic, if those rules can apply throughout the derivation, i.e., intertwined with other syntactic operations that apply to phrases, and if those rules apply to elements on the basis of properties that these elements have independently from the need to treat them as "morphological". A rule that applies to elements that have an autosegmental phonological form, such as Input Correspondence, is acceptable, because it does not apply exclusively to "morphological" elements. It also applies to element such as sentential negation or sentential adverbs in sign languages. A rule that only applies to elements that are marked as X^{-1} is not acceptable, because such a marking is arbitrary; it is not imposed by the system but by the linguist analysing the system.

The current proposal does not aim to do away with morphology altogether. It just aims to locate those processes that are typically considered to be "morphological" in the syntax, interwoven with traditionally "syntactic" operations.

[1] Cf. Siebert (1999) for a similar argument.

It also does not aim specifically to reduce all "morphological" operations to syntactic ones. I do not explicitly exclude the possibility that some operations only apply to "morphological" elements, as long as this limitation is triggered by an independent property of the elements in question. Interestingly, however, there do not seem to be any clear cases of such operations. Principles and rules that I propose tend to be applicable to both types of structures, which strengthens the idea that there is no discrete distinction between the them.

There is at least one concession to the syntactic system that must be made if we are to pursue a unification of syntax and morphology. Traditionally, syntactic theory has attempted to exclude all idiosyncrasies from the grammar. The idea has been that syntactic operations must be as general, as universal as possible, otherwise some generalisation is being missed. Morphological operations, on the other hand, do not have such a strong universality requirement. It is perfectly acceptable to propose a morphological rule that only applies to one particular language. In a unified syntax/morphology module, incorporating such language-specific rules into the syntax is inevitable. We cannot avoid the need to formulate rules that are specific to a certain language or certain forms within a language, for the simple reason that such rules are known to exist.

The situation is not as bad as these remarks suggest, however. Morphosyntactic rules are sensitive to morphosyntactic features in standard ways: selection and agreement. The only concession that must be made is that morphosyntactic features can be language-specific. This, however, is not only a descriptive necessity but also a conceptual inevitability, if the assumption that features cannot be innate is correct (cf. the discussion in chapter 1). Apart from this, no novel syntactic mechanisms will be proposed in the current chapter. As it turns out, selection and agreement, together with features, are the only morphosyntactic building blocks that we need.

4.3 Syntax-morphology synthesis

Lieber (1992, p. 21) states that:

> [t]he conceptually simplest possible theory would (…) be one in which all morphology is done as a part of a theory of syntax (…) A truly simple theory of morphology would be one in which nothing at all needed to be added to the theory of syntax in order to account for the construction of words.

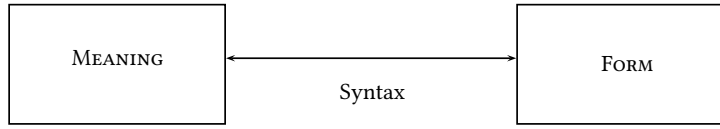
It is certainly true that the most elegant theory possible would not distinguish between a syntax module and a morphology module. There are, in fact,

several other conceptual reasons for preferring a theory with a unified syntax-morphology module. Even though none of these arguments is really fatal to a two-module theory, they do show that if we are able to formulate a unified theory that is not less elegant than a two-module theory, we should prefer the unified theory.

4.3.1 Form-meaning pairing

The idea that language is form with a meaning goes back to Aristotle. The individual signs (the forms) have a particular meaning, but when signs are combined, it is the syntactic structure that links the form of such a combination to its meaning. Syntax, in this view of language, is the glue that mediates the pairing of form (phonology) and meaning (semantics). We can visualise this situation as in (1):

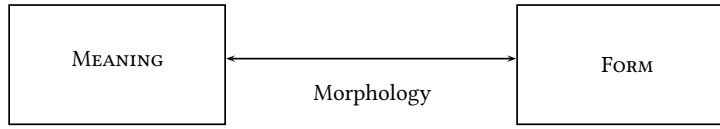
(1)



Interestingly, we can repeat the previous paragraph, replacing the word "syntax" with "morphology" and still maintain a true statement:

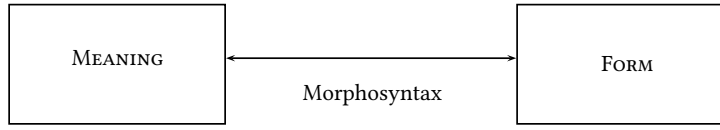
The idea that language is form with a meaning goes back to Aristotle. The individual signs (the forms) have a particular meaning, but when signs are combined, it is the morphological structure that links the form of such a combination to its meaning. Morphology, in this view of language, is the glue that mediates the pairing of form (phonology) and meaning (semantics). We can visualise this situation as in (2):

(2)



What this means is that from a bird's eye view, syntax and morphology perform the same function: they link form to meaning. From such a perspective, it seems strange to even assume there should be two different modules. It would make much more sense to have just one module responsible for both word structure and phrasal structure:

(3)



I will refer to the single morphology/syntax module as "morphosyntax" or simply "syntax". The obvious bias toward syntax in this nomenclature should not be taken as an indication that syntax is more important than morphology. It is simply the case that given the conventional connotations of these words, "syntax" more easily generalises to a hyperonym than "morphology".

4.3.2 Distinguishing structures

In traditional terms, the difference between morphology and syntax lies in the kinds of structures that they deal with: morphology operates below the word level, syntax operates from the word level upwards. The problem with such a definition of the distinction is that the term "word" is not defined (see also Haspelmath 2011). Syntax deals with heads, not with words, but we cannot say that morphology "operates below the head", because the standard view is that morphology operates on heads as well. Many of the heads that syntax deals with are realised overtly as affixes or merely as features of some word form (e.g. Infl or T/Asp in Germanic, Romance, Slavic, etc., C in languages with clitic complementisers, D in languages with clitic determiners, K (or Case) in languages with overt case morphology, etc.)

Nonetheless, as discussed above, it has generally proven impossible to formulate a theory of a unified morphosyntax that does not sneak the distinction between word-level and phrase-level structures back into the theory in some way or other. For Ackema and Neeleman (2004, 2007), this fact is an important reason for assuming that there *is* a separate morphology module. Such an approach, however, has to answer the question which structures are generated by the morphology module and which by the syntax module. Intuitively, it seems that it is not difficult to distinguish between words and phrases in most cases, but as Haspelmath (2011) shows, there is no general, cross-linguistic property that distinguishes words from phrases and our intuitions in this regard are all too often shaped by our writing system.

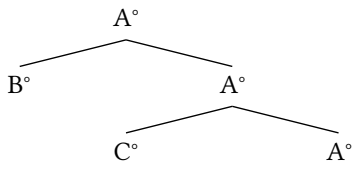
This is not to say that there have been no attempts at defining the notion "word" in the literature. One such attempt is made by Embick and Noyer (2001), working within the framework of Distributed Morphology. Embick and Noyer's proposal captures what many would probably consider the most straightforward and intuitive way of defining the word. Specifically, Embick and Noyer define two notions, the *Morphological Word* and the *Subword*:

(4) a. *Morphological Word*:
At the input to Morphology, a node X° is (by definition) a morphosyntactic word (MWd) iff X° is the highest segment of an X° not contained in another X°.

b. *Subword*:
A node X° is a subword (SWd) if X° is a terminal node and not an MWd.

The problem with this definition is that it presupposes that we are able to tell when a specific node in the tree is "the highest segment of an X° not contained in another X°". This, however, is only possible by stipulation. The kind of structures that Embick and Noyer have in mind is represented in (5):

(5)



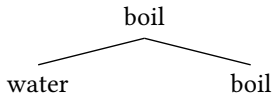
Here, the idea is that the top A° node corresponds to a word. However, in a framework that adopts Beard's (1988) *Separation Hypothesis*, which DM does in the form of *Late Insertion*, there is no way to determine on the basis of the syntactic tree which nodes are X° nodes and at which point an X° projects into an X′. The syntactic tree contains just the heads A, B and C. From the structure alone it is not clear whether C is the complement of A or whether it is an adjoined head.

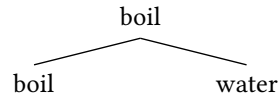
In actual fact, there *is* a way to tell which node is an X° node and which node is a projection X′. Chomsky's (1995) bare phrase structure proposal allows us to do this. However, from Embick and Noyer's perspective, it yields the wrong results for structures such as (5).[2] In bare phrase structure (cf. also section 2.5 in chapter 2), an X° category is a minimal projection, i.e., a piece of structure that is not projected from another element (in other words, not composed by Merge). Similarly, a maximal projection is the highest node in a projection line. For the structure in (5), this means that B and C are indeed X° categories, and so is the *lowest* A node. But the other two A nodes are non-minimal, i.e., they are *not* X° categories. The highest A node could in fact be a maximal projection, if it does not project any further. In other words, according to Embick and Noyer's definition, only the bottom A node is a morphological word, which is of course not what they want to say. Furthermore, the notion of "subword" becomes a contradiction, because every X° is a MWd.[3]

To make the problem a bit more explicit, consider the following two structures:

[2]Embick and Noyer (2001) explicitly adopt bare phrase structure, but they seem not to have noticed the incongruity between their proposal and Chomsky's.

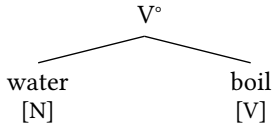
[3]Note that this technical argument applies only to systems that assume no separation between syntax and morphology, such as DM. However, the point that it is not possible to reliably distinguish between words and phrases is independent of the theoretical framework one chooses, as Haspelmath (2011) shows. That is, lexicalist theories and theories such as that of Ackema and Neeleman

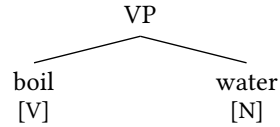
(6) a.


b.


(6a) is the base for the word form *water boiler*, while (6b) is the start of a derivation that leads to a verb phrase *(to) boil water*. The only difference seems to be the order in which both words appear, but this is not a property we would like to rely on, obviously: it will hardly do to argue that morphology branches to the left while syntax branches to the right. Besides, it is commonly assumed that syntactic structures do not define linear order.

The point here is that when two heads such as *boil* and *water* are merged, the structure is still indeterminate: it could be the start of a derivation yielding *water boiler*, but it could also yield *to boil water*. A possible response to this problem might be to argue that the trees should properly be represented as in (7a) and (7b), although it should by now be obvious that this does not bring us anything:

(7) a.


b.


Such a representation simply begs the question. The zero-level projection marker and the "VP"-label are really just mnemonic devices reminding us of the kind of structure that we are dealing with: a "morphological" and a "syntactic" structure, respectively. One can of course raise such mnemonics to the status of theoretical devices, but that does not change the fact that they are there to make a distinction that does not follow directly from the properties of the elements involved: the fact that we combine an N and a V head does not tell us whether we are dealing with morphology or with syntax.

This is a point that is well worth dwelling on for a second: our model of grammar is based on the intuition that there is a distinction between the word-level and the phrasal level. This notion pervades our thinking about language, it guides our analyses every step of the way. It is reflected in our descriptive machinery, which we use without realising that something like "V°" or "VP" is nothing more than a mere description of what we observe, not an explanation.

(2004) face the same problem, but it turns up as the problem of deciding which structures are built by which module.

The existence of languages with elaborate and highly regular morphology emphasises this point, as Jackendoff (2002) also points out. Take the Classical Nahuatl word form *ahōniccuah*, which contains the following morphs:

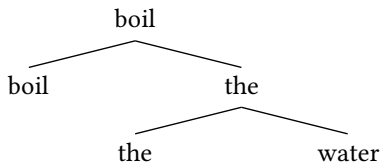
(8) *Nahuatl*
ah- ō- ni- c- cua- h
NOT ANT 1sg.S 3sg.O eat PERF
'I have not eaten it.'

The single Nahuatl 'word' *ahōniccuah* corresponds to a five-'word' sentence in English. There is no sensible way for us to claim that the Nahuatl word is constructed through mechanisms fundamentally different from the English sentence. Semantically, the two utterances must express the same structure, and it would be unlikely if the language faculty were to contain two fundamentally different mechanisms to generate functionally similar structures. At the very least we should proceed from the hypothesis that it does not.

Ackema's (1995) proposal and the others mentioned above are compatible with this point, in that they argue that there is no distinction between syntactic and morphological operations. The problem remains, however, that they cannot explain why the Nahuatl 'word', while built by the same types of operations, still looks different from the English sentence. Their explanation is that there are simply different kinds of elements to which the operations apply: $X^{<0}$ and $X^{\geq 0}$. But this is merely begging the question, a notational device introduced to describe the difference. If we drop the notational device, we have no way of deriving the empirical differences.

The only difference that seems empirically true is that the syntactic tree but not the morphological tree can be expanded with functional elements. A functional element such as *the* may not appear in a morphological tree. That is, in syntax, but not in morphology, we may have the following:

(9)



That is to say, *in English*. There are, however, many languages with affixal determiners. Similarly, can we really say that the Nahuatl example in (8) does *not* contain functional structure? But if it does, how are we to make sense of this fact? The usual assumption, that the affixes are either a morphological reflex of syntactic features on the lexical head, or that the lexical head moves to the

functional heads to "pick up" the affixes, is inspired by our preconception that morphology and syntax are different. It follows from the assumption, it does not prove it.

4.3.3 Making sense of the intuition

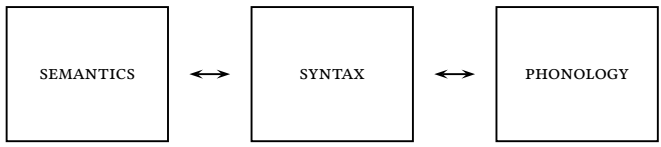
I do not wish to claim that the entire distinction between syntax and morphology is an illusion. There *is* a distinction, it is just not there where we suspect it. Our intuition that there is a difference between the level of the "word" (however vague that notion is) and the level of the clause is in some way correct, and our model of grammar should reflect it. It should just reflect it in a different manner than we are used to.

The notion of lexical item discussed in chapter 2 offers a starting point for an analysis that captures the intuition in a different way. First, consider again the representation of a lexical item in example (8) of chapter 2, repeated here as (10):

(10) $\lambda x(\mathbf{car}(x)) \leftrightarrow \begin{bmatrix} \text{N, sg} \\ \text{count} \end{bmatrix} \leftrightarrow /\text{kaɪ}/$

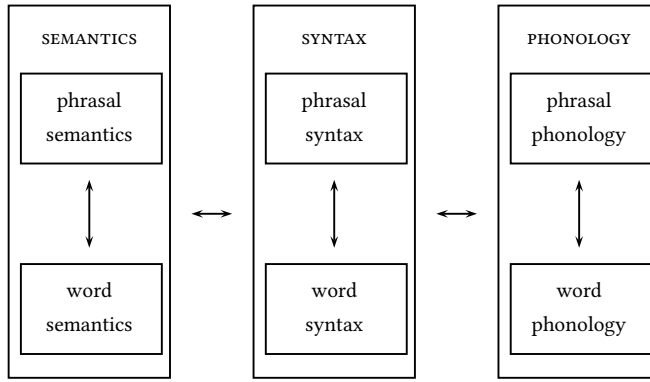
Important in this representation is that the three groups of features are separate, which is why they are visually separated here. We can in fact generalise this notion to the entire grammar:

(11)



Just like the representation for lexical items, this grammar model is inspired by Jackendoff (1997, 2002). What is relevant at this point is that the distinction between morphology and syntax is usually placed in the second box, the one that builds structures. Ackema and Neeleman (2004, p. 135) generalise this and assume that each module has two submodules, one for phrasal structures and one for word structures:

(12)



But why? Why do we need to distinguish a word level box and a phrasal level box in each of the macromodules? There is no reason why the semantics box would need it. Whether we use a word such as *ahōniccuah* or a phrase such as *I have not eaten it*, or whether we say in English *unbearable* or in Arabic *lā yuḥmal* (lit. 'it is not borne'), the semantics is the same. In other words, the semantics box can do without the two submodules.

However, given that we do not want to assume that the structure-building operations behind *I have not eaten it* are fundamentally different from the structure-building operations behind *ahōniccuah*, we should ask ourselves why we would want to assume two submodules in the SYNTAX box. Or, to phrase that question a bit more generally: what is the *actual* basis for making such a distinction? Why do we say that *boil water* is fundamentally different from *water boiler*?

The directly observable differences between syntactic and morphological structures are phonological in nature. That a *hót-dŏg* is different from a *hót dóg* is reflected in the stress pattern. That *boil water* differs from *water boiler* is obvious from the word order,[4] the phonological element /ɚ/ and perhaps also the stress pattern. We say that *-er* is a suffix and thus belongs to morphology because of its requirement to attach to a stem. But this requirement is really a *phonological* requirement: it is the syllable /ɚ/ that must attach to a prosodic word in order to be phonologically licit.

We have no direct access to the SYNTAX box in (12). Our assumption that it is divided into a word level and a phrase level box is based solely on observations made through the phonological system of language, i.e. on our observations of the phonology box. What is more, the distinction is very old and dates from a time when the structure of the phonology box was not very well understood.

[4]Which, in current minimalist thinking, is a PF phenomenon, cf. Chomsky (1995b), Nunes (2004), Richards (2004) Fox and Pesetsky (2005), Kremers (2009a), etc.

Nowadays, however, we know that the prosodic structure of an utterance is essential for shaping its phonology. It is generally recognised that there is a *prosodic hierarchy* in phonology (cf. Selkirk 1981, 1984, 1995, Nespor and Vogel 1986, Truckenbrodt 1995). This hierarchy has the Utterance at its top, and the syllable at its bottom:[5]

(13) Utterance (U)
Intonational phrase (IntP)
Phonological phrase (φ)
Prosodic word (ω)
Foot (Ft)
Syllable (σ)

The prosodic hierarchy is subject to the *Strict Layer Hypothesis* (SLH), which essentially states that the hierarchy is exhaustive and non-recursive. Each element at level *n* is completely contained in an element at level *n+1*, while containing one or more elements of level *n−1*. That is, an utterance contains one or more intonational phrases, and each IntP is contained in exactly one U (which it may share with other IntPs). In turn, each IntP contains one or more phonological phrases, etc. Furthermore, an element at level *n* cannot contain another element of level *n*. A prosodic word may not contain another prosodic word, a φ cannot contain another φ, etc.[6]

As discussed in chapter 2, there are rules governing the mapping of syntactic structure onto prosodic structure, for example the rule that a (lexical) XP in syntax corresponds to a p-phrase. These correspondence rules are not absolute, however. That is, there are cases where the syntactic structure and the phonological structure do not line up. This is in part due to conflicting correspondence rules, in part to the fact that syntactic structure is recursive, while prosodic structure is not,[7] and in part to the fact that after the correspondence rules establish an initial prosodic structure, readjustment rules can rearrange the boundaries of prosodic constituents.[8]

[5]The levels and the symbols referring to them are essentially the ones Nespor and Vogel (1986) introduced. I do not go into the heavily discussed question whether these are the correct levels, or whether we need more / fewer, etc.

[6]Here, I have factored out the SLH into two constraints. Sometimes it is argued that the SLH actually comprises more than two constraints. For example, Selkirk (1995) assumes four. Note that in OT approaches, the principles are assumed to be violable. These issues are not relevant to the point at hand, so I ignore them here.

[7]But see Ladd (1986) and others for a different proposal.

[8]Note that Scheer (2008) argues that if we were do to away with prosodic domains and would instead refer only to prosodic boundaries, the mismatches would disappear, because even though not every syntactic phrase may correspond to a prosodic domain, each prosodic boundary corresponds to the edge of a syntactic phrase. Although I believe Scheer is essentially correct, the

Phrase-level syntax maps primarily onto the φ-level in prosodic structure, to a lesser extent also onto the level of IntP. Morphology maps onto the levels of the prosodic word, the foot and the syllable. *Well, obviously*, one is tempted to say, because morphology operates below the word level. But this is a matter of perspective. If we look at it the other way around, things change: then we suddenly see that it is not the structure-building component (either syntax or morphology) that determines which prosodic levels its output is mapped onto. Rather, it is the prosodic level on which the effects of a certain operation are visible that directs —or rather *squints*— our view of this operation as syntactic or morphological.

Consider this for a moment. *Ahōniccuah* is a single prosodic word and is therefore considered a morphological structure. *I have not eaten it* contains (at least) two prosodic words and is therefore considered syntactic. That is, there is nothing inherent to the structure-building operations behind *ahōniccuah* that force it to come out as a "morphological" structure. It is the phonological form that the resulting morphosyntactic structure maps onto that prompts us to view it as a "morphological" structure.

Adopting this view leads us to an interesting conclusion. One crucial property of the prosodic hierarchy in (13) is that it is continuous: there is no break between any two levels. That is, if it is indeed the phonology that determines our view of the structure-building component, i.e., of the SYNTAX box in (12), then we have no basis for distinguishing between word-level syntax and phrasal syntax ("morphology" and "syntax" in traditional terminology, respectively). There should be only *one* structure-building component, and its operations should be oblivious of the kinds of phonological structures associated with the symbols to which they apply.

This is, of course, exactly what the prosodic syntax model also predicts. Syntax, the structure-building component, manipulates heads, bundles of morphosyntactic features that have no phonological form whatsoever. Unless we add a diacritic to heads that are "morphological" rather than "syntactic", we cannot determine whether a structure $[_{A}$ A B] is morphological or syntactic. Only when we look at the mapping to phonology, i.e., at the phonological component of $[_{A}$ A B], can we determine this. But at this point, it is not relevant anymore, at least for the syntax module.

prosodic syntax model should work equally well in his model, so I will stick to the more familiar prosodic hierarchy model.

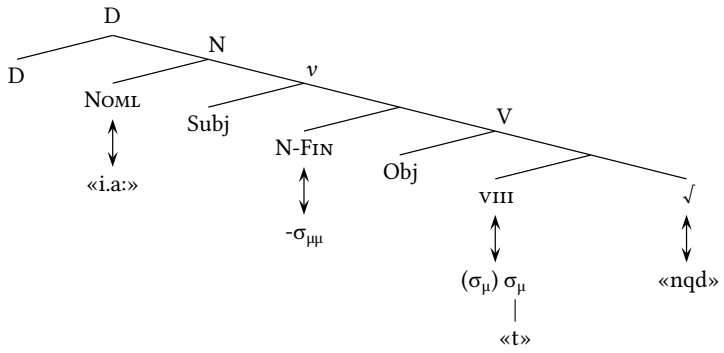
4.4 Data

The preceding chapters have basically presupposed the view sketched here. Let us return briefly to the data discussed there and see how exactly they support the view that syntax and morphology are not strictly separate but part of a continuous system. Let us first look at the derivation of the Arabic verbal nouns again, discussed in section 2.7.1. The basic structure is repeated in (14):

(14) ʔaqlaqa-nī -ntiqād-u -l-rajul-i li -l-mašrūʕ-i
annoy.3sg.m-1sg.OBJ criticising-NOM DEF-man-GEN to DEF-project-GEN
'The man's criticising of the project annoyed me.'

The tree that I proposed is (15):

(15)



As discussed, the word form *intiqād* is not created in a separate morphology module before being inserted in the syntactic tree. Rather, the different morphemes making up the word form *intiqād* are all part of the syntactic tree as independent elements.

However, unlike existing approaches that try to subsume morphology and syntax into a single module, the morphemes are not combined into a single word form by the syntactic module, either. Rather, they are combined by the phonological module. The details of the analysis are discussed in chapter 2, they do not need to be repeated here.

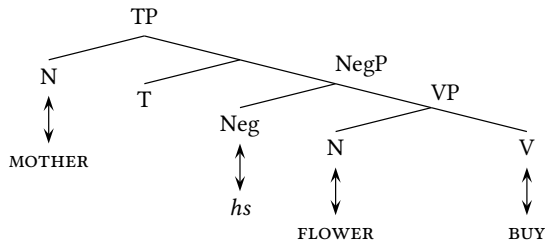
Next, consider the analysis negation in German Sign Language (DGS) in section 3.1.1. One relevant example phrase is repeated in (16):

(16) **head:** *neg*
hands: MOTHER FLOWER BUY
'Mother does not buy a flower'

The relevant element in these clauses is the negation, which is realised as a head shake accompanying the verb. The most straightforward analysis appears to be to treat the head shake as a morphological negation marker, which is exactly what Pfau and Quer (2002) do.[9] But this analysis is not necessary. The phonological form of the negation, Φ(Neg), is an autosegment consisting of a head shake. As such, it needs to be associated with some segmental material.

In other words, we can simply adopt an analysis in which the negation is a head in the syntactic structure occupying the normal position for Neg heads, along the lines of (17):

(17)



This tree accommodates the negation in the standard manner, as a sister of VP. Head movement of V or lowering of Neg is not necessary. Rather, phonology takes care of the realisation of Neg. The syntactic structure does not need to instruct the phonology where or how to realise it. It follows automatically from the fact that Neg selects V, which means that Φ(Neg) combines with Φ(V), and from the fact that the negation is non-segmental, which means that it must be associated with segmental material. Put the two together and the only structure that phonology can generate is (17).

In fact, negation in DGS is even more problematic for morphology than (16) suggests. The head shake can spread over the entire VP, even though this is less common than realisation with just V:

(18) **head:** *neg*
hands: MOTHER FLOWER BUY
'Mother does not buy a flower'
(Pfau and Quer 2002)

[9] See also Pfau (2002, 2008) for this view. Note, however, that such an analysis still faces the question why negation is sentential.

A morphological analysis is not able to handle these data, since the head shake is no longer associated with a single word. Data such as (18) therefore require a syntactic analysis, but once such an analysis exists, it is no longer satisfactory to have two different analyses for what is essentially the same phenomenon.

This problem is easily solved, of course, if we let phonology take care of the realisation of the head shake. There is data suggesting that the spreading domain is prosodic rather than (morpho)syntactic, as discussed in section 3.1.1 (see there for details). Therefore, we can account for the data by assuming that phonology can choose both the prosodic word and the phonological phrase as the domain to align the head shake with. In either case, the verb must be contained in the domain, so that the head shake is either aligned with the verb alone or with the phonological phrase that contains the verb.

The last case that I discuss here involves clitics, which have been discussed in section 3.3. The basic idea is that since clitics (appear to) stand somewhere between morphology and syntax, an analysis that does not assume a strict separation between the two modules stands a better chance of analysing them properly. Here I will limit myself to the Latin second-place clitic *-que* 'and', which illustrates the problem and the proposed solution best.

The main problem with *-que* is the fact that it can attach both to words and to phrases, depending on the prosodic structure of the relevant element. The relevant data are repeated here:

(19) a. circum-que ea loca
around-and those places
'and around those places'

b. contrā-que lēgem
against-and law
'and against the law'

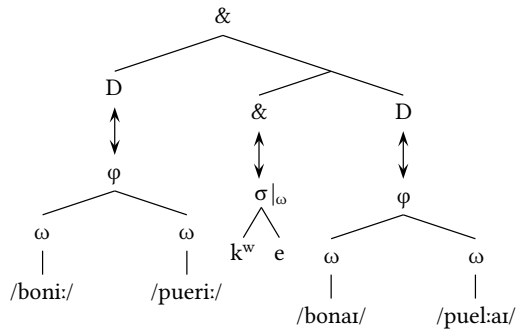
(20) a. in rēbus-que
in things-and
'and in things'

b. dē prōvinciā-que
from province-and
'and from the province'

As discussed in section 3.3.3, *-que* attaches to prepositions that are multisyllabic, as in (20a). If the preposition is monosyllabic, however, *-que* attaches to the word following the preposition. Here, too, the solution is to let phonology take care of the placement of *-que* and assume a syntactic structure with the &° head in its normal structural position:[10]

[10] Which I have assumed is as a sister to the second conjunct, even though it is not at all clear whether that is the best analysis. For our current purposes, this question is not relevant, however.

(21)

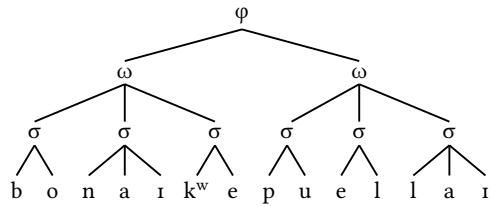


Recall from section 3.3.3 that I argued that the placement of *-que* is determined by the fact that it has a prosodic requirement in its lexical entry:

(22) &° ↔ σ|ω
kʷ e

This requirement forces phonology to realise Φ(*-que*) as an autosegment requiring association with some segmental material. The association is established by attaching Φ(*-que*) to the first prosodic word that follows it:

(23)



As with the previous examples, we see that by placing the process that builds the word form in the phonology, we are able to describe the structure and formation of conjunctions with *-que* without having to refer to a morphology module. We only need to assume a syntactic tree along the lines of (21) and a lexical entry for the conjunction along the lines of (22), and then the phonology does the rest.

However, my claim is in fact stronger than this: not only is it not necessary to refer to a morphology module, it is in fact a mistake to do so. One should not try to account for the form *bonaeque* by bringing the N° and the &° heads together in the hierarchical structure, after which the phonology is a relatively simple and "dumb" process that takes the tree and just plugs in the segments. After the

hierarchical structure is built, but before phonology processes the structure, it is not clear at all in what order the elements in the tree are going to end up. The phonology takes an active part in determining the final linear order.

This does not mean that the process of phonological composition is not deterministic. Given the same input structure, we expect the same output. It is just that the syntactic tree does not contain enough information to determine the output of phonology. The phonological forms of the elements in the syntactic tree influence the output of phonology as well and this information is (by definition) not part of syntax.

The area of grammar where this effect is felt most is no doubt morphology (in its descriptive sense), because constituent parts of words, i.e., affixes, are much more likely to determine their own alignment properties than constituent parts of phrases, i.e., words.[11] This is one of the factors that creates the impression of a separate module for word structure, but the impression is an acoustic illusion, in essence an epiphenomenon created by the interaction of several factors.

4.5 Counterarguments

As mentioned in the introduction, there have been many attempts to unify syntax and morphology into a single system. There are equally many criticisms of such proposals, which usually point out that there are differences between syntax and morphology that cannot be explained in a unified theory. In this section, I discuss several points that have been brought up in this discussion and show why I do not believe they pose a problem for the assumption that morphology is not a separate grammar module. The phenomena that I discuss are the difference between syntactic and morphological word formation, the lack of morphological movement and lexical integrity. A further argument that is often advanced in favour of a separate morphology module, the existence of autonomous morphological phenomena or, as Maiden (2004) calls it, *Morphology by itself*, is a much broader topic and is therefore discussed separately, in section 4.6 below.

The arguments that I present in this section and the next are, by necessity, plausibility arguments. The gist of each of the arguments is that although a morphology module *could* be a way of explaining the phenomenon at hand, it is not the only possible account, and in some cases it is even a suboptimal account, because it makes incorrect predictions. My aim is to show that an alternative, non-morphological account is possible in principle, either by arguing about the phenomenon in general, or by discussing one or two typical cases and show what the non-morphological analysis would look like in the current proposal. With

[11] Although such cases do exist, cf. *warm enough* vs. *how/too/so warm*, cf. Jackendoff 2002

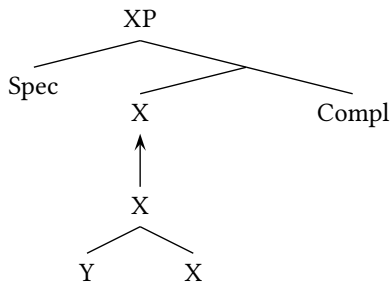
this, the burden of proof comes to lie with those arguing in favour of a morphology module, because a theory that includes such a module is by definition more complex than a theory that does not.

4.5.1 Syntactic vs. morphological word formation

Ackema and Neeleman (2007) present a number of arguments supporting their claim that it is empirically necessary to distinguish between a syntactic and morphological module. These arguments, however, do not apply to the current proposal, because they argue against a type of syntactic word formation that the current proposal does not assume. Since Ackema and Neeleman's concept of syntactic word formation is representative of the kind of theories that are presented in the literature, I will say a few words about it here.

In a morphological theory of word formation, according to Ackema and Neeleman, a complex word of the form $[_X$ X-Y] has the following structure:

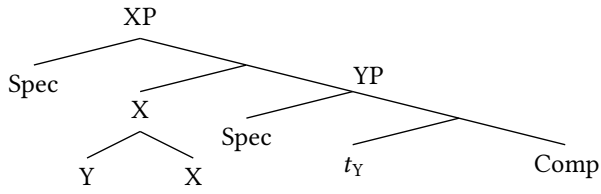
(24)



That is, the complex word itself is formed in the morphological module and then inserted into the syntactic structure (as indicated by the arrow), where it then projects a phrase.

Ackema and Neeleman contrast this structure with the following, in which word formation takes place through head movement in syntax:

(25)



In this structure, X takes YP as its complement and Y° subsequently moves to X°, forming a complex head. This head movement analysis, Ackema and Neeleman claim, makes several predictions that are not borne out by the linguistic

data. First, we would expect that the movement of Y can strand elements in YP. However, it is obvious that head movement that feeds word formation cannot do so:

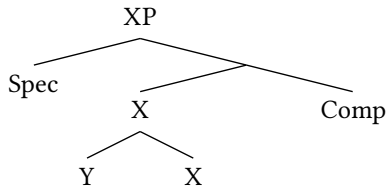
(26) a. *the [city$_i$ centre] [of [a prosperous medieval [t$_i$ [in Northern Italy]]]]
b. *[parent$_i$ hood] [(of) [a [responsible [t$_i$ [from Glasgow]]]]]

Also, if word formation is syntactic, we would expect that verbal nouns such as *driver* would still retain the licensing properties of the underlying verb. However, this is not the case, as witnessed by the fact that *driver* cannot take an accusative complement. Instead, it needs the preposition *of* to license a complement:[12]

(27) driver *(of) a truck

These data are not immediately problematic for the current proposal, however. What these data show is that a theory of word formation that employs syntactic movement is problematic. However, in a bare-phrase-structure approach, there is no need to assume movement. Current minimalist thinking has abandoned the X′-schema and adopted bare phrase structure, making it possible to merge two heads in syntax directly. That is, we can have a syntactic variant of the structure in (24):

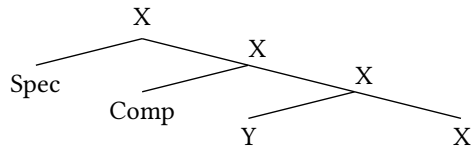
(28)



In this structure, the complex head [$_X$ Y-X] is simply formed in syntax, after which it projects in the normal manner. Note that BPS also does not make a fundamental distinction between complements and specifiers, and bar levels are merely descriptive. Therefore, the structure in (28) can also be represented as in (29):

[12] Obviously, verbal nouns often do retain the licensing capabilities of the underlying verb: gerunds can assign accusative case and allow adverbials. This means that we must argue that *-er* attaches low in the structure. See the discussion of *-ing* in section 3.3.2 in chapter 3 for a similar example.

(29)



This is in fact the same kind of structure that I proposed in section 2.7.1 for Arabic verbal nouns, exemplified in example (21) of chapter 2.

In bare phrase structure, complex heads can be formed in syntax without having to assume movement. The predictions that Ackema and Neeleman argue follow from a syntactic theory of word formation do not apply to a structure of this type. If we merge two heads directly, one of these does not project and we do not expect the non-projecting head to be able to license any dependents, let alone to be able to strand them.

4.5.2 Lack of morphological movement

One typical argument that is often heard (at least informally) in defence of a separate morphology module is the fact that there is no movement in morphology. This is hardly a convincing argument, however. In order to evaluate it properly, we first need to establish what kind of phenomenon would count as morphological movement, and this kind of movement must be theory-independent. In minimalism, for example, movement is a fundamental analytic tool and is therefore employed for theory-internal reasons in many instances. As a contrast, consider non-transformational frameworks such as HPSG, which explicitly do not employ movement as an analytic tool. In such frameworks, there are certain phenomena, generally called *long-distance dependencies*, such as *wh*-fronting, topicalisation, V2, etc., that are analysed by means of gaps, symbolised with so-called *slash* features.

This is the type of phenomenon that we should be looking for in order to find morphological movement: an element or category of elements that can appear in two different positions with respect to the other elements in the word form. There are two reasons, however, why such a phenomenon is unlikely to occur inside words. First, long-distance dependencies generally involve content words, not functional elements.[13] Word-internally, the equivalent would be movement of the lexical root of a word. Since most words contain only one lexical root, it is difficult to determine whether movement takes place: movement in syntax pied-pipes the morphemes attached to the moved root. If the same thing happens in

[13]Of course, in minimalism, movement of functional elements is possible and in fact quite ubiquitous, in the form of head movement, but this is exactly the kind of theory-internal movement that is irrelevant for the discussion.

morphology, movement cannot be detected.

This problem does not apply to compounds. Since compounds contain more than one lexical root, movement of one could theoretically be visible. Here, however, the problem arises that the relative order of the lexical roots in a compound is meaningful: a *car radio* is not the same as a *radio car*. A movement operation must obviously be detectable for the hearer. If the relative order of lexical items is relevant for establishing thematic relations, movement must be marked in some way. This is easier in syntax, because there are usually more lexical items available, which in addition have different categories. In English, for example, topicalisation of the object is easy to detect, because the default N-V-N order is disrupted and replaced by an N-N-V order. In contrast, topicalisation of the subject would be much harder to detect.

Obviously, if it were necessary to mark movement inside a word, language would find a way to do so, by adding a special functional marker, for example. The question is, however, why movement would take place at all. In syntax, movement is often related to discourse or information structure effects: marking of topic or focus, interrogatives, etc. Such effects are arguably irrelevant or even impossible at the word-level: since words do not convey full propositions, it is not clear what purpose topicalisation or interrogative marking inside a word would serve, for example.

The point here is not to show that movement inside words is fundamentally impossible. For the purpose at hand it suffices to show that the fact that we do not observe movement in morphological structures does not necessarily entail the existence of a separate morphology module. We can account for this observation in a way that is compatible with the idea that there is only a single morphosyntax module.

We can go one step further, however. It is not at all clear whether the observation that there is no word-internal movement is correct. The problem is that movement is often a defining characteristic of syntax. Notions such as "internal fixedness" and "uninterruptability" are typically part of the definition of the term "word" (cf. Haspelmath 2011, p. 38), which means that if we find movement, we automatically assume that the relevant phenomenon is syntactic. This tendency makes it impossible to ever prove or disprove morphological movement, obviously.

Romance object clitics illustrate this point. These clitics appear before the verb if the verb is finite and after the verb if it is non-finite:

(30) a. Lo veo. *Spanish*
3sg.m see.1sg.pr
'I see him.'

b. ver-lo
see.INF-3SG
'to see him'

Crucially, Romance object clitics must appear directly adjacent to the verb and adhere to a fixed order if there is more than one. Furthermore, there are co-occurrence restrictions (cf., e.g., Desouvrey 2005). These are properties that are more typically associated with morphological constructions. There are good reasons, therefore, to treat Romance object clitics as morphological.

Nonetheless, the fact that placement is variable and depends on the finiteness of the verb means that Romance object clitics are generally analysed syntactically (as suggested by the term *clitic*). Rather than considering the option that either the clitic or the verb stem moves word-internally, we assume that clitic placement is a syntactic process.

All in all, the argument that there is no morphological movement is very weak. The kinds of movement (long-distance dependencies) we see in syntax make no sense in morphology, and it is not even clear how to identify morphological movement if the sheer fact that an element can appear in different positions is taken as an indication of its syntactic nature.

4.5.3 Lexical Integrity

It is often claimed that words are subject to some principle that makes them opaque in syntax. Usually called the *Lexical Integrity Principle* (cf. Anderson 1992; Lapointe 1980), it states that the structure of words is not accessible to syntax and that constituent parts of words cannot be separated in syntax. This is taken as evidence for a separate morphology module. It is true that this would indeed be a possible explanation for lexical integrity effects: if words are generated in a separate module, it makes sense that the syntactic module treats them as atomic.

As a first problem, however, note that the lexical integrity requires a proper definition of the term "word", which, as already noted, is not unproblematic. If we cannot define "word", then the Lexical Integrity Principle becomes meaningless. A definition of words as lexico-semantic units (i.e., lexemes) does not suffice, because particle verbs would certainly fall under that definition, even though they are separable in syntax. Nor can we define words as indivisible units in syntax, because that would make Lexical Integrity circular.

Secondly, even proponents of lexical integrity admit that it is not absolute. For example, Booij (2009) concludes:

> In conclusion, this paper has shown that the principle of Lexical Integrity should be formulated in such a way as not to exclude the

> different modules of the grammar from ever having access to word-internal structure. Moreover, Lexical Integrity as the prohibition on syntactic manipulation of word-internal constituents is not an absolute universal, but rather the default situation. (Booij 2009, p. 98)

Similarly, Lieber and Scalise (2006) state that "[…] we know that morphology and syntax interact, and that this interaction is not a one way affair: morphology sees syntax and syntax sees morphology" (p. 10). Neither Booij (2009) nor Lieber and Scalise (2006) wish to argue that the Lexical Integrity Principle should be abandoned altogether, however. Rather, they claim that only a limited interaction between syntax and morphology should be allowed, because, as Lieber and Scalise state, "[…] this possibility predicts far more interaction than we find" (p. 30). They do not, however, explain what sort of interactions we would expect.

In fact, I believe we would not in fact expect much more interaction than what we see. The discussion around lexical integrity usually revolves around words and their opacity to syntactic operations. Words, however, are not the only structural units that are opaque to operations initiated at higher levels of structure. Syntactic phrases are often opaque in similar ways. Take, for example, the claim that constituent parts of words cannot be moved:

(31) Tea$_i$, I have bought a t_i pot.

It is true that English does not allow the fronting of a single member of a compound. However, a similar form of subextraction is disallowed for noun *phrases* as well:

(32) Blue$_i$, I have bought a t_i tea pot.

Why would we ascribe the impossibility of (31) to Lexical Integrity, when the apparently very similar fact in (32) cannot be ascribed to it? This question becomes even more compelling when we consider the following colloquial Russian data from Pereltsvaig (2008, p. 8, 10):

(33) a. Čërnogo$_i$ ja rešila ne pokupat' [$_{NP}$ t$_i$ xleba]! *Russian*
black I decided not to.buy bread
'I decided not to buy black bread.'

b. V vagon ona xodila restoran obedat'.
to carriage she went restaurant to.dine
'She used to go dine in a carriage restaurant.'

These Russian examples show exactly the kind of displacement that English does not allow in (31) and (32). In (33a) an attributive adjective, *čërnogo* 'black' is moved from its base position inside the noun phrase to sentence-initial position. In (33b), part of a lexical compound is displaced.[14] Pereltsvaig (2008) argues that the displaced element can be either a topic or a focus, the crucial aspect that enables displacement being contrastivity. For our purposes, it suffices to observe that displacement is possible.

These data suggest that "word status" is not the decisive property that determines whether subextraction is possible or not.[15] Rather, the possibility of subextraction appears to depend at least in part on language-specific factors. Another contrast, in this case between English and German, suggests the same thing:

(34) a. Which book$_i$ were you looking for t_i?

b. *Welch-em Buch$_i$ hast du nach t_i gesucht? *German*
which-DAT book have.2sg you to looked
'Which book were you looking for?'

English allows preposition stranding in *wh*-questions, German does not: in English, the preposition *for* remains in its base position while its complement, *which book*, is *wh*-fronted. In the equivalent German clause, the preposition must be fronted along with the noun phrase. Stranding *nach* 'to' is not possible.

Lexical Integrity is also associated with the fact that parts of words are not available for processes such as (adjectival) modification, coordination, and pronominal reference. This is illustrated by the following Hebrew examples (from Borer 2009):

(35) a. beyt morá *Hebrew*
house.CNSTR teacher
'the teacher's house'

b. beyt sefer
house.CNSTR book
'school'

Both phrases in (58) are examples of construct state constructions (see also section 3.2.5): a particular nominal construction in which a head noun is modified by a dependent noun. The head noun is a bare noun and cannot take a

[14] Interestingly, it has pied-piped the preposition *v* 'in(to)', which suggests the displacement may be phonological in nature, as it targets a phonological unit, not a syntactic unit. I will not pursue this matter here, though.

[15] In fact, it is not clear how it could be, given the lack of a sound definition of the notion "word".

definite article. The modifying noun is either definite or indefinite and its definiteness determines the definiteness of the entire construction. The construct state construction is historically a possessive construction and can still be used in this way, as (35a) shows.[16] Hebrew has many lexicalised construct state constructions, however, which often have a non-transparent meaning. (35b) is an example of this. Literally, *beyt sefer* means 'house of book', but the lexicalised meaning of the construction is 'school'.

As Borer shows, the two types of construct state constructions have different properties. The lexicalised construction is opaque to a number of operations: the constituent elements are not referential, cannot be modified, coordinated or referred to pronominally. The construction in (35a) is transparent to all of these processes. These examples show that integrity effects can even cut across a syntactic construction: although both examples in (35) are construct states, (35b) shows stronger integrity effects.

One might argue that (35b) is a typical example of a *lexical* integrity effect, because *beyt sefer* 'school' is obviously stored in the lexicon, while *beyt morá* 'a teacher's house' is not. The point is, however, that there is no *structural* difference between the two examples. It is not the case that *beyt sefer* is a 'word' just because it is subject to a strong structural integrity effect. As far as its structure is concerned, it is still a phrase.

Summarising, we can conclude that lexical integrity cannot be defined in structural terms. Similar structures in different languages may have different extraction options, and even within a single language identical structures may be opaque to a different extent. From the Hebrew examples it is clear that referentiality plays an important role in determining the opacity of a structure. The construct state *beyt sefer* as a whole is referential, but its component parts are not: *beyt sefer* refers to the concept SCHOOL, it does not refer to the concepts HOUSE and BOOK. In contrast, the referent of *beyt morá* consists of two entities, a teacher and a house, and both these concepts are referred to by the structure *beyt morá*. The constituent parts of the structure are therefore themselves referential, which makes the structure less opaque.

Similar effects can be observed in Dutch and German particle verbs. In these languages, it is possible to front or scramble the particle of a particle verb under certain circumstances (Zeller 2001; Lüdeling 2001; Müller 2002):[17]

[16] In Modern Hebrew, possession is often realised with the element *šel*, which can be translated with 'of'. In Standard Arabic, the construct state is still the normal way to express possession.

[17] In these examples, the relevant particles are in boldface.

(36) a. An der Haltestelle stiegen hübsche Frauen *German*
at the bus.stop climbed pretty women
ein. **Aus** stiegen nur Männer.
in out climed only men
'At the bus stop, pretty women got on. Only men got off.'
(Zeller 2001, p. 89)

b. Ich weiß, dass die Sonne **auf** im Osten und
I know that the sun up in.the East and
unter im Westen geht.
down in.the West goes
'I know the suns goes up in the East and down in the West.'
(Lüdeling 2001, p. 50)

However, fronting or scrambling is not always possible:

(37) a. *?**Ab** ist Nixon 1974 getreten. *German*
off is Nixon 1974 stepped
intended: 'Nixon resigned in 1974'
(Zeller 1999, p. 64)

b. *... dat Jan het meisje **op** vaak merkte. *Dutch*
that Jan the girl up often noticed
(Neeleman and Weerman 1993, p. 436)

Müller (2002, p. 136) looks at these structures in detail and reaches the following conclusions:

> It is clear that in many instances fronting, separation, and modification are impossible, but this is not due to general properties of particle verbs. In the case of particle fronting and intraposition, other factors interfere, like information structure and the possibility of establishing a contrast. Whether a particle can be modified or not depends on semantic factors. In my opinion, it is the right approach to allow fronting, intraposition, and modification and to account for the appropriate syntactic structures. Additional constraints like those discussed for fronting then rule out or specify the markedness of certain constructions.

Müller's conclusions are clear: the structural integrity effects that we observe with particle verbs are not the result of their structure. They result from external factors: information structure, contrast, semantics. The examples in (35) suggest

that the same is true for Hebrew construct states as well: it is not the structure itself that yields integrity effects, they result from external factors.

Note that this does not mean that *all* integrity effects result from external factors. Structural factors certainly do play a role. Hebrew construct states do not allow extraction, for example, which is presumably a structural restriction related to the fact that the head noun is in construct state form. Similarly, the impossibility of preposition stranding in German may be a result of the fact that the preposition assigns (overt) case to the noun phrase. All in all, the factors involved in integrity effects are diverse and to some extent language-dependent. It is beyond the scope of this section to explain all integrity effects, but what is clear from the discussion is that integrity effects do not distinguish morphological structures from syntactic ones. For this reason, they do not provide sufficient reason to assume that syntax and morphology constitute separate modules.

It is nonetheless interesting to speculate on the cause of integrity effects for a moment. One thing that stands out is that in general, structures that are transparent are semantically complex, while semantically simplex structures are syntactically opaque, even if they are syntactically complex. In other words, syntactic transparency is dependent on semantic complexity.

This is best illustrated by the Hebrew examples in (35). The construct state *beyt sefer* 'school' is opaque to a number of operations; semantically, it is simplex. This contrasts with *beyt morá*, which is transparent to the same operations and is semantically complex: it references two concepts rather than one. Similarly, those particle verbs that allow fronting of the particle are to a certain extent semantically transparent.

There thus seems to be a general integrity principle, which we may formulate as follows:

(38) *Structural Correspondence*
Semantic units correspond to syntactic units, and vice versa.

Obviously, we would need to specify what exactly constitutes semantic and syntactic units. Relevant syntactic units could be all levels of structure, all maximal projections, or possibly only those maximal projections that are also phases (at least CP, *v*P, DP). On the semantic side, the term *unit* is even more difficult to define, since not only concepts are intended, but also larger structures. Possibly, relevant structures are all structures that do not contain free variables.[18]

[18]The name *Structural Correspondence* is deliberately reminiscent of Linear and Input Correspondence. In a fundamental way, the three principles all do the same thing: keeping together what belongs together across module boundaries.

Whatever the definitions, however, the intuitive idea is that syntactic elements can only be transparent to operations that originate outside their containing phrase if they themselves constitute a semantic unit of some kind. Movement, for example, only applies to elements that form a meaning unit. Obviously, they may be part of a larger meaning unit (ultimately, they are part of a sentence / proposition), but they are semantically 'complete' in some sense.

If these speculations are on the right track, they have direct consequences for the morphology / syntax discussion. The smallest semantic units, those that are simplex, will tend to correspond to indivisible syntactic units, i.e., to heads. In turn, syntactic heads tend to correspond to prosodic words. These two tendencies strengthen the syntax / morphology illusion, but this is not a deep property of the language faculty. Both tendencies are interface effects: if semantic units would not correspond to syntactic units, it would be impossible to encode any complex meaning in a reliable, reconstructable manner. Similarly, the prosodic word is the lowest constituent of the prosodic hierarchy that can provide enough forms to accommodate the wide variety of concepts that the semantic system provides, therefore, syntactic heads, especially content words, will tend to correspond to (at least) a prosodic word.[19]

All of this is speculation, of course, and delving into these questions further would go well beyond the scope of this section. One consequence of a principle such as Structural Correspondence is worth noting, however. Since Ross's (1967) seminal work on restrictions on transformations, one of the leading questions that syntactic research has been trying to answer is why certain structures are "islands", i.e., do not allow extraction. Given Structural Correspondence, however, we would expect opacity to be the norm rather than the exception. That is, the question that we should ask is perhaps not why certain structures are opaque, but rather why certain structures are transparent. I will leave these questions open for future research, however.

4.6 Morphology by itself

4.6.1 Introduction

The final argument often put forward in favour of a separate morphology module is what might be called *autonomous morphology*: the idea, put forward by Aronoff (1994), that there are generalisations that can only be accounted for in terms of morphology. Aronoff argues that there is an additional level of morpho-

[19] In fact, some languages, such as Chinese, seem to have opted for the syllable rather than the prosodic word, which is probably made possible by the lexical use of tone, which of course greatly expands the number of possible phonological forms.

logy between the levels of syntax and phonology and that every mapping from syntax to phonology passes through this morphological level. Maiden (2005) introduces the term *morphome* for phenomena of this kind. A morphome is an abstract morphological function that can be applied to a word form to yield a morphologically derived form. The crucial point of the morphome is that it cannot be reduced to purely syntactic or purely phonological properties.

In this section, I discuss a number of phenomena that fall in this category and I show how they can be analysed without the assumption of a morphology module sitting between syntax and phonology. It is inevitable that in doing so, we need to accept the fact that syntax can deal with idiosyncrasies and that there are syntactic rules with a limited domain of application. Crucially, however, I believe that the goal of unifying syntax and morphology can be achieved without sneaking in a morphology module into the syntactic system, as other proposals have been forced to do. What this means is that when morphological principles are introduced into syntax, previous proposals have been forced to limit their application to certain types of constructions in some artificial manner, e.g., by distinguishing between $X^{>0}$ and $X^{\leq 0}$ heads.

4.6.2 The Romance N-pattern

A well-known example of autonomous morphology is the so-called N-pattern, discussed by Maiden (2004). The N-pattern is a pattern occurring in Romance verbal morphology, demonstrated in table 4.1 with the Spanish verb *jugar* 'to play'.

	sg	pl
1st	***jueg***-o	jug-amos
2nd	***jueg***-as	jug-áis
3rd	***jueg***-a	***jueg***-an
impt	***jueg***-ue	jug-ad
inf	jug-ar	

Table 4.1: N-pattern: *Spanish*

The N-pattern is a verb stem alternation that contrasts the singular and 3rd person plural forms with the 1st and 2nd person plural forms. It occurs in many Romance languages and applies to different categories of verbs. Furthermore, the exact stem alternation differs from verb to verb. In Spanish, for example, two frequently occurring alternations are *o*/*ue* and *e*/*ie*, but there are a few others,

which occur in limited sets of verb, such as *u/ue*, which only occurs in *jugar* and *oler* 'to smell'.[20]

The alternation can also be suppletive, as in the French example *aller* 'to go', which uses the stem *v(a)-* in the singular and 3rd person plural forms, and the stem *all-* in the 1st and 2nd person plural forms, as demonstrated in table 4.2.

	sg	pl
1st	***v***-ais	all-ons
2nd	***va***-s	all-ez
3rd	***va***-	***v***-ont
impt	***va***	all-ez
inf	all-er	

Table 4.2: N-pattern: *French*

In short, whether a verb is subject to an N-pattern alternation or not is not predictable from the phonological form of the verb. In Spanish, for example, the *e/ie* and *o/ue* alternations are fairly common, but not all verbs with *e* or *o* partake in them (e.g., *tensar* 'to tighten', *poner* 'to put'). This means that a purely phonological explanation is impossible.

Likewise, there is no syntactic or semantic property that distinguishes the two sets of forms. Therefore, so the argument goes, we need a separate morphology module to account for such facts. If morphology were just syntax, stem alternations such as this one would require a syntactic explanation, i.e., in terms of some syntactic feature or a combination of several features. There is no single feature or set of features that describes 1st, 2nd and 3rd person singular and the 3rd person plural. Only a disjunctive feature set could describe these forms, but disjunctive feature sets do not correspond to natural classes.

However, this argument only holds on a specific view of syntax: it presupposes that there is only a limited (presumably innate) set of syntactic features. This view automatically relegates any features not part of this set to morphology and/or the lexicon. This is indeed a common assumption in the minimalist literature, but it should be noted that it is an assumption based on *theoretical* considerations. There is no *empirical* evidence that supports it. Obviously, we must assume that the N-pattern is learnt on the basis of input forms, no-one would propose that the N-pattern is innate. However, this fact in itself is not sufficient to argue that the N-pattern is not syntactic, unless one adopts this as an axiom.

[20]Note that the stem variant of *oler* is spelt *huel-*. The *h* is added for orthographic reasons.

In fact, as argued for in chapter 1, it is unlikely that syntax operates with a set of innate features. Rather, we must assume that during language acquisition, an I-language is developed by extracting a set of features from the input and, crucially, that this is the *only* source of features. As a consequence, syntactic features are language-specific, and this fact means that the problem we are facing disappears. Although the singular and 3rd person plural forms do not form a natural class, they clearly pattern alike in the Romance languages, which means that they provide sufficient basis for the syntactic system to generalise a feature [±NPAT] that captures these forms (cf. Kremers 2014, for details).

For Spanish, an alternative analysis makes some sense. Harris (1985) assumes that an alternating verb such as *jugar* has a phonological representation with two slots, unlike verbs that do not have the relevant stem alternation. Only one of these slots is associated with a vowel, except under certain circumstances, specifically, when the syllable is stressed. In these cases, the second slot is also associated with a segment, resulting in a diphthong. This solution works for Spanish, because it can account for the fact that the relevant vowel alternation appears to be linked to stress, since the same alternation also shows up in other contexts (e.g., *viéjo* 'old' vs. *vejéz* 'old age'). Obviously, however, it will not work for suppletive N-pattern alternations such as French *aller*.

For this reason, I prefer the analysis with a [±NPAT] feature. Based on the observable forms of the verb, those forms that show the alternative stem form are marked with [+NPAT]. Since [±NPAT] has the task of matching person/number endings with verb stem forms, I assume that both the relevant T heads and the relevant verb stems are marked with [±NPAT]: a T head that has a [sg] feature or the features [3pl] has an additional [+NPAT] feature, and the verb stem *jueg-* has a [+NPAT] feature. Ignoring the semantic components, the lexical entries for a [2sg] T head and the verb stem *jueg-* are the following:[21]

(39) a. [T, 2sg, +NPAT] ↔ /s/
b. [V, +NPAT] ↔ /xueg/

When the verb stem *jueg-* occurs in the structure, it carries [+NPAT], which must be matched by the person/number ending. [2sg] is compatible with *jueg-*, but [2pl] is not, because it has a [–NPAT] feature.

The feature [NPAT] is a binary feature because we need to be able to distinguish between "normal" verb stems of verbs that alternate (i.e., *jug-*) and verb stems of verbs that do not alternate (e.g., *pon-* of *poner* 'to put, place'). To see why, consider what would happen if NPAT were a privative feature and the stem

[21] The actual pronunciation of the verb stem *jueg-* is closer to [xʊ̯eɣ]. The lexical item in (39b) reflects the underlying form.

alternant *jug-* were selected in combination with a [2sg] T head. The T head in this case would not be incompatible with the verb stem, resulting in the ungrammatical form **jugas*. Obviously, T has a feature that V lacks, but it would require an additional stipulation to rule out this combination: T has other features that V lacks (such as φ-features) that are not problematic for the derivation.

Therefore, the verb stem alternant *jug-* must be explicitly marked as being incompatible with an NPAT feature, as in (40):

(40) [V, –NPAT] ↔ /xug/

Now, when *jug-* is combined with a [2sg] T head, the two heads do not match, because they have opposite values for the NPAT feature. On the other hand, verbs that do not participate in the N-pattern alternation simply do not have an NPAT feature. When a verb stem such as *pon-* is combined with a [2sg] T head, which has a [+NPAT] feature, there is no clash. Although T has an NPAT, it is still compatible with V, because the latter lacks NPAT altogether.

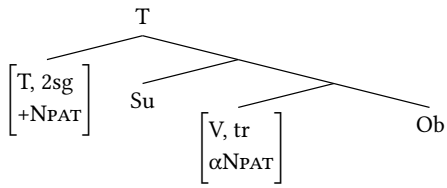
From a minimalist perspective, this analysis may raise a few questions. NPAT is obviously an uninterpretable feature, i.e., it cannot be interpreted semantically. However, as pointed out by Jackendoff (1997), this is not a real issue. Because the semantic system cannot deal with the NPAT feature, it is the task of the syntax/semantics interface to filter out this features. Interfaces are needed when systems that are fundamentally different (because they perform different tasks) need to communicate. Because of this fundamentally different nature, it is inevitable that the representations employed in the two systems are at best partially compatible. Syntax, for example, deals with heads such as N, V and P. These, however, do not have direct equivalents in semantics. Not every N is an entity, not every V is an event and not every P corresponds to a relation (take prepositional objects, for instance).

By their very nature, then, interface systems must ignore any property in one representation that has no equivalent in the other representation. Therefore, if syntax employs an NPAT feature, because the phonological component of a head depends on it, the interface to the semantic system can simply ignore this feature when it encounters it. There is no need to have the derivation crash for this reason. To put this in somewhat different terms, a feature is essentially just a formalisation of a property. Some verbs have the property that they alternate according to the N-pattern and this property must be marked formally. Under certain circumstances, we must invoke this property to ensure that the derivation yields the correct word form, but the feature serves no purpose beyond that.

Another problem that arises from the perspective of minimalism is how the agreement between T's and V's NPAT features is established. Current consensus

in the minimalist literature seems to be that in order to initiate Agree, the probe must be *active*, i.e., it must have at least one unvalued feature. For NPAT this is not entirely unproblematic. Consider the constellation under which Agree with respect to NPAT could take place:

(41)



In this example, I have left the NPAT feature of V unvalued (as indicated by the α). It is clear that it cannot be T that has an unvalued NPAT feature, because the 2sg ending is invariably [+NPAT], which means that T cannot initiate Agree based on NPAT. T does have other unvalued features (φ-features, for instance), but it is not immediately clear whether they would suffice to establish an Agree relation that can value V's NPAT feature, because on standard assumptions, there is no other Agree relation that exists between T and V.[22]

A more intuitive alternative would be to say that when T is merged, its features must be compatible with the features of V, by virtue of the c-selectional relation between T and V. That is, when a category A selects and merges with a category B, A and B cannot have incompatible feature sets. Two feature sets are incompatible if they both contain the same feature but with different values. So if T has [+NPAT], V is incompatible with it if it has [–NPAT]. V is compatible if it does not have an NPAT at all, if it has a valued [+NPAT] feature, or if it has an unvalued NPAT feature. In the latter case, merger with T will result in V's NPAT being valued.[23]

Such a feature matching requirement should be an uncontroversial addition to the system, especially because it is required anyway. Determiners, for example, usually have restrictions on the type of nouns they can be merged with: indefinite pronouns cannot occur with non-count nouns, many languages mark determiners for number, gender and case, which must match the corresponding features on the noun, etc. Although some of these matches could be handled

[22] Furthermore, Agree between T and V would be Agree between a head and its complement (or the complement of its complement, if we factor in *v*). While technically not impossible, it would be an uncommon form of Agree.

[23] That is, depending on theoretical preferences, one may assume that verbs that are subject to an N-pattern alternation have an unvalued NPAT feature, which is valued during the derivation and subsequently forces the selection of one of the stem alternants, or one may assume that V is merged with a valued NPAT feature which T must then match.

by Agree,[24] case matching between D and N is more problematic, because case is assigned from outside the noun phrase. Even if D's unvalued CASE feature could initiate Agree, it would only find another unvalued CASE feature on N. By standard assumptions, no Agree relation can be established in such cases.

While it would be possible to subsume this kind of feature matching under Agree, it would require extending the definition of Agree. That being the case, there is no conceptual harm in opting for a different relation, say Match, that must exist between a lexical head and its extended projection: the feature set of each head in the extended projection line of a lexical head must be compatible, in the sense described above, with the feature set of the lexical head.[25]

For present purposes, it suffices to show that a separate morphology module is not the only way to account for autonomous morphology. How exactly individual cases can be analysed is a different question and is obviously beyond the scope of this paper. The example suggests that relevant phenomena can be analysed syntactically or phonologically. Note that I am not claiming that all relevant phenomena can be analysed uniformly. Some phenomena may require a syntactic analysis, while others can be handled in phonology. This lack of uniformity does not mean that positing a separate morphology module is a better analysis, however. There is no guarantee that all relevant phenomena can be handled uniformly in a morphology module: we may need different morphological mechanisms to deal with different kinds of phenomena.

The only point of assuming a separate morphology module would be that it enables us to isolate language-dependent idiosyncrasies from syntax. This is a *conceptual* issue, however, one that is not relevant in a framework such as minimalism, which aims to relegate all idiosyncrasies to the lexicon. Note that assuming an [NPAT] feature is compatible with this aim: [NPAT] is a feature of individual lexical items. There is no need to adopt a special syntactic rule to accommodate it (cf. Kremers 2014, for details).

4.6.3 English past participles

The second example of autonomous morphology that I discuss could be called a text-book example: past participle formation in English. There are two points about past participles that Aronoff (1994) argues cannot be accounted for without

[24] If we allow for Agree between a head and its complement.

[25] Note that there is also a concept often called 'concord', agreement between a noun and its adjectival modifiers. This, however, is not a single process but rather the result of two independent processes, an Agree relation inside the AP between the adjective and its AP-internal subject and a binding relation between D and the AP-internal subject. See Kremers (2005) and Struckmeier and Kremers (2014) for details.

reference to an independent morphology module: their formation cannot be described as a simple suffixation, and both passive and perfect participles are identical for each verb and moreover, if historical change affects the past participle of a particular verb, both the passive and the perfect participle are affected simultaneously. The latter fact strongly suggests that the passive and perfect participles are in fact one and the same form, which is indeed a common assumption and the reason why we usually use the term *past participle* to refer to it. Aronoff's claim is that there is no syntactic reason why these two functions are performed by the same word form and that therefore there is a morphological generalisation at work.

However, the claim that there is no syntactic reason that a single word form can perform both functions is not correct. Haider (1986) shows that the past participle can simply be described as a word form with a suppressed external argument. The choice of auxiliary, either *to be* or *to have*, determines how this form is used. The auxiliary *to have* unblocks the external argument, thus recreating the argument grid of the underlying verb and allowing the past participle to be used in an active perfective construction. The auxiliary *to be* does not unblock the external argument. Because English clause structure requires a subject, the highest argument in the remaining argument grid of the past participle is raised to subject position. This argument of course retains its internal theta role, which means that the external theta role remains suppressed.

The other issue that Aronoff mentions is the fact that past participle formation is not a single morphological process. It can be formed by suffixation (*work ~ worked*, *beat ~ beaten*), by ablaut (*sing ~ sung*) or by a combination of both (*break ~ broken*). Again, the argument is that there is no syntactic nor phonological reason why some verbs use suffixation while others use ablaut. Therefore, a separate morphology module must be involved.

In English, past participle formation has a lot in common with past tense formation: there is a regular suffix *-ed* that most verbs take, and there are a number of subregularities, such as *ring ~ rung*, in addition to truly irregular forms. The discussion of the past tense debate in section 1.2.4 should therefore provide some insights into a possible analysis. The main conclusion that Westermann and Ruh (2012) draw is that the best way to model past tense formation is to use a single-mechanism model, i.e., a model that does not make a strict distinction between a rule-based mechanism for regular forms and a lexicon-based look-up method for irregular forms. Rather, a single-mechanism model is more appropriate, which handles regular and irregular forms in much the same way.

For a high-level analysis such as the one discussed here, this means that we need a single mechanism for describing both the general rule and the irregular forms. The format that I proposed in section 1.2.4 is a schema of the following form:

(42)
$$\begin{bmatrix} \textsc{syn}|\textsc{head} & \textit{past-verb} \\ \textsc{phon} & \boxed{1} \oplus \langle \text{d} \rangle \\ \textsc{h-dtr} & \begin{bmatrix} \textsc{syn}|\textsc{head} & \textit{verb} \\ \textsc{phon} & \boxed{1} \end{bmatrix} \\ \textsc{n-h-dtrs} & \langle [\text{past}] \rangle \end{bmatrix}$$

This format is obviously inspired by HPSG, but can easily be adapted to the present model:

(43)
$$\begin{bmatrix} \text{ppp} \\ u\text{V} \end{bmatrix} \leftrightarrow /\text{d}/|_{\omega}$$

Here, ppp stands for *past participle*, the verb form that the structure builds. The structure in (43) is a lexical entry whose syntactic component consists of a past participle feature and an unvalued V feature, which means that it c-selects a verb.[26] The phonological component consists of the consonant /d/, which has a prosodic alignment requirement. The combination of a c-selectional feature and a prosodic requirement means that (43) is a suffix.

For the past tense subregularity of the type *ring* ~ *rang* I proposed the following schema:

(44)
$$\begin{bmatrix} \textsc{syn}|\textsc{head} & \textit{past-verb} \\ \textsc{phon} & \langle \boxed{1} \mid \text{æ}, \text{ŋ} \mid \boxed{2} \rangle \\ \textsc{h-dtr} & \begin{bmatrix} \textsc{syn}|\textsc{head} & \textit{verb} \\ \textsc{phon} & \langle \boxed{1} \mid \text{ɪ}, \text{ŋ} \mid \boxed{2} \rangle \end{bmatrix} \\ \textsc{n-h-dtrs} & \langle [\text{past}] \rangle \end{bmatrix}$$

The equivalent subregularity in past participle formation is *ring* ~ *rung*. Adapting the schema to the current format is slightly more difficult, because (44) contains a restriction on the phonological form of the selected verb. In principle, we have two options open to us. We could either encode the restriction in syntax or in phonology.

[26] Obviously, the entry is simplified in several ways. Apart from the missing semantic component, I have also not indicated how the structure suppresses the verb's external argument. The answer to this question depends on the way the external argument is implemented, which is an issue that is orthogonal to the discussion at hand.

The syntax option would be fairly straightforward. We would need to assume a feature that marks the verbs to which this pattern applies, which would be a language-specific feature much like the NPAT feature of the previous section. It is obvious that the linguistic input provides enough information for establishing this feature, therefore it would be a safe assumption. Let us label this feature as SUBREGULARITY(ing) or SR(ing) for short. Then we can adopt a lexical entry of the following form:

(45) $\begin{bmatrix} \text{ppp} \\ u\text{V[SR(ing)]} \end{bmatrix} \leftrightarrow \begin{bmatrix} \text{+low} \\ \text{+central} \end{bmatrix}$

The syntactic component of (45) contains a c-selectional restriction that is more specific than just a category. It specifically requires an element of category V that has the feature [SR(ing)]. Such an 'extended' c-selectional restriction is admittedly an addition to the theory, but it is not a novel property. It is essentially an extension of an existing property, the ability of a head to select the category of its complement. On the assumption that the category is basically just a feature, it is not a fundamental change to assume that other morphosyntactic features can also be selected for.

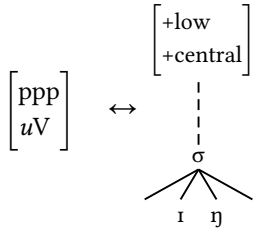
The phonological component of (45) consists of two distinctive features, [+low] and [+central]. These are the features that distinguish [ɪ] from [ʌ], which is the vowel change that occurs in this particular subregularity. The idea is that the lexical entry in (45) adds these features to the phonological structure, overwriting the corresponding features of [ɪ].

The features [+low, +central] are features of segments, but they are not segments themselves and therefore the phonological component of (45) is a prosodic morpheme. As such, the lexical entry in (45) is subject to Input Correspondence, which means that the features [+low, +central] must be added to the phonological component of the head that (45) selects, which is obviously the verb. Since the verbs to which (45) applies are typically monosyllabic, the features are associated with the vowel, changing [ɪ] into [ʌ].

This syntactic analysis has an important drawback, however. Unlike the N-pattern discussed in the previous section, there is a phonological aspect to the *ring* ~ *rung* subregularity. The lexical entry in (45) does not capture this intuition, however. The feature SR(ing) could in principle be added to *any* verb, even those that have a different phonological form. In order to express this intuition, the model would require a rule that determines to which kinds of verbs the feature SR(ing) can be added. It is not clear, however, where such a rule would be located in the grammar. Since its target is individual lexical items detached from any syntactic context, its most natural location is the lexicon. A lexicon with

rules, however, especially with rules that are relevant for syntax, is essentially a morphology module.[27] Therefore, if possible, we should try to avoid them.

These considerations leave us with the phonological option. If we can encode the phonological restriction in the phonological component of the lexical entry for the past participle, we would not need this additional rule. Interestingly, the phonological option is quite feasible:

(46) 

Again, I have simplified the lexical entry. The syntactic component still has the feature ppp and the c-selectional feature, but the latter no longer needs the additional requirement for the feature SR(ing). The phonological component is more complex now. It still contains the features [+low, +central], but these are now attached to a syllable that contains the segments [ɪŋ]. The association line is dashed, which indicates that it is a newly established association.

The dashed association line is actually crucial for the representation in (46). A dashed line is commonly used in the phonological literature to indicate that an association is newly formed as the result of a phonological process. In the lexical entry in (46), it indicates that the features [+low, +central] are added to an existing syllable […ɪŋ…]. That is, the phonological component of a lexical entry may specify two things: phonological material that is added to the structure, and phonological material that must already be present.

In other words, the phonological component specifies both the form of a lexical entry and the context in which it must be used. The context acts as a phonological requirement, specifying when the lexical entry is licit. In order to distinguish the two in the representation, I adopt the convention that the structure above the dashed association line is the material to be added and the structure below it is the context in which it is added.

Distinguishing between these two parts of the structure is crucial. The phonological component of a lexical entry must be able to specify the phonological material that it contributes itself, but it must also specify the context in which it can be used. Up until now, the context has consisted of prosodic boundaries or prosodic constituents, but the possibilities are not limited to prosodic structure.

[27] Lexical redundancy rules would be an exception to this, since they do not alter or create structure.

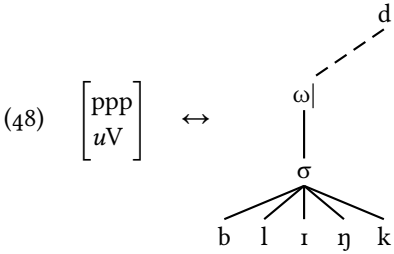
In principle, anything that can be expressed phonologically can serve as context in the prosodic component of a lexical entry.

Note, by the way, that the entry in (46) has the two properties that define an affix: the syntactic component has a c-selectional restriction and the phonological component has a prosodic requirement. The idea that such 'non-segmental' morphemes are affixes is not entirely new (e.g., Hayes and Lahiri 1991), but in the present model it becomes clear that from a technical point of view, they really *are* affixes.

As discussed in section 1.2.4, the lexical entry in (46) would incorrectly capture the verb *blink*, which forms its past tense and past participle form with the regular suffix *-ed*. In order to capture the fact that in does not form its past tense with the ɪ~æ ablaut, it was necessary to adopt a schema specifically for the verb *blink*:

(47)
$$\left[\begin{array}{ll} \textsc{syn}|\textsc{head} & \textit{past-verb} \\ \textsc{phon} & \boxed{1} \oplus \langle \text{d} \rangle \\ \textsc{h-dtr} & \left[\begin{array}{ll} \textsc{syn}|\textsc{head} & \textit{verb} \\ \textsc{phon} & \boxed{1} \langle \text{b, l, ɪ, ŋ, k} \rangle \end{array}\right] \\ \textsc{n-h-dtrs} & \langle [\text{past}] \rangle \end{array}\right]$$

Converting this into the format used here, we obtain (48):

(48)


In (48), the context in which the segment /d/ is to be inserted is the right edge of a prosodic word that is monosyllabic and contains the segments /blɪŋk/. The /d/ is associated with the PWd boundary to ensure that it is suffixed to the stem.[28]

In a rule-based approach, there would be no need to adopt a specific rule for *blink*, but as Westermann and Ruh's 2012 model shows, a neural network tends to

[28] Strictly speaking, it would suffice to associate /d/ with the stem. The phonotactic rules of English do not allow /d/ to be positioned anywhere other than at the end. However, I prefer to encode the fact that /d/ is a suffix directly into the lexical entry, since this is what I have done for all other suffixes.

find a different solution, one in which forms such as *blinked* are stored, because at the low-level model that an artificial neural network provides, the form *blinked* is hard to learn due to interference from the other word forms that the network must be able to produce. In terms of the higher-level model developed here, the lexical entry in (46) would apply to *blink* because it matches the required context of the entry. The only way to prevent this is to adopt the extra entry in (48).

Note, however, that in order to ensure that (48) is indeed applied in the case of *blink* and that (46) is blocked, we also need to invoke the *Elsewhere Principle*. The Elsewhere Condition is a staple of morphological theory, which, according to Anderson (1992), has roots going back to Pāṇini. It essentially states that the application of a specific rule blocks the application of a more general rule. In the example at hand, the lexical entry in (48) is more specific than the one in (46), and hence if the former can apply, it does. Given that the Elsewhere Condition is such an important principle in morphology, it stands to reason that we need to adopt some version of it in a model that unifies syntax and morphology. I postpone a more detailed discussion of this principle until section 5.1. For the moment, it suffices to say that the Elsewhere Condition ensures that the lexical entry in (48) is applied to *blink*, rather than the entry for the ɪ~ʌ ablaut in (46) or even the general entry in (45).

Past participle formation is taken as an example showing the need for a separate morphology module, since its formation seems to depend on an abstract feature, say [ppp], that may be realised as a suffix (*-ed* or *-(e)n*), as ablaut, as a combination of both, or irregularly. However, on the assumption that syntactic features are not innate but acquired, there is no reason why [ppp] could not be a syntactic feature, as long as we are able to encode the relevant morphological processes as lexical entries. This is indeed possible, as I have argued.

The past participle is also thought to be subject to a generalisation that can only be expressed in morphological terms. The same form is used both for the passive and for the perfective. Again, this argument is not very strong. The past participle is formed by a process that suppresses the external argument. This verb form can then be used to construct the passive but also the perfective, on the assumption that in the latter case, the auxiliary *to have* contributes an external argument.

The conclusion that we can draw from the discussion is that the past participle does not provide a sufficient argument for a separate morphology module. The only real concession that we need to make in order to accommodate it in syntax is the assumption that syntactic features can be language-specific, acquired by generalising over input forms. Since we have already made this assumption for independent reason (cf. chapter 1), there are no obstacles for adopting a syntactic analysis of the English past participle.

li	AUX	DAT	ACC/GEN	*se*	*je*
Q (question particle)	auxiliaries (except *je*)	dative pronoun	accusative/ genitive pronoun	REFL (reflexive pronoun)	3SG AUX

Table 4.3: Serbo-Croatian clitic cluster (Browne 1974)

4.6.4 Templatic morphology

General considerations

Another phenomenon that is often considered to provide a strong argument in favour of a separate morphology module is what is sometimes called *templatic morphology*. When multiple clitics target the same position, they usually appear in a fixed order that is arguably not predictable from their syntactic functions nor from their phonological properties. A well-known example is the clitic cluster in Serbo-Croatian, which is discussed extensively in the literature (see, e.g., Schütze 1994 and Bošković 2012 and references cited there). Browne (1974) proposes the template displayed in table 4.3 for the clitic cluster in Serbo-Croatian.

Schütze (1994) argues that clitic ordering within the clause is purely phonological, while clitic ordering within the clitic cluster is subject to morphological constraints:

> While it is conceivable that a clitic's need for a host to its left or right could be a morphological requirement, I will argue later that the satisfaction of that requirement is subject to purely phonological constraints and does not show the usual characteristics of a morphological process. On the other hand, ordering within the clitic cluster shows these characteristics very clearly (p. 48).

Obviously, the prosodic syntax model cannot appeal to morphology in order to account for the ordering of templatic clitic clusters. Bošković (2012) argues that clitic ordering is determined by syntax, which makes his model more compatible with prosodic syntax, but he has difficulty explaining the exceptional position of the 3sg auxiliary *je*.

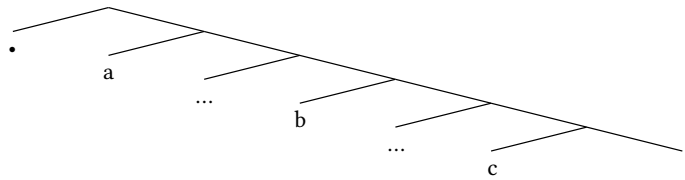
Before we look at this case in somewhat more detail, I would like to discuss the requirements and restrictions that an analysis in line with prosodic syntax would be subject to. In principle, two types of account would be possible: a phonological one and a syntactic one.

A phonological account of clitic ordering would be based on the assumption that phonological composition places all clitics in the same position in the linear

string, after which phonological properties of the elements in question determine how they are ordered. There are not many phonological rules that would be capable of mandating an order among clitics, however. Obviously, each language has general phonological rules and specific phonotactic rules that refer to segments, but violations of such rules are commonly solved by modifying, deleting, inserting or reordering *segments*, whereas clitics are (generally) full syllables. Rules that affect and possibly reorder syllables, on the other hand, are usually not sensitive to the segments contained in those syllables. In other words, a purely phonological account of the order of clitics in a templatic clitic cluster is implausible.

A syntactic account of the order within a clitic cluster would require that the clitics appear in the linear string either in the same order in which they occur in the syntactic structure, or, depending on assumptions, in the reverse order. To see this, consider the following abstract tree:

(49)



Suppose that a, b and c are clitics, with other material intervening between them, indicated by the ellipses. The dot represents an element whose phonological component provides a host for the clitics. By the theory developed so far, the clitics would stay in their base positions and their phonological components specify their attachment site. Their lexical entries would look like (50):

(50) a ↔ α

{ ... ω| ... }$_{\text{IntP}}$

The phonological material that is to be inserted is represented here as α. It is associated with the right boundary of a prosodic word that is contained in an intonational phrase. This IntP is the association domain of the clitics a, b and c, which means that the domain in which they are placed is the IntP. Note that although the clitics under discussion are second-place clitics, I have left the position of the PWd within the IntP unspecified in (50). The reason for this is that the principle of Left-to-Right Association will ensure that α is associated with the first PWd. It is therefore not necessary to specify this in the lexical entry.

A bit more needs to be said about the association domain at this point, because the analysis sketched here may appear to violate Input Correspondence: on the assumption that the auxiliary clitics are T heads, for example, they select V (or *v*). Input Correspondence would then require them to be associated with Φ(V). Recall, however, that in section 3.1.1 I discussed the fact that the negative head shake in German Sign Language can be associated either with the verb, i.e., Φ(V), or with the entire verb phrase, Φ(VP). I accounted for this fact by assuming that there are actually two lexical entries for Neg°, one associating it with the PWd and one associating it with the p-phrase. The latter option means that although Neg° syntactically selects for a (projection of) V, its phonological component can be associated with the entire p-phrase corresponding to the VP. This, I argued, is not in violation of Input Correspondence, because the head shake must still be associated with the p-phrase containing Φ(V).

Although the head shake is not a clitic,[29] its association domain is nonetheless important, serving the same function as the association domain of a "real" clitic. Of course, clausal second-place clitics do not necessarily have a c-selectional restriction, but those that do are automatically subject to Input Correspondence, because as special clitics, they also have a prosodic requirement. However, because they have the IntP as their association domain, they can appear in second place while still obeying Input Correspondence, as long as they are contained within the IntP that also contains the phonological component of the verb they c-select.

Having said this, let us return to the tree in (49). If we assume that the clitics a, b and c do not move, and have lexical entries along the lines of the example in (50), then there are in principle two orders that the clitics could take, as demonstrated in (51), where H is the phonological host and the =-sign indicates phonological cliticisation:

(51) a. H=α=β=γ

b. H=γ=β=α

Which of these two is correct is an empirical issue, and it is in principle possible that different languages display different orders. Any order that deviates from them requires a specific explanation. For example, if one of the clitics has different alignment properties than the others —if all clitics are enclitics and one is a proclitic— then the proclitic would appear first in the cluster. Another option may be that one or more of the clitics actually move in syntax, which could

[29] In fact, as an autosegmental morpheme it has a prosodic requirement and because in c-selects V in syntax, it actually has the two properties that are characteristic of affixes. As discussed earlier, since I do not accord any theoretical status to the term 'affix', one may label it as such, or refrain from doing so, depending on one's preferences.

change the hierarchical order among them, leading to a different order in the cluster.

Serbo-Croatian clitics

With this discussion in mind, let us look at the Serbo-Croatian clitic cluster in somewhat more detail. First, as shown in table 4.3, the general clitic order in the cluster is *li*-AUX-DAT-ACC/GEN-*se*-*je*. Ignoring *je* for the moment, this order essentially reflects the expected order of these elements in the clausal structure. The question particle *li* is generally associated with the C domain, presumably C°. The auxiliaries that can appear in the AUX slot are clitic forms of the verb *htjeti* 'will, to want', expressing future tense; clitic forms of the aorist of the verb *biti* 'to be', expressing the conditional ('would'); and clitic forms of the copula verb *jesam*,[30] which is also used together with participles to create analytic tense forms. Since all three auxiliaries can (or even must) appear in combination with a lexical verb, it stands to reason to treat them as instantiations of T°.[31] As such, they are below *li* in the projection line of the clause.

After the AUX slot there are slots for arguments: dative, accusative/genitive[32] and reflexive.[33] These, too, appear in the order in which we might expect them based on general assumptions about clause structure: verbs that have both dative and accusative arguments usually assign accusative to their lowest internal object. Lastly, reflexive elements such as *se* are cross-linguistically often closely associated with the verb, because they appear to change the argument grid of the verb rather than simply fill an argument position. For this reason, it is expected that *se* should appear below the arguments.

The order in the clitic cluster thus appears to reflect the hierarchical order of the clitics' base positions in the clause.[34] The only obvious exception is *je*, which is the 3sg form of the copula verb *jesam*. This element always appears last in the clitic cluster, regardless of the fact that it is functionally the same as the other forms of the copula verb. Ignoring *je* for the moment (I will return to it below), we can maintain that Serbo-Croatian instantiates the order in (51a), i.e., H=α=β=γ. In this respect, the clitics appear to contrast with affixes, which are subject to

[30] *Jesam* is a defective verb with only present-tense forms. The form *jesam* is actually the 1st person singular form, there is no corresponding infinitive. (Although *jesam* is sometimes analysed as the imperfective of the verb *biti* 'to be'.)

[31] Bošković (2012) treats them as V_{aux} heads that c-select a lexical V and raise to AgrS, a position above T.

[32] As Schütze (1994, fn. 3) points out, the accusative and genitive pronoun clitics are largely homophonous and the data on their relative order are murky.

[33] Note, though, that the dative reflexive *si* appears in the dative slot. The element *se* can, as Schütze (1994, fn. 4) notes, either be a "true reflexive object pronoun or a particle associated with particular verbs."

[34] A conclusion also reached by Bošković (1995, 2012), whose data and analysis I discuss below.

the Mirror Principle (Baker 1985),[35] which has the practical effect that affixes whose syntactic function is positioned lower in the tree are closer to the root they attach to. As a result, the relative order of suffixes reflects their hierarchical order *in reverse*: in a suffix cluster, the hierarchically lowest element comes first. In the Serbo-Croatian cluster of second-place enclitics, the hierarchically lowest element comes last. This raises an important question, because both types of elements share their direction of attachment: they both attach to a preceding PWd.

A possible explanation for this discrepancy may be found in the association domain of the relevant elements. Suffixes have the PWd as their association domain. Since they must be associated with the prosodic domain containing the phonological component of the head of the structure they c-select, they are structurally outside the domain they must be contained in. If we assume a bottom-up derivation of phonological structure, reflecting the bottom-up derivation of syntactic structure, it is possible to argue that an element that is structurally outside its association domain can, and therefore must, be handled immediately. The phonology thus attaches it to its host before the next suffix is considered.

For elements that have a larger domain, things are different. Clause-level special clitics have the IntP as their association domain, which means that when they are encountered (again assuming a bottom-up process of phonological composition), their association domain has not been established yet. By the time it has (i.e., when the CP is mapped onto phonology), all clause-level special clitics are still awaiting placement. It is possible, depending on the details of the relevant process, that they are then placed in hierarchical order.

Let us take a closer look at the Serbo-Croatian data. As a start, consider (52), which shows that the clitic cluster can appear after the first PWd of the clause:[36]

(52) Taj **mi je** pesnik napisao knjigu.
that me AUX poet written book
'That poet wrote me a book.' (Schütze 1994, p. 377)

In this example, the clitic cluster *mi je* breaks up the subject NP *taj pesnik* 'that poet'. The initial impression is that the clitic placement in (52) can be accounted for in prosodic terms, and Schütze (1994) argues that it should be. There are indications, however, that a prosodic account may not be sufficient, or even necessary.

[35] Baker formulates the Mirror Principle as follows: "Morphological derivations must directly reflect syntactic derivations (and vice versa)".

[36] The relevant clitics are marked in boldface throughout.

First, as noted by Diesing, Đurđević, and Zec (2009, p. 62), structures such as in (52) are marked: they only occur in contexts in which "contrastive emphasis on [*taj*] 'that' is appropriate". Similarly, Bošković (2012) notes that the material before the clitic cluster can always be moved syntactically (examples are given below), which means that a syntactic analysis should be possible.

Furthermore, the right edge of the first PWd is not the only position in which the clitic cluster can appear. There is an alternative placement, which has the clitics follow the first constituent of the clause:

(53) Taj pesnik **mi je** napisao knjigu.
that poet me AUX written book
'That poet wrote me a book.' (Schütze 1994, p. 377)

In this example, the clitic cluster *mi je* does not appear after the first PWd *taj* 'that' as in (52), but after the first constituent *taj pesnik* 'that poet'. Schütze uses the terms 1W (first word) and 1C (first constituent) placement in order to distinguish between the two options, which I adopt here. Crucially, there are cases in which 1C placement is obligatory, at least for many speakers.[37] The relevant cases are multi-word proper names, conjoined NPs, post-head genitives and post-head PPs:

(54) a. %Lav **je** Tolstoj veliki ruski pisac.
leo AUX Tolstoy great Russian writer
'Leo Tolstoy is a great Russian writer.'

b. %Sestra **će** i njen muž doci u utorak.
sister will and her husband come in Tuesday.
'My sister and her husband will come on Tuesday'

c. %Prijatelji **su** moje sestre upravo stigli.
friends have my.GEN sister.GEN just arrived
'My sister's friends have just arrived.'

d. %Studenti **su** iz Beograda upravo stigli.
students have from Belgrade just arrived
'Students from Belgrade have just arrived.'
(Halpern 1992, pp. 94–95)

[37] Schütze (1994, p. 66) points out that there is much inter-speaker variation regarding the acceptability of the structures in (54), which may be due to regional, dialectal or stylistic differences. Here, I label all structures with %, following Schütze, indicating their marked status.

In each case, placing the clitics after the NP renders the sentences grammatical. These examples are problematic for a purely phonological account, because such an account would predict that the clitic cluster can occur in all the positions indicated in (54) without triggering any markedness effect.

There are also cases of obligatory 1W placement, and these show quite convincingly that 1C placement must be mediated by syntax:

(55) a. Jako **mi je** dosadna njegova posljednja knjiga.
very me AUX boring his last book
'His last book is very boring to me.'

b. *Jako dosadna **mi je** njegova posljednja knjiga.

(Browne 1975, p. 118)

In (55), the clitic cluster *mi je* must be placed after the first (prosodic) word. Crucially, *jako dosadna* 'very boring' forms a constituent in the clause, since it is the predicate of the copula *je*. As such, it likely constitutes a separate p-phrase in phonology. Therefore, if 1C placement were phonological (e.g., place the clitic cluster after the first p-phrase), we would expect (55b) to be grammatical: it has the clitic cluster immediately following the first p-phrase of the clause. Since it is ungrammatical, we must conclude that phonology is not responsible for 1C placement.

Note that the clause in (55) starts with the predicate. Diesing, Đurđević, and Zec (2009) argue that 1W placement is strongly preferred for clauses that are predicate-initial, while 1C placement is the normal placement for clauses whose first constituent is an argument.[38] This means that in their assessment, (55b) is not ungrammatical, contrary to Browne (1975). Regardless of the grammaticality of this example, however, the simple fact that clause structure affects the preferred placement of the clitic cluster (1C for argument-initial clauses and 1W for predicate-initial clauses), shows that a purely prosodic account is not tenable.

Schütze's (1994) account for the contrast in (55) is based on the assumption that the clitics move to C° and that they can be supported by an element in Spec,CP. If Spec,CP is empty, the clitics must undergo *Prosodic Inversion*, a process proposed by Halpern (1992), which inverts the clitic cluster and the first PWd, yielding 1W placement. In (55), Spec,CP is empty because the predicate cannot be in Spec,CP, on the assumption that Spec,CP is a topicalisation position in Serbo-Croatian and a predicate cannot be topic. By contrast, the NPs in (54) are all in Spec,CP.

The gist of Schütze's model is the following:

- Clitics move to C° (for reasons that are still unclear).

[38] Diesing, Đurđević, and Zec say nothing about clauses that start with an adjunct.

- C° may already be occupied by the verb, in which case the clitics adjoin to it; in phonology, the clitics take the verb as their host.
- If C° is empty, Spec,CP may be filled, in which case the clitic can take the last PWd of the constituent in Spec,CP as its host.
- If neither situation occurs, phonology can rescue the structure by applying Prosodic Inversion. The goal of this operation is to provide the clitic cluster with a phonological host.

The analysis is not without its problems, since the clitic placement options are more varied than the model predicts.

Bošković (1995, 2001, 2012) argues strongly against any kind of PF involvement in the placement of clitics and he rejects the notion of Prosodic Inversion. His argument, following Franks (1997), is that only elements that can be shown to be movable independently in syntax can host clitics. That is, when the clitic cluster appears in 1W position, the PWd before the cluster can actually be moved in syntax, even if it is not a syntactic constituent. Serbo-Croatian is a language with very free word order and it readily allows splitting up of constituents:

(56) U veliku Jovan ulazi sobu.
in big Jovan enters room.
'Jovan enters the big room.' (Schütze 1994, p. 400)

In (56), the constituent *u veliku sobu* 'into the big room' is split up by the subject and the verb. Since the constituent is not split up by a clitic cluster but by phonologically non-reduced material, the operation that yields structures such as (56) must be syntactic.[39] Bošković now argues that whenever a clitic cluster appears in 1W position, the element preceding the clicic cluster has actually been moved in syntax. In particular, he cites facts about split names to support this argument. Consider (57):

(57) a. Lava Tolstoja čitam.
Leo.ACC Tolstoy.ACC I.read
'Leo Tolstoy, I read.'

b. Lav Tolstoja čitam.
Leo Tolstoy.ACC I.read
'Leo Tolstoy, I read.' (Bošković 2012, p. 48)

[39]It is interesting that the preposition is moved along with the adjective: the moved structure is not a syntactic unit, but it *is* a phonological unit, given that (monosyllabic) prepositions such as *u* do not constitute separate PWds. Cf. also (33b) for a similar Russian example.

Complex names can be case-marked either on both the first and last name (57a), or just on the last name (57b).[40] Interestingly, only if the first name is case-marked can it be split from the last name, as in (58a). If the first name is not case-marked, splitting it from the last name is ungrammatical, as (58b) shows:[41]

(58) a. Lava čitam Tolstoja.
Leo.ACC I.read Tolstoy.ACC
'Leo Tolstoy, I read.'

b. *Lav čitam Tolstoja.
Leo I.read Tolstoy.ACC
'Leo Tolstoy, I read.'

Crucially, clitics can also split up a proper name (at least for some speakers), but again, only if the first part is case-marked:

(59) a. Lava sam Tolstoja čitala.
Leo.ACC am Tolstoy.ACC read
'Leo Tolstoy, I read'

b. *Lav sam Tolstoja čitala.
Leo am Tolstoy.ACC read
'Leo Tolstoy, I read'

These and similar examples, Bošković claims, show that even the 1W placement is syntactic and no PF process is necessary in order to account for it.

Bošković makes another claim: contrary to what Schütze (1994) and a number of other authors claim (see both Schütze 1994 and Bošković 2012 for references), the clitics do *not* all move to C°, or even to a single position somewhere high in the tree. In Bošković's analysis, there is no clitic cluster in syntax. Rather, each clitic occupies a specific position in the tree in the functional domain above VP, with no other material intervening. One fact that supports this analysis (Bošković 2012 provides several more) is the observation that adverbs that can have both a subject-oriented and a manner reading disallow the subject-oriented reading when they follow *some* clitics, but not all:

[40] Bošković glosses *Lav* in (57b) as nominative. I would argue that the form is not marked for case.

[41] Note that the judgements below contrast to a certain extent with those in (54a) above. But as remarked in footnote 37, splitting of proper names, although marked, is not universally ungrammatical.

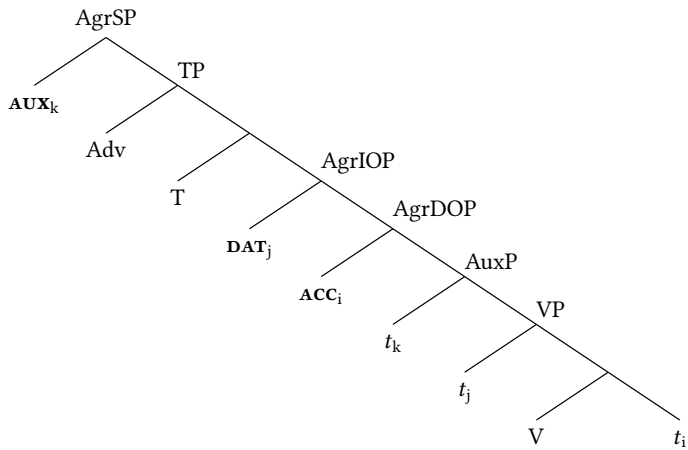
(60) a. Oni **su** pravilno odgovorili Mileni.
they AUX correctly answered Milena.DAT
'They did the right thing in answering Milena.'
'They gave Milena a correct answer.'

b. Oni **su** **joj** pravilno odgovorili.
they AUX her correctly answered
*'They did the right thing in answering her.'
'They gave her a correct answer.' (Bošković 2012, p. 56)

On the assumption that subject-oriented adverbs attach to TP while manner adverbs attach to VP, the judgements in (60) can only be accounted for if we assume that the dative clitic *joj* 'her' occurs in a lower position than the 3pl auxiliary *su*: if only *su* appears in the clause, the adverb *pravilno* 'correctly' can have a subject-oriented reading, but when *joj* 'her' appears in the clause, this subject-oriented reading is lost. If *joj* had moved to C° (or a position close to it), this fact could not be explained: *pravilno* would have the option of adjoining to TP, yielding the subject-oriented reading. But if *joj* is in a lower position, it forces *pravilno* to be in a lower position as well, with the result that the subject-oriented reading is no longer available.

Bošković (2012, pp. 61, 68) proposes the following structure for the Serbo-Croatian clause, with the clitic positions indicated by boldface AUX, DAT and ACC:

(61)



Bošković is not entirely clear on the question whether the clitics are in head or in specifier positions, but since the relevant projections cannot contain any other overt material except for the clitic, we can ignore this question here. As

indicated, Bošković assumes that the argument clitics move from positions inside VP to agreement positions in AgrDOP and AgrIOP. Similarly, the auxiliary clitics move from a position labelled Aux°, which is essentially a V° that c-selects another VP, to the subject agreement position.

In Bošković's analysis, then, the clitic cluster is to a certain extent an accidental occurrence: in syntax, there is no cluster. The order of the clitics in the cluster therefore directly reflects their order in syntax, and phonology plays no direct role in clitic placement. This does not mean that phonology is not involved, however. Clitics are subject to two prosodic requirements in Bošković's analysis: they must appear as closely as possible to the left edge of the intonational phrase in which they are contained, and furthermore, since they are clitics, they cannot appear in first position. The combination of these two constraints means that clitics must appear in second position.

Crucially, it is up to syntax to establish a structure in which both these requirements can be met. Since in Bošković's analysis, PF cannot displace clitics, syntax must ensure that they appear in the right position and that there is some overt material preceding them inside the IntP, so that they have a host. Syntax, of course, cannot fulfil these requirements directly. All it can do is generate a structure that conforms to its own requirements. It is then up to phonology to filter out any structures that would lead to a violation of any prosodic requirement. For example, a structure that has both *pravilno* 'correctly' in the TP-adjoined adverb position and the dative clitic *joj* in AgrIOP would violate the requirement that *joj* is as closely as possible to the left edge of the IntP. As a result, the structure would be filtered out at PF.

The clitic *je* is problematic in Bošković's account, for obvious reasons: being an auxiliary it is positioned high in the clause and should therefore appear first in the clitic cluster. At first sight, the structure in (61) appears to offer an elegant explanation for this problem. Note that *je* is actually the stem of the copula *jesam*, so one could argue that it does not move out of its base position in Aux°, because it lacks a subject agreement morpheme. Bošković (2012, p. 64) shows, however, that *je* behaves identically to the other auxiliary forms with respect to the tests that he uses to determine the clitic's position. For example, it allows a subject-oriented reading for *pravilno*, just as the 3rd plural copula *su* does:

(62) a. Jovan **je** pravilno odgovorio Mileni.
Jovan AUX correctly answered Milena.DAT
'Jovan did the right thing in answering Milena.'
'Jovan gave Milena a correct answer.'

b. On **joj** **je** pravilno odgovorio.
he her AUX correctly answered.
*'He did the right thing in answering her.'
'He gave her a correct answer.' (Bošković 2012, p. 60)

Just as *su* in (60) above, when *je* appears as the only clitic, as in (62a), *pravilno* can have a subject-oriented reading. The clitic must therefore be in the same high position that *su* occupies in (60a) above. When the clitic cluster also contains the dative clitic *joj*, the subject-oriented reading is lost. Bošković therefore assumes that *joj* in (62b) is in the same AgrIOP position as in (60b), forcing *pravilno* to be in the lower, VP-adjoined position, where only the manner reading is available.

(62a) and (b) therefore provide conflicting information with regard to the position of *je*. When *je* is the only clitic, it is in a high position, but when it is combined with other clitics, it is in a low position. Bošković solves this conflict by adopting Chomsky's (1993) *copy theory of movement*, by which a moved element is remerged in the tree, rather than leaving behind a trace.[42] When PF finds multiple copies of a single element in the tree,[43] generally only the highest copy is spelt out. Under specific circumstances, however, it is possible to spell out a lower copy, especially when spelling out the highest copy would violate some restriction. Bošković, following Franks (1998, 2000), argues that this is what happens in the case of *je*. A PF requirement of *je* forces it to appear at the right edge of the clitic group. If the highest copy in ArgSP were to be spelt out in (62b), this requirement would be violated, because *joj* would appear after *je*. Therefore, in such cases, PF has the option to spell out the lowest copy of *je*, which sits in Aux°, just above VP.

Bošković's analysis of *je* clearly depends on theory-internal considerations and therefore cannot be adopted wholesale into the current model. The general analysis and the argument that clitics must occupy different positions in the tree seem basically sound, however. In fact, following Bošković's account, we can provide fairly straightforward explanations for the placement of the clitics.[44] Firstly, all clitics except *je* have a very simple lexical entry without any prosodic requirement. Here, I use the 3pl copula *su* as an example, but the other clitics are similar:

[42] As Chomsky (2008, p. 158, fn. 17) notes, the term 'copy' in the copy theory of movement has lead to some confusion in the literature. Essentially, what is meant is that the original element is merged at a second position in the tree, so-called *remerge*, not that an actual copy is made.

[43] Or, more correctly, finds that an element has been merged in more than one location.

[44] I remain agnostic on the question whether the clitics are base-generated in the positions in which they appear, or whether they are moved there, as Bošković assumes. Since Serbo-Croatian does not have clitic-doubling, the pronominal clitics seem to be genuine arguments, which suggests a movement (long-distance dependency) account.

(63) $\begin{bmatrix} \text{AUX} \\ \text{3pl} \end{bmatrix} \leftrightarrow$ /su/

The crucial point of the phonological component of (63) is that the auxiliary is simply mapped onto the string /su/ and does not have any specific prosodic requirement. Because of this, it is not an autosegment and therefore does not have the ability to appear anywhere other than where the syntax puts it. However, *su* does not constitute a separate PWd, so it is phonologically deficient and requires a PWd to cliticise onto. On the assumption that the direction of Stray Adjunction in Serbo-Croatian is leftward, *su* requires a host to its left. It is up to syntax to provide this host. If it is not available, the structure is ungrammatical, because clitics in Serbo-Croatian (apparently) do not have the option to procliticise onto the following word.

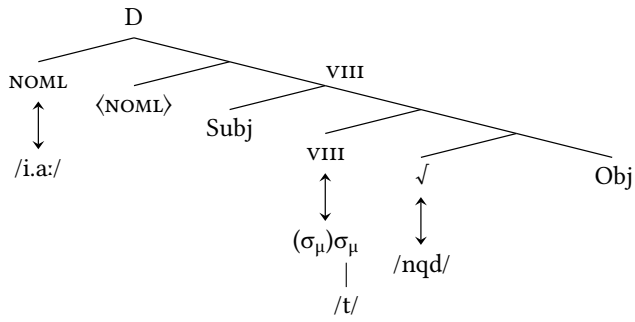
The one exception is of course *je*. Within the current framework, however, it is not necessary to assume that PF can spell out the lower copy of the clitic (and consequently no need to assume that there *is* a lower copy). Rather, we can instead assume that *je* has the following lexical entry:

(64) $\begin{bmatrix} \text{AUX} \\ \text{3sg} \end{bmatrix} \leftrightarrow \text{je}|_{\omega}$

That is, *je* has a phonological form similar to the suffix *-ing* discussed in section 3.3.2: it consists of a syllable with the additional prosodic requirement that it aligns with the right edge of a PWd.

Note that in Bošković's analysis, the auxiliaries are V° heads that c-select another V. Under this analysis, we would expect *je* to attach to V, since Input Correspondence requires a head to take the head of the structure it c-selects as host in phonology. However, the auxiliary has moved to AgrS, which selects T. We have seen a similar case in section 3.2.5, where I assume that the head N moves to D in Arabic verbal nouns. The relevant example is repeated here:

(65)



In this example, I assume that the nominaliser NOML moves to D in order to provide D with an index. NOML maps onto a prosodic morpheme and since it selects its complement, the stem VIII marker, Φ(NOML) must take Φ(VIII) as its phonological host. Movement of NOML to D does not affect this, because D selects the (moved) NOML, which in turn selects VIII.

The situation with *je* is different, however, since AgrS, the landing site of *je* according to Bošković's analysis, selects T and T itself is not a prosodic morpheme. In fact, it is not even clear what the phonological component of T is, but in the case of *je*, it seems safe to assume that it is empty. In a way, then, *je* has moved too far away from the element it c-selects and we may assume that because of this, it is no longer possible for it to take Φ(V) as its phonological host.

What exactly the correct definition of *too far away* is, is an open question, but the case of *je* offers several options. For one, it could be the case that movement of any kind blocks Input Correspondence from applying. This would mean that the NOML head in the Arabic verbal noun does not actually move to D but that instead D and NOML form a combined head of some sort.[45] Another possibility is that because *je* has moved, Input Correspondence now applies to its landing site, AgrS. AgrS c-selects T, but T has no phonological form, so that *je* cannot take it as its host. And since Φ(T) is empty and therefore not by itself subject to Input Correspondence, it is not possible to, so to speak, pass on the hosting responsibility, in the sense that *je* must attach to whatever Φ(T) must attach to. As a result, *je* is not able to attach to the verb and must attach to whatever is available. Given the position of *je* in the tree, this is the word preceding the clitic cluster plus the clitic cluster itself, which incorporates into its host's PWd. *je*, due to its phonological form, must appear at the right edge of this PWd.

The 3sg copula *je* is involved in another phenomenon that is often accounted for in terms of templatic morphology: co-occurrence restrictions. When *je* co-occurs with the 3sg feminine accusative clitic, which is also *je*, the 3sg feminine accusative clitic shows up as *ju*. Furthermore, the combination of reflexive *se* plus copula *je* usually surfaces as *se* (that is, *je* is dropped). Schütze (1994, p. 420) mentions that colloquially one also finds reductions of the combination of the 1st and 2nd sg accusative clitics (*me* and *te*) plus *je* to just *me* and *te*.

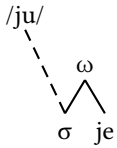
The change from *je* to *ju* for the 3sg feminine accusative clitic could be captured with a phonological rule along the lines of (66):

(66) /e/ → [u] / j___je

[45] Although this would of course raise the question again why the Arabic verbal noun appears phrase-initially. See section 3.2.5 for discussion.

However, this rule would predict that *every* /e/ is realised as [u] in the relevant context. In other words, it is not possible to restrict the rule to the 3sg feminine accusative clitic. To capture this fact, we need a way to refer to the syntactic context as well. For this purpose, I propose a pair of lexical mapping rules for the 3sg feminine accusative clitic, one for the general case and one for the form *ju* in the context of *je*:

(67) a. $\begin{bmatrix} \text{3sgf} \\ \text{ACC} \end{bmatrix} \leftrightarrow$ /je/

b. $\begin{bmatrix} \text{3sgf} \\ \text{ACC} \end{bmatrix} \leftrightarrow$ 

The mapping rule in (67b) maps the 3sg feminine accusative clitic onto the string /ju/ associated with a syllable inside a PWd that ends in /je/. The rule does not explicitly state that this /je/ must be the 3sg copula, because it is not clear how that could be done. Although it is in principle possible to include syntactic context in the mapping rule, the relevant head in this case is the Aux° head, which sits higher in the tree and may be separated from the accusative clitic by the dative clitic.

The prosodic context, however, should generally assure that *je* does not surface as *ju* preceding any other occurrence of the syllable *je*, because it specifies that /ju/ must be followed by /je/ within the same PWd. Therefore, even if the clitic cluster ends with the 3sg feminine clitic *je* and the following word starts with this syllable, the clitic should still surface as *je*.

The fact that the auxiliary *je* is dropped after the reflexive *se* can be captured in a similar way. Since Browne (1974) notes that when *je* is dropped after *se*, the vowel of *se* may be lengthened, I assume that the mapping rule actually insert the vowel /e/:

(68) $\begin{bmatrix} \text{AUX} \\ \text{3sg} \end{bmatrix} \leftrightarrow$ 

A purely phonological rule then reduces the two /e/ to a long /e:/ or even to single short /e/.

Summarising, the discussion suggests that the prosodic syntax model is in principle able to account for clitic cluster ordering without resorting to the use of a separate morphology module. In Serbo-Croatian, clitic ordering is primarily syntactic, which means that there is indeed no need to assume a morphology module that orders the clitics after they have been moved into a single position in syntax. Deviations from the syntactic order, i.e., cases such as the 3sg copula *je*, and co-occurrence restrictions can be handled through lexical mapping rules. Obviously, the data and the analysis discussed here do not prove that every kind of clitic cluster can be handled without recourse to a morphology module. To the extent that the Serbo-Croatian clitic cluster is typical of templatic morphology, the analysis does suggest, however, that a morphology module is not strictly necessary in order to account for templatic phenomena. A combination of syntactic ordering and lexical mapping rules seems sufficient.

4.7 Concluding remarks

In this chapter, I argue that the grammar of human language does not include a separate morphology module. Rather, there is only a single module responsible for the linking between form (phonology) and meaning (semantics), which I call "syntax". The empirical differences between what are traditionally called syntactic and morphological structures can be explained if we take into account the contribution that the phonological module makes to the construction of linguistic forms. In essence, the syntax/morphology distinction is an acoustic illusion: a syntactic structure that is mapped onto a prosodic word is generally considered to be "morphological", while a syntactic structure that is mapped onto phonological phrases is considered "syntactic". In reality, however, the structure-building mechanism is the same.

There is also a slightly different way of looking at this. Language users store associations between form and meaning, as represented by the format of lexical entries adopted here (repeated from example (8) in chapter 2):

(69) $\lambda x(\mathbf{car}(x)) \leftrightarrow \begin{bmatrix} \text{N, sg} \\ \text{count} \end{bmatrix} \leftrightarrow /\text{kaɪ}/$

The point is that it is immaterial whether the phonological component of such stored associations is a PWd, a syllable, or a segment. We have seen examples of such cases in this and the previous chapter. We may tentatively assume that a language-learning child actively seeks out such form-meaning pairings. Subsequently, the brain constructs generalisations over these form-meaning pairs. It does not stop there, however: the brain then constructs generalisations over

those generalisations (cf. O'Reilly et al. 2012). What we call "morphology" may be thought of as the body of lower-order generalisations, often still with references to lexical items. "Syntax", then, is the body of higher-order generalisations, often with any reference to specific lexical items generalised away. *Morphology by itself*, in this view, is the result of generalisations over forms alone, without a meaning component.

Crucially, these generalisations are expressed in our model as the structure "in the middle", i.e., as the syntactic feature bundles that stand between the semantic and the phonological component in lexical entries such as (69). Clearly, it is immaterial whether these generalisations are lower-level or higher-level generalisations, at least for the formalisation. In either case they are formalised in the same way. In other words, there is no discrete distinction between morphological and syntactic generalisations. They are part of a single system.

5
Competition-based interaction

The term 'competition' is not uncommon in linguistic theories. Morphological theories often employ some form of competition to describe cases where two morphs could in principle fill a specific slot but only one actually does. Optimality Theory generalises the idea of competition, arguing that all linguistic forms are the result of competition between different possible candidates.

In this chapter, I look at some forms of competition from the perspective of the prosodic syntax model discussed in the preceding chapters. It is essentially an exploratory chapter and therefore much of what is said is speculative, but the general idea is that with only a few extensions, the model can be used to describe competition in a straightforward and simple way, without making the theory more complex and more difficult to implement in a lower-level model.

The chapter starts by looking at the Elsewhere Condition, discussing what it would take to implement it in as simple a manner as possible. I then reprise the Serbo-Croatian clitic cluster in an attempt to tie up a few loose ends in the discussion in the previous chapter. I end the chapter with a discussion of a typical OT-style model and how it could be implemented in the model sketched here.

As stated, the discussion in this chapter is speculative and will end up raising more questions than it answers. Nonetheless, I believe the basic idea is worth exploring further.

5.1 The Elsewhere Condition

Sections 1.2 and 4.6.3 made reference to the *Elsewhere Condition*, the principle that states that the application of a specific rule blocks the application of a later, more general rule. This rule is invoked whenever two (or more) rules compete

due to the fact that there are contexts in which both potentially apply. The examples discussed involve English past tense and past participle formation. Here, I focus on the latter formation, repeating the rules from section 4.6.3, but the analysis of past tense is essentially the same, given that the formation of both verb forms is very similar in English.

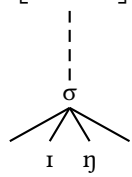
The regular past participle is formed by attaching the suffix *-ed* to the stem of the verb. This suffix has three phonologically conditioned allomorphs, /d/, /t/ and /ɪd/, of which /d/ is usually considered the underlying form. The regular past participle formation can then be captured with a lexical mapping rule of the form in (1):

(1) $\begin{bmatrix} \text{ppp} \\ u\text{V} \end{bmatrix} \leftrightarrow /\text{d}/|_{\omega}$

This rule simply states that the syntactic features [ppp, *u*V] correspond to a suffix /d/. Postlexical phonological rules then adjust this form if necessary.

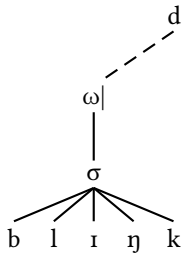
As discussed, the subregularity that changes verbs such as *ring* into *rung* and *drink* into *drunk* is captured by the following rule:

(2) $\begin{bmatrix} \text{ppp} \\ u\text{V} \end{bmatrix} \leftrightarrow \begin{bmatrix} \text{+low} \\ \text{+central} \end{bmatrix}$



This rule adds the features [+low, +central] to the syllable containing /ɪŋ/, effectively changing the vowel /ɪ/ to /ʌ/. Syntactically, the two rules in (1) and (2) apply in the same context: both introduce a [ppp] feature and both select a verbal complement. Phonologically, however, there is a difference: (1) places no requirements on the phonological context, while (2) does: it only applies to words that contain the segmental sequence /ɪŋ/.

A third rule vies for the same syntactic context as the two previous ones:

(3) $\begin{bmatrix} \text{ppp} \\ u\text{V} \end{bmatrix} \leftrightarrow$ 

Like the two rules above, the rule in (3) applies to syntactic contexts that introduce a [ppp] feature and select for a V head. The phonological component of (3) is even more specific than in (2), however: it requires the syllable /blɪŋk/, to which the suffix /d/ is added. The word *blink*, although meeting the context for the ɪ~ʌ ablaut rule, forms its past tense with the "regular" suffix /d/. As discussed in section 1.2.4, we need a specific rule to encode this fact.

The increasing specificity of the phonological context is the property that orders the rules with respect to each other. Although this principle in itself is fairly straightforward, the question is how we can implement it in the grammar model. The simplest way would be to say that the grammar contains a rule that mandates the Elsewhere Condition. This, however, is not a very attractive solution, because it requires much additional machinery that complicates the grammar considerably.

If the Elsewhere Condition were a grammar rule, the grammar would have to be able to inspect and compare lexical mapping rules. It would need the ability to consider the feature content of the three rules in (1)–(3) and establish that the second of these is more specific than the first and the third is more specific than the second. This requires that the grammar has an additional module for inspecting and comparing rules. Furthermore, it requires that rules such as those in (1)–(3) are represented *explicitly* in the grammar, meaning that they are available to the grammar as *data*. That is, the rule is not only available to the system as an operation or test that can be applied to certain data, but also as data itself, to which other operations can be applied.

Representing a rule explicitly in the grammar, even lexical mapping rules of the kind assumed here, is something to be avoided if possible, because it complicates the grammar, or more specifically, it complicates the implementation of the grammar at a lower level of analysis. To see why this is the case, it is useful to look at the implementation of a much simpler system, because that makes it easier to see what is required of such an implementation. I will use the Boolean relation OR as an example. The entire "grammar" of Boolean OR is given in the truth table in (4):

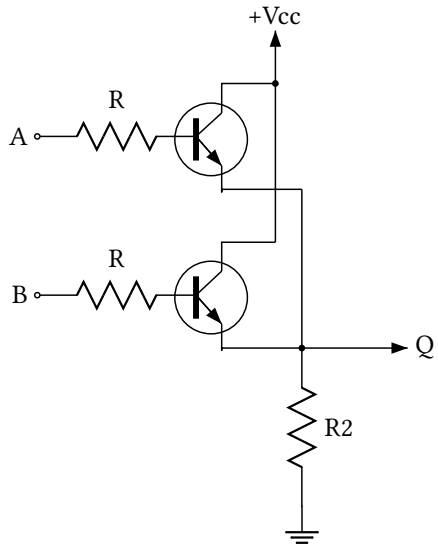


Figure 5.1: Boolean OR circuit

(4)

A	B	Q
0	0	0
1	0	1
0	1	1
1	1	1

The table in (4) can be summarised as the rule in (5):

(5) *If either A or B is 1, then Q is 1.*

Boolean OR can be implemented in a logic circuit such as the one displayed in figure 5.1.[1] The way this circuit implements the rule for Boolean OR can be loosely described as follows: if a current flows from either A or B, the corresponding transistor (the circled structures) causes a large current to flow from the power supply voltage (+Vcc) to the output line Q. Crucially, the circuit is set up in such a way that a current on *either* A or B is sufficient to let a current flow from Q.

The circuit in figure 5.1 implements the Boolean OR rule in (5), but it encodes it *implicitly*. The rule is not available as data to the system. The only data that

[1] See `http://www.electronics-tutorials.ws/logic/logic_3.html`.

the system handles are the input currents on A and B and arguably the output current on Q.

If we would want to construct a logic circuit that has the ability to, say, compare Boolean OR and Boolean AND, we would need a more complex system, one that has the ability to place identical inputs on an OR gate and an AND gate and compare the results. Note, however, that the circuitry to do this would undoubtedly contain a number of AND and OR gates beyond the two that it needs for testing. In such a setup, the fact that the system is applying the very same rules that it is testing is *not* a piece of information that is available to the system itself. Making this information available to the system is not a theoretical impossibility, but obviously requires an even more complex system.[2]

Moreover, it requires a system that has an internal representation of the rules it employs, the way they are put together, and the means to manipulate these representations. At this point, the system has reached the complexity of an electronic computer, which can indeed do these things. But it should be obvious that if we want to implement the Boolean OR rule, we do not need a computer. The circuit in (5) will do.

The point of this analogy is of course not to claim that the brain is similar to a computer or to a complex circuit board. It mainly serves to make clear that a system that *implements* a certain rule does not need to have access to the *content* of that rule, and furthermore, that being able to inspect and otherwise manipulate rules requires that the relevant rules are represented in the system as information that the system can process (i.e., as *data*). Systems that have these introspective abilities are by necessity more complex than systems that lack it.

In the same way, a grammar model that implements lexical mapping rules and higher-order rules such as the Elsewhere Condition implicitly will be simpler than a grammar model that treats rules as data. Note that in principle, three types of grammar model are possible, with three different levels of complexity. The most complex one represents both lexical mapping rules and higher-order rules such as the Elsewhere Condition explicitly. A system that encodes higher-order rules implicitly but lexical mapping rules explicitly is less complex, but still more complex than a system that encodes both types of rules implicitly.

If we wish to eliminate the Elsewhere Condition as an explicit rule of grammar, we need another way of implementing its effects. To see how this might be done, consider the three lexical mapping rules above, represented here in simplified form:

[2] There does seem to be a limit to this kind of introspection, however. The circuits implementing the introspection are not available for introspection themselves.

(6) *English past participle formation* (partial)

(I) Default rule: add /d/ to the stem.

(II) ɪ~ʌ ablaut: add [+low, +central] in the context of /ɪŋ/.

(III) *blink*-rule: add /d/ to /blɪŋk/.

The rules are numbered in order of specificity, with rule (III) being the most specific. The Elsewhere Principle states that rule (III) blocks rule (I) and (II), while rule (II) blocks rule (I). We can encode these blocking relations directly into the grammar in the following way:

(7) a. (III) $\not\rightarrow$ (I)
(III) $\not\rightarrow$ (II)
b. (II) $\not\rightarrow$ (I)

These blocking relations should be read as follows: if rule (III) applies, rule (I) and rule (II) do not apply. Note that such blocking relations are not transitive: although rule (III) blocks rule (II) and rule (II) blocks rule (I), it is not the case that rule (III) blocks rule (I) through rule (II). The fact that rule (III) blocks rule (I) has to be recorded separately. The reason is that if a rule X has a blocking relation with a rule Y, X only blocks Y when X is active. In the current case, if rule (III) is active (because its context matches), it blocks rule (II). Since rule (II) is therefore not active, it cannot block rule (I). Instead, rule (III) needs to block rule (I) directly.

Implementing the Elsewhere Condition in this manner throughout the grammar obviously entails adding a large number of such blocking (or inhibitory) relations to the model. Such an approach seems to fly in the face of established methodology, which holds that a theoretical model should capture the generalisations that underlie the domain it aims to describe. Encoding the effects of the Elsewhere Condition directly into the model in the form of inhibitory relations between rules seems to do the exact opposite.

The reason behind the idea that a theoretical model should capture appropriate generalisations is that the model must be kept maximally simple. The Elsewhere Condition regulates the behaviour of morphemes that target the same context, and a grammar model that does this by means of a single rule for any set of competing morphemes is simpler (in the sense that it has to make fewer assumptions) than a model that has to state the condition for each set of competing morphemes separately.

It is important to note, however, that this simplicity exists at the level of analysis at which the model is located, which in this particular case is at a relatively

high level. As the discussion above should make clear, it is by no means a certainty that this simplicity can be maintained when the model is implemented at a lower level of analysis. In the case of the Elsewhere Principle, it seems unlikely that it can.

For this reason, I prefer to implement the Elsewhere Condition as a set of inhibitory relations, as just outlined. Instead of being a principle of the grammar, the Elsewhere Condition becomes a principle that guides the construction of the model. It is still an important principle, but it is what we might call a meta-principle: a principle that the grammar adheres to implicitly, without being required to do so by an explicit rule.[3]

Adopting the notion of inhibitory connections between lexical mapping rules obviously raises the question whether the opposite, excitatory connections, are possible and indeed necessary as well in our model. It should probably come as no surprise that the answer to this question is yes. In fact, the model already *has* such connections: each lexical mapping rule is in fact nothing more than a set of (mutual) excitatory connections. Take, for example, the lexical item *car* in (8):

(8) $\lambda x(\mathbf{car}(x)) \leftrightarrow \begin{bmatrix} \text{N, sg} \\ \text{count} \end{bmatrix} \leftrightarrow /\text{ka}\text{ɹ}/$

Basically, this formula states that if the phonological structure /kaɹ/ is activated, the syntactic structure [N, sg, count] and the semantic structure $\lambda x(\mathbf{car}(x))$ also become active. Similarly, the semantic structure $\lambda x(\mathbf{car}(x))$ will activate the syntactic structure [N, sg, count] and the phonological structure /kaɹ/.

The representation in (8) suggests that syntax mediates between semantics and phonology. In fact, the representation is based on this assumption. However, if we interpret the arrows as excitatory relations, it becomes clear that this cannot be the whole story. After all, many nouns have the syntactic structure [N, sg, count], which means that this syntactic structure has connections to many phonological and semantic representations, not just to the representations given in (8).

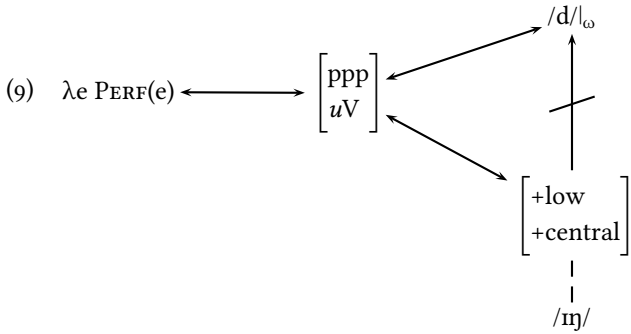
If /kaɹ/ would just activate [N, sg, count], there is no way for the system to know which semantic representation to activate. We are therefore forced to assume that in addition to the two mutual excitatory relations in (8), there is an

[3] The proposed analysis of course raises the question why grammars adhere to such meta-principles. The answer to this question probably depends on the meta-principle. For example, in the case of the Elsewhere Condition, it may be a basic principle of a self-organising system: two elements targeting the same context must necessarily conflict and this conflict needs to be resolved. The NoCoda constraint, which I argued in section 3.3.2 is a similar meta-constraint, possibly results from architectural constraints of the brain.

additional mutual excitatory relation between λ*x*(**car**(*x*)) and /kaɹ/. For our purposes, this relation is not relevant and therefore I will continue to use representations such as the one in (8), but it should be noted that the semantic-phonological connection is implied.

Given the fact that lexical entries that have the same syntactic structure actually *share* this structure (i.e., it is not the case that each entry has its own copy), the inhibitory relations in (7) must be made more precise. (7b), for example, states that the ɪ~ʌ ablaut rule inhibits the default rule. But the syntactic structures of both rules are identical, and hence shared. This makes sense, of course, since there is no syntactic distinction between the two structures, but it also means that it cannot be the entire rule (II) that inhibits rule (I). Rather, it must be the phonological component of rule (II) that inhibits the phonological component of rule (I). This can be schematically represented as in (9):

(9)



The semantic representation in (9) is obviously simplified. What is important is that both rule (I) and rule (II) share the same semantic and syntactic representations. They only diverge in phonology, where rule (I) is linked to a /d/ suffix, while rule (II) is linked to the features [+low, +central] in the context of the structure /ɪŋ/. The latter structure has an inhibitory connection to /d/, which means that if the context /ɪŋ/ is present, the suffix /d/ is suppressed.

A quick note on terminology: I use the term *schema* (plural *schemata*) for the chunks of structure in a rule such as (9). That is, the semantic component λe PERF(e) is a schema, the syntactic component [ppp, *u*V] is a schema, and so are the phonological chunks in the rules. Schemata are connected to each other by excitatory and inhibitory connections. I refer to a set of connected schemata such as that represented in (9) as a *rule*. Such rules are essentially part of the lexicon, just as lexical entries (or *lexical mapping rules*) such as in (8) above are part of the lexicon.[4]

[4]The term *schema* is obviously adopted from HPSG, but it should be noted that its meaning here is quite different.

Note that according to the rule in (9), the two phonological schemata /d/ and [+low, +central] are both activated by the syntactic schema [ppp, *u*V]. This seems to result in a conflict, because only one of them can be active: a verb either has the /d/ suffix, or it uses the ɪ~ʌ ablaut, it does not have both. The structure [+low, +central] inhibits /d/, but this obviously cannot happen on the basis of activation from [ppp, *u*V] alone: /d/ should only be inhibited if the context for [+low, +central] is present. This means that the dashed line can also be interpreted as a unidirectional excitatory relation, with the context /ɪŋ/ activating the structure [+low, +central]. This in turn suggests that the excitation from [ppp, *u*V] to [+low, +central] is not as strong as the activation from [ppp, *u*V] to /d/ and that [+low, +central] is only able to inhibit /d/ if it receives excitatory input from both structures.

If these considerations are correct, it appears that excitatory connections can vary in strength and that excitatory inputs can be additive. The same may of course be true of inhibitory connections. I will not pursue this matter here, because more data is needed in order to be able to examine these options in more detail and to determine how best to implement them. For the moment, the dashed line indicating that /ɪŋ/ is the context for [+low, +central] suffices to capture the intention of a rule such as (9), at least informally.

Interestingly, the rule in (9) indicates that what was presented earlier as a collection of three rules, given in (1)–(3), is really a single rule with multiple phonological components (schemata), with specific inhibitory connections between them. Note that the number of phonological schemata is not limited to two or even three, since there are other subregularities in the system and of course other irregular forms.

The rule in (9) essentially contains all the information required for past participle formation in English, and can informally be read as a series of if-then statements: the past participle semantics and syntax trigger (or: are realised by) the suffix /d/ *if* the context /blɪŋk/ is present, otherwise the features [+low, +central] *if* the context /ɪŋ/ is present, otherwise the suffix /d/. Interestingly, the entire past participle formation can be captured in a *single* rule. In particular, there is no need for a dual mechanism in which a general rule is distinguished from listed irregular forms (cf. the discussion in section 1.2.4).

This fact recalls the claim of Westermann and Ruh's (2012) connectionist model of the English past tense, discussed in chapter 1, that the English past tense is best modelled by "a structured single-mechanism, multiple-representation system" (p. 15). It should be noted, however, that the current model and Westermann and Ruh's model are quite different. For one, it would not be correct to say that (9) contains multiple representations, or even dual representations, as Westermann and Ruh argue is the case for their connectionist model. Their model has two representations of each verb, both of which are based on the verbs' phonological

forms. One representation is encoded in the direct input-output connections, the other is encoded in the hidden layer, where (groups of) units each represent a single verb or a range of similar sounding verbs that share the same past tense form. The rule in (9), on the other hand, does not contain *any* representation of the verbs to which it applies, with the exception of the irregular verbs. In this sense, it is more in line with familiar theoretical models, which separate the lexicon from the grammar. Where (9) differs from traditional models is that it encodes the irregular and semi-irregular forms in the rule, rather than mark the relevant verbs in the lexicon.

This difference between the model in (9) and Westermann and Ruh's connectionist model is the result of the fact that both models have different aims. Westermann and Ruh aim to model the *production* of past tense forms, given the verb and given the task of producing the past tense, in a manner that is compatible with a neural architecture. The rule in (9), on the other hand, is not meant as a production model. Its aim is to describe the knowledge of a speaker of English about the formation of the past participle in a manner that is deliberately kept as simple as possible, whereby 'simplicity' is defined from a certain perspective, as discussed above and in chapter 1.

Even though the model in (9) sports excitatory and inhibitory connections, which are typical elements of connectionist models, there is no claim that it constitutes a connectionist model, nor that it is compatible with one, or even that the best way to implement it at a lower level of analysis is through a connectionist model. The only claim behind (9), apart from the fact that it is an attempt to describe a native speaker's knowledge about past participle formation, is that it *should* be simpler to implement the knowledge contained in it at a lower level of analysis (i.e., in a lower-level model) than a model that explicitly incorporates meta-principles.

I have emphasised the word 'should' here, because the claim is not based on a mathematical analysis of the complexity of an analysis with an implicit vs. an analysis with an explicit Elsewhere Condition. Rather, it is based on the assumption that an explicit incorporation of the Elsewhere Condition requires being able to treat rules as data, while an implicit incorporation of the same condition does not, and that treating rules as data adds complexity to the system. Despite the lack of a mathematical foundation, I believe the assumption is a plausible one and I therefore adopt the idea that pieces of structure can have excitatory and inhibitory connections to other pieces of structure. Crucially, such connections are possible between structures in different modules and they can, but do not need to be, mutual. This system allows us to account for a basic form of competition between structures: if two structures compete for the same context, such as the two past participle formations captured in (9), one has the ability to inhibit the other. As discussed, this implements the Elsewhere Condition for the case at

hand.

In the following sections, I take a look at some cases where I believe the type of model outlined here can be beneficial to the analysis. They involve cases where phonological information appears to affect the syntactic structure and cases where competition between rules plays a role. The first example that I discuss is the Serbo-Croatian clitic cluster, which was already discussed in the previous chapter, but which has a phonological component that requires some more attention. After that, I take a look at the German middle field, where various factors compete in establishing the order of constituents.

5.2 Serbo-Croatian clitics reprised

In chapter 4, I looked at the Serbo-Croatian clitic cluster in the discussion of templatic morphology. My aim at that point was to argue that templatic morphology does not provide sufficient reason to assume a separate morphology module, because the order of clitics in the cluster can be described syntactically, with the exception of the 3sg copula *je*. Following arguments from Bošković (2012), I adopted the position that the clitics are ordered in syntax and that there is in fact no *syntactic* clitic cluster at all.

Things are different in phonology, however. In the phonological structure, the clitics are dependent elements, i.e., they do not constitute PWds and need to incorporate into one. In other words, in phonology, the clitics *do* form a cluster, and this fact has consequences for the syntax. For instance, it inhibits adverbs from adjoining to TP, as data from Bošković (2012) shows:

(10) Oni **su** **joj** pravilno odgovorili.
they AUX her correctly answered
*'They did the right thing in answering her.'
'They gave her a correct answer.' (Bošković 2012, p. 56)

This example, repeated from (60b) in chapter 4, shows that when a dative clitic is present in the cluster, here *joj* 'to her', the adverb *pravilno* 'correctly' can only have a manner reading, even though in principle it can also have a subject-oriented reading. On the assumption that the subject-oriented reading is obtained by adjunction to TP and the manner reading by adjunction to VP, (10) shows that the presence of the dative clitic forces the adverb *pravilno* to be adjoined to VP. Adjunction to TP would mean that the adverb intervenes between the copula clitic *su* and the dative clitic *joj*, preventing them from forming a (phonological) cluster.

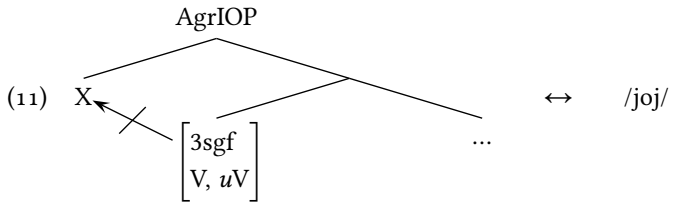
The question is how phonology is able to block TP-adjunction of *pravilno* in this case. Bošković assumes that the clitics are subject to two phonological constraints: they must appear as close to the left edge of the containing IntP as possible, but they cannot appear in the first position of the clause (i.e., they need a host to their left). The latter constraint simply results from the clitics' phonological form: they are single syllables and do not constitute independent PWds. On the assumption that the direction of Stray Adjunction in Serbo-Croatian is leftward, we can account for the observation that the clitics require a host to their left.

It would seem that the first of these two constraints can be implemented by associating the clitics with the left IntP boundary. This is not unproblematic, however, since such an association provides them with an autosegmental property, making them subject to Left-to-Right Association. Within their association domain, the IntP, they would then be associated with the first possible position. Since the clitics are syllables and Stray Adjunction favours leftward adjunction, the first position they can be associated with is the right boundary of the first prosodic word in the IntP. Such an analysis would therefore predict that the Serbo-Croatian clitics always appear in 1W position, which is obviously not correct.

For this reason, the proposal made in section 4.6.4 that all Serbo-Croatian clitics (except *je*) have no prosodic requirement must be maintained. The effect that no overt material is allowed between C° and the lowest overt clitic has to be derived in another manner. The idea that I will pursue here is that the clitics prohibit anything from appearing in a c-commanding position within the maximal projection in which they are contained. In other words, they must be the highest element within their maximal projections.

Although intuitively obvious, we need to be careful about implementing this idea in the current framework. An inhibitory relation with some syntactic structure seems the most straightforward method, but there is no obvious syntactic structure that we can inhibit. Basically, we need to inhibit any syntactic structure that contains syntactic material below Spec,CP that c-commands the clitic. At first sight, it might seem that we can do this by using a generic label such as 'X', which generalises over labels such as N, V, etc. However, even if we were to adopt such a label,[5] it is still not clear what exactly the inhibition rule would have to look like. As a first approach, one might suggest something along the following lines:

[5]In actual fact, I am not an advocate of such a generic syntax label. Labels such an N, V, A, etc. are shorthands for (collections of) properties that the members of there categories share. A generic label would not designate any property except for the property of being syntactic.



Unfortunately, this inhibition does not yield the result that we are after. First, note that it is restricted to the specifier position of AgrIOP, the projection in which *joj* is located in Bošković's (2012) analysis. This means that (11) cannot inhibit the adjunction of *pravilno* to TP. It is pointless to replace the node AgrIOP in (11) with a generic label YP, however, because that would exclude *any* syntactic material from appearing before *joj*, even other clitics. In sum, there does not seem to be a good way to inhibit the appearance of syntactic material in the relevant domain. The domain needs to be larger than the maximal projection of the clitic, in order for dative clitics to block adjunction to TP, but it should be smaller than the projection that houses the next clitic.

Rather than specifying in which kinds of contexts a clitic *cannot* appear, it is easier to specify in which contexts they *can* appear. A clitic must appear in a syntactic context in which the next-higher element in the structure is itself a clitic, or C°. This of course requires that we can identify the relevant clitics. In principle, it would be possible to list every clitic that could precede *joj*, but we would have to do this for every clitic combination, not a very attractive prospect.

The discussion of the Elsewhere Condition above suggests that if there is no other way to implement this idea without complicating the grammar, it is the preferable option. However, in this particular case, there is a better way. In the discussion of the Romance N-pattern in section 4.6.2, I argue that verbs that display an N-pattern alternation are marked with an [NPAT] feature. This is a language-specific feature that is motivated by the fact that N-pattern verbs share a distinct pattern of behaviour that sets them apart from other verbs. There is good reason to assume that this property of verbs is generalised in a speaker's I-language (in the sense of Chomsky 1995b) and should be represented as a feature in a grammar model.

This same reasoning applies to the Serbo-Croatian clausal second-place clitics. The fact that they appear in a cluster in the second position of the clause is a property that they share and that sets them apart from other lexical items. It therefore stands to reason that clitics are marked with a morphosyntactic feature, which we can simply call [CL]. To a certain extent, [CL] is a remarkable feature, because it cuts across categories: there are auxiliaries, arguments and a question particle that have this feature. But because this feature describes a particular morphosyntactic behaviour, we are nonetheless justified in adopting it.

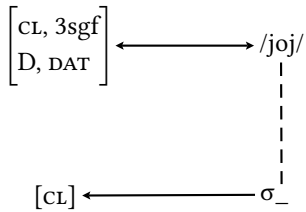
Using this feature, we can identify clitics in syntax. What we need now is a way to identify clitics in the right context. Essentially, a clitic must always appear strictly adjacent to another clitic preceding it, or it must appear strictly adjacent to Φ(C). Strict adjacency translates structurally to immediate c-command. However, if we were to rely on this, we would run into the same complication that we ran into above with the inhibition structure in (11): even though the clitics form a cluster in phonology, they do not do so in syntax.

In other words, even though a clitic must follow either Φ(C) or another clitic in the phonological string, it is not the case that a clitic must be immediately c-commanded by another clitic or by C. As discussed, the argument clitics, in Bošković's (2012) analysis, are in specifier positions of Agr projections, which means that Agr heads may intervene between clitics, and the auxiliary clitics are in AgrS, which means that T intervenes between the auxiliary clitic and the argument clitics, since AgrS is above T, while AgrIO and AgrDO are below it.

Note that adopting Chomsky's (1995) proposal that Agr positions should be abandoned in favour of having T and *v* mediate agreement will not solve the intervention issue. It would work for the argument clitics, because if they are specifiers of *v*, there would no longer be any Agr heads intervening between them. However, T would still create intervention effects: the facts regarding the interpretation of *pravilno* 'correctly' clearly indicate that the auxiliary clitics appear above T, while the argument clitics appear below it.

Crucially, however, even if there are heads intervening between clitic positions, these heads must be phonologically null. This follows from the fact that the clitics form a cluster in phonology. Any intervening, phonologically non-null head would break up this cluster, which would result in ungrammaticality. It is this fact that we must exploit in order to formulate a rule that captures the facts correctly. The rule should state that the phonological component of a clitic should immediately follow a syllable that is itself the phonological component of a clitic. The following lexical mapping rule for *joj* expresses this idea:

(12)



This rule says that the clitic /joj/ should be inserted after a syllable that is associated with a syntactic element that is itself a clitic.

The rule in (12) raises several questions. On the phonological side, the roles of /joj/ and the structure σ_ should be clear. As per the convention adopted in the

previous chapter, /joj/ is the structure that the rule contributes, while σ_ is the context in which it is to be inserted. Specifically, the underscore indicates where /joj/ is to be placed with respect to the context.

This context is associated with a syntactic structure [CL]. This association is in actual fact a strong excitatory relation, which I assume applies in one direction only, from phonology to syntax, because we want to express the idea that the syllable preceding /joj/ must be a clitic, not that every clitic must be a syllable preceding /joj/. In other words, the excitatory relation expresses a requirement. Excitatory relations can generally be understood this way: if element A is active and has an excitatory relation with element B, then element B must also be active. If it is not, the structure is ungrammatical.[6]

In essence, the structure [CL] in (12) is the syntactic context in which the fully specified clitic structure [CL, 3SGF, D, DAT] must be inserted, although the exact relation between the two syntactic structures is not specified. There is no need to specify this relation, since the important part of the rule, the fact that a clitic may only follow another clitic, is expressed in the phonological component. Linear Correspondence will ensure that the syntactic relation is one of c-command, but this does not need to be expressed in the rule.

5.3 Scrambling in German

The German clause can be analysed descriptively using a template (cf. Eisenberg 2013). This template consists of (usually) five fields, each of which contains specific types of constituents. Table 5.1 gives an overview of the template and the elements that occupy each field in different clause types. It also gives the equivalent position in a standard minimalist clause structure analysis. The template is based on two brackets (German *Satzklammern*), which in a way define the skeleton of the clause. The left bracket contains the complementiser in subclauses and the finite verb in main clauses, and corresponds to the C° position. The right bracket contains the verb complex (minus the finite verb in main clauses, of course, since it moves to the left bracket) and corresponds to the V° and T° positions (on the assumption that both are head-final).

The two brackets delineate three fields, simply called the prefield, the middle field and the postfield (German *Vorfeld*, *Mittelfeld* and *Nachfeld*). The prefield corresponds to Spec,CP and contains at most one constituent, but remains empty in subclauses and in yes/no questions. The postfield contains extraposed elements, usually CPs or PPs, occasionally also other kinds of constituents. The middle

[6]Note, however, that it is in principle possible that B is inhibited by another part of the structure. If this inhibition cancels out the excitation coming from A, a structure containing A can be grammatical even if B is not present.

	prefield	left bracket	middle field	right bracket	postfield
subclause	∅	Compl	XP*	V	CP / PP
main clause	XP	V_{+fin}	XP*	V_{-fin}	CP / PP
rel. clause	rel. pron.	∅	XP*	V	CP / PP
equivalent	Spec,CP	C	Spec,TP+VP	V+T	extraposition

Table 5.1: German clause structure

field is where the action is, so to speak. It contains all constituents not already positioned elsewhere, and corresponds to all argument and adjunction positions of TP and VP.

This template defines a strict and fixed order in the German clause, but the order of constituents in the middle field is essentially free. It is generally accepted that there is a default or unmarked order, but given the right context, essentially any order is possible (cf. Lenerz 1977), a phenomenon that is known as *scrambling*. Several factors have been identified that influence the order in the middle field, such as case, definiteness, focus, etc. Since these factors may have conflicting ordering requirements, they are in competition. Below, I discuss an Optimality Theory analysis of a number of these factors and consider how it can be implemented in the model sketched here. Before that, I take a look at stress placement in the middle field, which also places a restriction on scrambling, albeit in a different way.

5.3.1 Stress placement

In a paper arguing that syntax does not interact with information structure directly, Fanselow (2007, p. 211) states that the main factor conditioning scrambling in the German clause is the position of stress, which is itself linked to focus. Fanselow mentions several theories on stress placement in German,[7] "which all more or less imply that the 'main' accent should be as far to the right as possible" (p. 211). As Fanselow points out, this constraint must be violable, because the unmarked word order in German does not always have main stress at the right edge, but it constrains scrambling, because a scrambling operation "must not make the structure worse with respect to accent placement" (*ibid*).

The following pair of questions, together with the answers in (14) below, all taken from Fanselow (2007, pp. 210-211), illustrate the effect:

[7]Specifically, he refers to Cinque (1993), Féry and Kügler (2006), and Samek-Lodovici (2005).

(13) a. Wer hat den Hubert eingeladen?
who-NOM has the.ACC Hubert invited
'Who invited Hubert?'

b. Wen hat der Gereon eingeladen?
who-ACC has the.NOM Gereon invited
'Who did Gereon invite?'

Both questions in (13) are *wh*-questions, putting the *wh*-constituent in focus. (13a) questions the subject, (13b) the object. Consider now the following two answers:

(14) a. Ich denke, dass der Gereon den Hubert eingeladen hat.
I think, that the.NOM Gereon the.ACC invited has
'I think Gereon invited Hubert.'

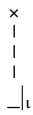
b. Ich denke, dass den Hubert der Gereon eingeladen hat.
I think, that the.ACC the.NOM Gereon invited has
'I think Gereon invited Hubert.'

The answer sentences use embedded verb-final clauses in order to avoid the verb-second effect, which complicates the picture. Both sentences in (14) have the same meaning, but the order of subject and object differs. Both are acceptable as answers to question (13a), but only (14a) is an appropriate answer to (13b). Put differently, in (14a), which has the unmarked subject>object order, both subject and object can be in focus, despite the fact that if the subject is in focus, it carries the main accent, violating the constraint that the main accent should be as far to the right as possible. In this case, the default order licenses the violation. A similar stress placement is not possible in (14b): if the object is in focus, and hence has main stress, it cannot be scrambled in front of the subject. The order in (14b) is only appropriate if the subject *der Gereon* is stressed.

Fanselow claims that the relevant constraint is stress alignment and not focus alignment, which is a plausible assumption given the fact that focus is a semantic category (possibly with an accompanying syntactic feature), whereas alignment is a phonological notion. Alignment is therefore only possible with phonological categories: a semantic or a syntactic category cannot be aligned, only its phonological component can. If we assume that stress is the phonological component of focus, then it should be stress that is aligned.[8]

[8]Féry (2010) argues that stress is only one possible correlate of focus. Féry (2013) develops a theory of focus alignment, arguing that a focused constituent is universally aligned with an IntP or a p-phrase boundary, but her examples suggest that what is actually aligned is not focus itself but the phonological correlate of focus, which would be in line with assumptions made here.

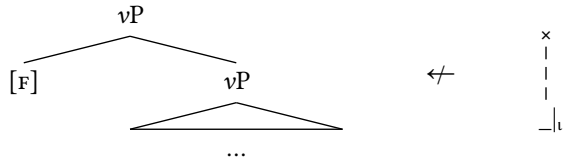
We may therefore assume a rule of stress alignment along the following lines:

(15) 

Here, the symbol × stands for stress, the ι for an intonational phrase, and the pipe bar for a prosodic boundary. The underscore indicates the position with which stress is associated. In addition, we need a rule mapping a syntactic focus feature [F] onto stress:

(16) [F] ↔ ×

The question that we now need to answer is how, on the basis of these two rules, we can ensure that scrambling does not move the stressed constituent to a position where (15) is violated. As stated above, the rule of stress alignment is not strong enough to force the stressed constituent out of its base position to a position closer to the right boundary of the IntP, which suggests that the stress alignment rule does not force a syntactic reordering. Rather, it seems that stress alignment simply prohibits the stressed constituent from being scrambled. Assuming that scrambling involves adjunction to *v*P, we can achieve the desired result by positing an inhibitory relation between the stress alignment rule and a syntactic structure that puts the focus feature [F] in a *v*P-adjoined position:

(17) 

The result of this rule is that any syntactic structure that adjoins a constituent with a focus feature [F] to *v*P is inhibited. Note that it is impossible in the present model to express that the *stressed* constituent cannot be adjoined to *v*P. Stress is a phonological property, whereas the structure that must be inhibited is a syntactic one. Under the current assumptions, it is not possible to identify the syntactic constituent whose phonological component has received stress, since stress assignment is handled by one rule and the inhibition of adjunction to *v*P of the stressed constituent by another.

The current model does not provide a way to identify the stressed constituent in syntax, because stress is a phonological property. Therefore, the rule in (17) can only target the constituent with a focus feature [F], given that this feature

generally correlates with stress. The only way to block the *stressed* and not the focused constituent from adjoining to *v*P would be to mark the stressed constituent with a syntactic feature, but this would be an undesirable solution, since it would introduce a morphosyntactic feature for the sole purpose of identifying the stressed constituent.

Because the rule in (17) identifies the constituent with an [F] feature, the analysis predicts that a constituent that has main stress but does not have an [F] feature (i.e., is not in focus), should be able to occur adjoined to *v*P, i.e., it should be possible to scramble it. I am not aware of any test cases for this prediction, however, since a stressed constituent is generally associated with focus.

The reverse is not always true, as Féry (2010) shows. A focused constituent in German can be deaccented under certain conditions, e.g., because of stress clash. According to (17), such a constituent should still be prohibited from adjoining to *v*P, despite the lack of stress. The stress clash cases that Féry (2010) gives cannot be tested, however, because when the focused constituent is scrambled, the stress clash disappears. Féry provides another example, however, involving *Second Occurrence Focus* (SOF), a combination of focus and givenness, that might be testable. In the example in (18), which is an adaptation from Partee (1999), *nur* 'only' is a focus marker, indicating that the constituent it is associated with is focused. In (18a), its associated constituent *Gemüse* 'vegetables' is stressed. In (18b), the phrase *nur Gemüse* 'only vegetables' has a SOF and is consequently deaccented:

(18) a. Jeder wusste, dass Anna nur Gemüse$_{F}$ isst.
everyone knew, that Anna only vegetables eats
'Everyone knew that Anna only eats vegetables$_{F}$.'

b. Wenn sogar Paul$_{F}$ wusste, dass Anna nur Gemüse$_{SOF}$ isst, hätte
when even Paul knew, that Anna only vegetables eats had
er ein anderes Restaurant vorschlagen sollen.
he another restaurant suggest should.INF
'If even Paul$_{F}$ knew that Anna only eats vegetables$_{SOF}$, he should have suggested another restaurant.'

If we attempt to adapt the examples above in (13) and (14) to this scenario, it seems that scrambling of a SOF is indeed not possible:[9]

[9]We cannot adapt (18) to test the possibility of scrambling, because *Gemüse* 'vegetables' is both inanimate and indefinite, two further reasons why its preferential placement is after *Anna*, which is animate and definite.

(19) a. Wen hat der Gereon eingeladen?
Who did Gereon invite?

b. Ich denke, dass der Gereon nur den Hubert$_{F}$ eingeladen hat.
I think that Gereon invited only Hubert.

c. Wenn sogar Paul$_{F}$ wusste, dass nur den Hubert$_{SOF}$ der
If even Paul knew that only the.ACC Hubert the.NOM
Gereon eingeladen hat, hätte er auch selber kommen können.
Gereon invited has, had he also self come can.
'If even Paul$_{F}$ knew that Gereon invited only Hubert$_{SOF}$, he could have come himself as well.'

(19c) does not appear to be particularly good, suggesting that despite the fact that *nur den Hubert* has been deaccented, it cannot be scrambled. However, it is possible that scrambling here is bad for another reason: scrambling in (19c) reverses the order of subject and object in comparison to (19b), which the (c) clause essentially repeats. Possibly, then, scrambling is bad here because an expectation of symmetry is violated.[10]

Summarising, the idea that stress alignment inhibits the adjunction of a focused constituent to *v*P, essentially prohibiting scrambling of the focused constituent, may be an indirect way to account for the observation discussed by Fanselow (2007), but there does not seem to be anything immediately speaking against it, and the analysis shows how an inhibitory relation may be used to express a cross-modular prohibition.

There is a more fundamental conclusion to be drawn here, as well. In the system as outlined here, it is not possible to state general prohibitions. It is, for example, not possible to state that a stressed (or focused) constituent cannot be adjoined to *v*P (i.e., cannot be scrambled). The only way to include such a prohibition is to connect it to a piece of structure that inhibits it. In the current example, we use the stress alignment rule, because it intuitively makes sense to link the prohibition against scrambling of the stressed constituent to it. In principle, we could have linked it to any other piece of structure, but it is important to link it to a structure that is active in the right context. That is to say, it will not be very helpful to link the prohibition against scrambling of the stressed (focused) constituent to a structure that describes the CP, for example, because at the point where that structure is active, the *v*P is presumably too deeply embedded to be affected. Association with the stressed constituent should trigger the prohibition at the right moment / in the right context, however.

[10] The rule in (17) makes another prediction: if the stress alignment rule itself is suppressed, it cannot inhibit a syntactic structure with an [F] feature adjoined to *v*P. Again, however, I am not aware of any possible test cases.

5.3.2 The order of constituents

The order of constituents in the German middle field, or, in structural terms, constituents within or adjoined to *v*P/VP, is influenced by a number of factors, as is well-known. In the remainder of this section, I look at the way some of these factors interoperate, in order to determine how they could be encoded in the current system. I will base myself primarily on Müller (1999, 2000), who develops an Optimality-Theory based analysis of scrambling in the German middle field. Müller adopts the following constraints:[11]

(20) a. NOM: [+nom] before [–nom]

b. DEF: [+definite] before [–definite]

c. AN: [+animate] before [–animate]

d. FOC: [–focus] before [+focus]

e. DAT: [+dat] before [+acc]

The constraints are ranked in the order given here, with NOM being the highest-ranking constraint.[12] Müller specifically argues that the features used in the constraints are syntactic, even those that are linked to semantic properties, because there is no one-to-one correspondence. For example, the syntactic focus feature is associated with one constituent, but focus projection can create a semantic focus domain that is larger than just the element carrying the syntactic focus feature.

As a first step toward developing an implementation of the constraints in (20), it is important to note that we cannot make reference to Spec,*v*P, as we did in the analysis of stress alignment above. The constraints governing word order in the middle field also apply to the base-generated order, unlike the constraint on stress alignment. They can, in fact, be triggers for scrambling.[13]

The first question that we need to answer is whether a syntactic or a phonological analysis would be preferable. The constraints in (20) all use the term

[11] As Müller notes himself, these do not capture all the factors that play a role in establishing word order in the German middle field. Like Müller, I am not trying to develop a comprehensive analysis of scrambling in German. Rather, I aim to demonstrate how the model outlined here can be used to analyse scrambling.

[12] Müller (1999) assumes two more constraints, ranked lower than the ones in (20). The first states that NPs precede adverbials, which should not pose a problem for the current model but which I ignore here, because the facts are more complicated, since the unmarked position of an adverbial in the middle field also depends on its type. The second additional constraint is a general permutation constraint, which forces scrambling and which exists for theory-internal reasons that do not apply in the current model.

[13] At least in a meta-theoretic sense. Whether the constraints trigger scrambling in the grammar model as well is a different matter, one that I will not go into here.

"precedes", which suggests that the relevant notion is phonological, and since we used a phonological context to analyse the Serbo-Croatian clitics as well, this might seem a viable option. There are, however, two problems with this option that the clitic case does not suffer from.

The first problem is that precedence in the constraints in (20) does not mean *immediate* precedence. The phonological context can therefore not be defined as strictly as for the Serbo-Croatian clitics: other material may intervene. Although traditionally, there is a phonological notation that allows for non-adjacency, it is not suitable for our purposes. For example, in an SPE-type rule for final devoicing as in (21), C_0 indicates that zero or more consonants may appear in between the position of the segment to which the rule applies and the word boundary:

(21) $[+\text{voice}] \longrightarrow [-\text{voice}] \;/\; \underset{[-\text{son}]}{\underset{|}{\_\_}}\; C_0\ \#$

Note, however, that even in (21), the segment to which the rule applies is still related to a (prosodic) *boundary*. More contemporary representations of final devoicing express this fact by using an autosegmental representation in which the [–voice] feature is associated with the coda of the syllable. The precedence requirement of the constraints in (20) is quite different. The constraints order constituents with respect to each other, independently of any prosodic boundaries. There is nothing in current phonological theory that would allow us to formulate such a general precedence requirement.

The second problem is related to the first: in the Serbo-Croatian clitic example, it is clear which prosodic category the rule should target: clitics are all syllables. In the current case, it is not at all clear what sort of prosodic categories we would have to target. We need to order syntactic constituents, but a single constituent may correspond to a prosodic word, to a prosodic phrase, or even to multiple prosodic phrases. Furthermore, due to prosodic readjustment and to mismatches between syntactic and prosodic phrasing, it is not always a certainty that a syntactic constituent corresponds to a prosodic constituent. The prosodic constituent may be smaller or larger than the syntactic constituent. In short, the idea of analysing the precedence requirements of the constraints in (20) in phonology seems futile.

We therefore have to analyse the precedence relation syntactically. This can be done fairly straightforwardly on the assumption that c-command maps to precedence, at least in the relevant domain. The relevant domain here is the middle field, in which all constituents except at most one occupy specifier positions. This one possible exception is the verb's complement, which is base-generated in Comp,VP. Depending on one's theoretical assumptions, it may remain there,

but even then, the relation between c-command and precedence holds, at least with respect to the other constituents in the middle field.[14]

With these considerations in mind, let us look at the first constraint in (20a), which states that [+nom] should precede [–nom], which, as argued, must be implemented as a rule requiring that a [+nom] argument should c-command a [–nom] argument. We now have two options to proceed. We can either specify a context that places specific requirements on insertion of [+nom] or of [–nom] arguments (or both), or we can add a schema to the grammar that specifies a particular c-command relating between [+nom] and [–nom] and specify that this structure is required or prohibited.

Specifying a context for [+nom] is not an option, because there is no requirement that [+nom] c-command [–nom]. It is only required to do so if there *is* a [–nom] argument in the clause, but it is not possible to express this by specifying a syntactic context. In order to allow clauses that only contain a nominative argument, we would also have to specify a context for [+nom] in which it does not c-command a [–nom] argument. This, however, would make it possible for [+nom] to be c-commanded by a [–nom] argument.

The alternative, specifying a context for [–nom], is equally problematic. Again, the problem is that there is no requirement that [–nom] be c-commanded by a [+nom] argument. Although most clauses contain a nominative subject in German, there are verbs that do not select a nominative argument:

(22) Ich sagte, dass mir vor ihm graut.
I said that me.DAT before him scares
'I said that I'm scared of him.'

Verbs such as *grauen* 'to scare' do not require a nominative subject. Obviously, if there were a constraint on [–nom] arguments that they be c-commanded by a [+nom] argument, clauses such as (22) would be expected to be ungrammatical or at least marked.

Verbs of this type are quite rare in German, but the same considerations also apply to passives of verbs that assign inherent (or lexical) case to their objects, which are more common. A verb such as *helfen* 'to help' assigns dative to its internal argument. When such a verb is passivised, the inherent case is not suppressed. The result is a verb that resembles an impersonal passive, in the sense that it does not have a nominative-marked argument, but it retains its internal, inherently case-marked argument, as shown in (23):

[14]Not necessarily with the verb, but that is irrelevant for the present case.

(23) Er meinte, damit wäre ihm nicht geholfen.
He meant, there-with were him.DAT not helped.
'He said that would not help him.'

All in all, we cannot claim that [–nom] arguments must appear in a context in which they are c-commanded by a [+nom] argument, because it is obvious that this is not always the case.

Since it is impossible to specify a context for either [+nom] or [–nom], we need to use a different method to implement the constraint NOM. The only other option that the system provides is to explicitly require or prohibit a schema. Requiring a schema in which [+nom] c-commands [–nom] is not possible, for essentially the same reasons that specifying a context is not possible. We cannot adopt such a requirement because it is perfectly possible to have clauses that do not satisfy it. We could license such clauses by additionally adding a schema to the grammar that contains only a [+nom] argument, but this schema would then be able to license structures in which a [–nom] argument c-commands a [+nom] argument: the [+nom] argument would be licensed by the second schema, the [–nom] argument would be licensed by a more general rule allowing non-nominative arguments, and the structure would be fine.

The option that we are left with is therefore to prohibit structures in which [–nom] c-commands [+nom]. In a way, this makes sense, because this is exactly the structure that the constraint in (20a) aims to prevent. As discussed above, however, the model does not provide a mechanism for adopting general prohibitions. A prohibition is only possible if it is connected to a piece of structure that triggers it. As it turns out, this is not a bad thing in the current case, because it would not be possible to have a general prohibition against a non-nominative argument c-commanding a nominative argument. After all, Spec,CP can be occupied by a non-nominative argument without difficulty:

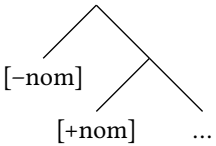
(24) Diesen Film möchte mein Bruder gerne sehen.
this.ACC film would.like my.NOM brother gladly see
'My brother would really like to see this film.'

(24), with the accusative-marked object *diesen Film* 'this.ACC film' in Spec,CP and thus c-commanding the nominative subject *mein Bruder* 'my.NOM brother' is not marked or marginal in any way. The preference for nominative arguments to precede non-nominative arguments only applies to the middle field, it does not apply to the structure as a whole.[15]

[15] This is in fact true for all the constraints in (20).

In order to accommodate this fact, we can link the constraints in (20) to C. The C head closes off the middle field, in the sense that all arguments to which the constraints in (20) apply occupy a position c-commanded by C. The actual structure that needs to be inhibited can be kept fairly simple:

(25) C *↛*



For this rule to work, we need to make one particular assumption, however: in the inhibited structure in (25), the two [nom] features are close, no other specifier is intervening. The structure must be interpreted more loosely, however, since structures in which a [−nom] argument c-commands a [+nom] argument that is lower in the structure should also be inhibited. Therefore, we are forced to assume that a structure such as in (25) is a template of sorts, that can be augmented with additional material as needed.

This issue is of course the syntactic variant of the same issue that we encountered above in phonology. Although the solution is admittedly not perfect, it is nonetheless more likely that a solution can be found in syntax than in phonology. Phonological relations are always local, usually requiring adjacency or near-adjacency. In syntax, relations can be non-local, c-command being a typical example.

So far, I have not said much about the way in which the type of rules proposed here can be incorporated into an existing grammar framework. The current case offers a few insights. In a derivational framework, linking the constraints in (20) to C can be interpreted to mean that when C is merged, the constraints become active, ruling out, or at least disfavouring, any structure that does not comply.[16] Once C has been merged and the constraints evaluated, they are not evaluated a second time. This makes it possible to have a [−nom] argument in Spec,CP.

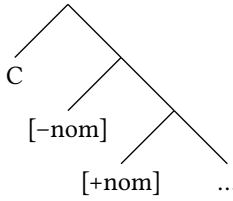
In a representational framework, things are somewhat different. Because of the representational nature of such a framework, it is probably not possible to restrict a structure to a specific domain if it is interpreted as a template. The rule can only be checked against the structure as a whole, because this is the only structure that the grammar model provides. There are no "intermediate" structures in the way that they exist in a derivational framework. Therefore, the

[16] Obviously, many details would need to be worked out, but they depend on the grammar framework one assumes. In minimalism, for example, one may be tempted to connect the evaluation of such constraints to phases, on the assumption that Spec,TP is not transferred with T and its complement.

rule in (25) would predict that a [–nom] argument cannot appear in Spec,CP if the clause also contains a [+nom] argument.

There is a simple, straightforward way in which we can make the rule work, however. All we need to do is to include the C head into the template, as in (26):

(26) C ↛



Due to the presence of C, the structure cannot apply to a [–nom] argument in Spec,CP that c-commands a [+nom] argument lower in the structure. Only a structure in which a [–nom] argument below C c-commands a [+nom] argument is ruled out by (26). Note that this is another reason why we need the assumption made above that the syntactic structure in (26) is a template: we want to rule out any structure in which [–nom] c-commands [+nom] under C, not just those structures in which [–nom] appears immediately below C.[17]

At this point, we have only accounted for one of the constraints in (20). The other constraints can be handled in much the same way: for each, we need a restriction on the element that comes last to c-command the element that comes first and each of these restrictions needs to be tied to C in order to limit its application to the middle field.

The more interesting question at this point is of course how we can incorporate the relative ranking of the constraints in (20). The ranking of constraints in OT expresses that a violation of a constraint C1 that outranks a constraint C2 will rule out a structure S, even if S does not violate C2. In other words, if C1 is active, C2 must be suppressed, so that it is no longer able to rule in S.

This idea is on the right track, but it is not quite there. In the current model, constraints are of course abstractions, they are not data and they are not encoded explicitly. This means that it is not possible for one constraint to activate or suppress another. The system only represents pieces of structure, i.e., schemata, so only schemata can activate or suppress other schemata.

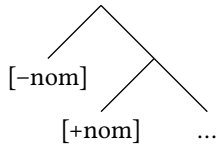
In order to make the discussion a bit more concrete, consider the first two constraints of (20), repeated here in (27) and converted into the current format in (28):

[17] This of course means that we need to restrict templates such as in (26) from applying across clauses. German has embedded V2 clauses, and we do not want to inhibit a [–nom] argument from appearing in the embedded Spec,CP due to the fact that there is a matrix T head.

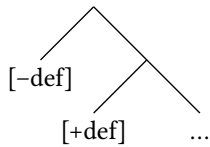
(27) a. NOM: [+nom] before [–nom]

b. DEF: [+definite] before [–definite]

(28) a. C ↛



b. C ↛

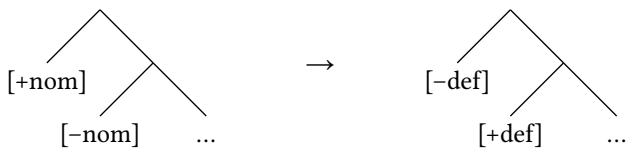


Looking at the issue at hand in terms of excitation and inhibition, what we want to say is that the constraint in (27a) / (28a) inhibits the constraint in (27b) / (28b). Since we cannot say that constraints excite or inhibit each other, our first idea might be to say that the structure after the arrow in (28a) inhibits the structure after the arrow in (28b). This, however, is also not possible. If a clause is generated in which a [–nom] argument c-commands a [+nom] argument, the inhibition relation in (28a) will inhibit this structure. As I argued above in the discussion on the Elsewhere Condition, an inhibited structure cannot excite or inhibit another structure, because it is inactive.

Considering the issue more closely, it becomes obvious that this is not even what we want. It should not be the case that the structure [–nom] ≫ [+nom] (where ≫ indicates c-command) inhibits the structure [–def] ≫ [+def], because that would mean that if a [–nom] argument c-commands a [+nom] argument, a [–def] argument cannot c-command a [+def] argument. The point is that we want the latter case to be true in general, *except* when adhering to that constraint would yield a structure in which a [–nom] argument c-commands a [+nom] argument.

Put differently, we want a structure in which a [+nom] argument c-commands a [–nom] argument —that is, the exact opposite of the structure in (28a)— to *allow* a [–def] argument to c-command a [+def] argument. In structural terms:

(29)



As it turns out, then, we need *two* schemata in order to implement a constraint such as NOM and to integrate it into a constraint hierarchy. The first, which we may call for ease of reference the negative counterpart of the constraint NOM, is the structure that is to be inhibited. The second, the positive counterpart of NOM, is the structure that inhibits the negative counterpart of DEF.

Obviously, NOM needs such excitatory relations with all the constraints in (20). Additionally, DEF needs similar excitatory relations with all the constraints in (20) that it outranks, and so on for the other constraints as well. Crucially, however, these excitatory relations go in one direction. There are no excitatory relations from any of the constraints in (20) to a constraint that outranks it.

Let us see how the rules proposed here are able to replicate the result of Müller's OT analysis. The relative ranking of NOM and DEF is shown by (30) (from Müller 1999, p. 797):

(30) a. dass eine Frau den Fritz geküsst hat.
that a.NOM woman den.ACC Fritz kissed has
'that a woman kissed Fritz.'

b. ?dass den Fritz eine Frau geküsst hat.
that den.ACC Fritz a.NOM woman kissed has
'that a woman kissed Fritz.'

As is generally the case when it comes to scrambling orders in the German middle field, both orders in (30) are grammatical, but there is a clear preference for (30a), making the structure in (30b) marked. This markedness is indicated here with the question mark.

Müller gives the following tableau for the clauses in (30):[18]

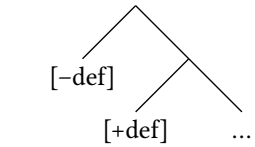
(31)

	NOM	DEF	AN	FOC	DAT
☞ C_1: eine Frau den Fritz		*			
C_2: den Fritz eine Frau	*!				

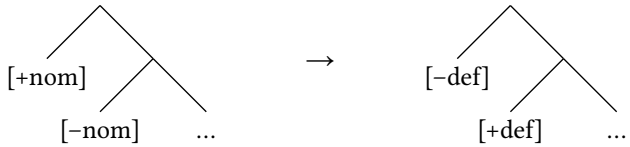
[18] In actual fact, Müller argues that the tableau in (31) represents a *subhierarchy* that governs the order in the middle field. In his model, a subhierarchy is a hierarchy of constraints that occupies a single slot (i.e., corresponds to a single constraint) in the main hierarchy. Grammaticality is determined by the main hierarchy, and for this purpose, a single constraint from the subhierarchy must be substituted for the corresponding constraint in the main hierarchy. The constraint that is substituted for the main constraint can be freely chosen from among the subconstraints, meaning that for all of the candidates of the subhierarchy, there will be at least one constraint that renders them grammatical in the main hierarchy. The subhierarchy does not determine grammaticality but markedness. The details of Müller's analysis do not concern us here, however. For our purposes, it suffices to look at the subhierarchy in isolation.

The analysis is straightforward: candidate C_1 violates DEF, because the subject is indefinite, whereas the object is definite, but this violation is not fatal, because the alternative C_2 violates NOM, which is a higher-ranked constraint.

In the current proposal, the three rules shown above in (28a/b) and (29) are relevant. Let us first consider the sentence in (30a), the unmarked structure, which has the order [nom] > [acc] and [–def] > [+def] (where > stands for precedence). This structure is inhibited by the rule in (28b), repeated here:

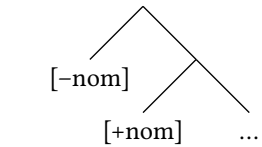
(28) b. C 

This rule should actually be enough to rule out (30a) (or in this case, degrade its acceptability), but its effect is counteracted by the rule in (29), also repeated here:

(29) 

The clause in (30a) matches the first schema in this rule, which means that the schema [–def] ≫ [+def] is activated again, making it available for licensing (30a).

The clause in (30b) is subject to the rule in (28a), also repeated here:

(28) a. C 

Since (30b) has a non-nominative argument preceding the nominative argument, its acceptability is degraded, resulting in a marked structure. In this case, there is nothing that counteracts this result, because there is no rule that links the schema [–def] ≫ [+def] to a schema that allows a non-nominative argument to c-command a nominative one.

Optimality Theory explicitly ranks constraints. When converting such a constraint ranking to the current model, the ranking is encoded implicitly by the excitatory connections between schemata encoding the constraints. As we have seen, we actually need two schemata to encode a single constraint, at least for

the constraints in (20). This means that converting an OT-style analysis to the current model is not a straightforward affair. Each constraint will have to be considered carefully to see what schema or schemata are necessary in order to implement it, and what excitatory and inhibitory relations are needed.

In general, however, the current model seems a good fit for implementing OT-style analyses. The ability to establish inhibitory and excitatory relations between pieces of structure (schemata) allows us to encode the relative ranking, although just as with the Elsewhere Condition, we are required to assume quite a number of such relations in order to encode the ranking of even a small constraint set. For the five constraints of (20) we need 4+3+2+1=10 excitatory connections in order to encode the hierarchy. Add a sixth constraint and we need 5+4+3+2+1=15 connections, etc.

However, such connections are straightforward and simple and allow us to represent constraints implicitly. An implementation of OT-style constraints and constraint ranking would arguably be more complex, since the constraints would have to be represented explicitly, as data, in order to attach a ranking to them.

It is important to note at this point that the underlying idea of the model being sketched here is that when a schema S is "inactive", any structure that matches S is ruled out. In other words, an inactive (inhibited) schema does not behave as if it were not there. In general, formal grammars require that every element of a structure be licensed, but there is no need to license a larger structure if all of its parts are licensed. For example, a minimalist model does not have to license every XP as a whole, because the projections from X to XP are all licensed individually. The same is true, *mutatis mutandis*, for HPSG-style grammars.

One might therefore be inclined to think that if a schema is inactive, it is essentially irrelevant for the question whether a specific structure can be generated / licensed or not. The current proposal makes the explicit assumption that this is not the case. An "inactive" structure is not truly inactive. Rather, it acts to inhibit any structure that matches it. This is why it is necessary in the analysis of (30a) to reactivate the schema [–def] ≫ [+def]: as long as this schema is inactive due to (28b), any structure in which an indefinite argument c-commands a definite argument is inhibited.

This of course raises the question how it is determined which schemata are part of the grammar, available for activation and/or inhibition. Why is it that the schema [–def] ≫ [+def] is available, but a schema such as [+def] ≫ [+def], which could be used to inhibit any structure that contains two definite arguments, is not? A seemingly trivial answer to this question is that the grammar does not contain a schema of the type [+def] ≫ [+def] for the simple reason that it does not. There is in fact more truth to this answer than one might think at first sight. The grammar does not establish any generalisations over structures that contain two

definite arguments because there is no need to do so. Although clauses with two definite arguments are undoubtedly fairly common in day-to-day language use, and hence in the input received by a language-learning child, if such structures do not exhibit any patterns that set them apart from other kinds of structures, there is no need to create a generalisation for them.

This seems intuitively obvious, but it raises several questions that have no immediately obvious answers. For instance, if [−def] ≫ [+def] is a schema that is needed to describe word order in the German middle field, and if such schemata are essentially generalisations, what input would allow such a schema to be established as part of the grammar? Given the observation that [+def] arguments generally precede [−def] arguments, structures that have [−def] ≫ [+def] should be relatively rare. Is this nonetheless sufficient to establish a generalisation? It seems unlikely that the absence of a particular structure is sufficient to add a rule to the grammar that prohibits its occurrence, because that would most likely lead to massive redundancy. Perhaps what is relevant is the fact that a structure, even though it is rare, typically occurs in a very specific context. In the case of the schema [−def] ≫ [+def] this would be a structure in which the [−def] argument is [+nom] and the [+def] argument is [−nom]. Such a situation may be sufficient to include the rule in (29), which states that [−def] ≫ [+def] is allowed in the context of [+nom] ≫ [−nom], but it would also have to trigger the inclusion of a rule that inhibits the schema [−def] ≫ [+def] in the general case.

The proposal sketched here raises other questions as well. For example, if a schema S can be inhibited by a schema A and activated by another schema B, what determines whether S is active or inactive when both A and B are active? This situation occurs in the analysis of (30) above, where I argue that the schema [+nom] ≫ [−nom] actually activates [−def] ≫ [+def], so that the DEF constraint can be violated. It is clear that in this particular case, the activation from [+nom] ≫ [−nom] must override the inhibition from C (which, as I argued above, triggers the DEF constraint), but nothing in the model sketched here ensures this. One option that immediately presents itself is the idea that inhibitory and excitatory connections are not binary but can differ in strength. If the inhibitory connection between C and [−def] ≫ [+def] is weaker than the excitatory connection between [+nom] ≫ [−nom] and [−def] ≫ [+def], then the latter outweighs the former and we obtain the required effect.

Another question that is raised and that may in fact be accounted for in the same way is how ungrammaticality can be distinguished from markedness, or how degrees of ungrammaticality can be accounted for. Here, too, the solution might be found in weighted connections, although this should probably not be taken to mean that stronger ungrammaticality is simply the result of stronger inhibitory connections, since ungrammaticality, especially different degrees of ungrammaticality, may also be the result of contrast. How exactly such contrasts

can be implemented in the current system is also an open question.

These are just a few of the questions that the current chapter cannot answer and that will require a more substantial investigation of the system sketched here. One final note is in order, though. A system of excitatory and inhibitory connections, especially if one were to incorporate weighted connections, may look superficially similar to a neural network. It is important to realise, however, that such a system does not constitute a neural network, nor is there any claim that it is especially well-suited to be implemented in a neural network. The only claim that I make is that this system allows us to reason about the grammar of natural language at a high level of analysis in a manner that should make it easier to implement it at a lower level of analysis. This lower level is almost certainly not the neural network, but rather an intermediate level, perhaps the level of *neural computation*, as in Pulvermüller (2003, 2010).

One may also look at this from the other direction. Although a neural network is a system consisting of units that can be activated and that have excitatory and inhibitory connections to other such units, it does not automatically follow that in order to model what such a system does at a higher level, one needs a similar system of activation units with excitatory and inhibitory connections. To use the analogy of the electronic computer again: although computers are binary machines at a very low level, analysing how a web search is executed does not necessarily require a "high-level binary model" (whatever that may be).

5.4 Concluding remarks

The basic idea of the current chapter is that we can describe grammar rules as links, both excitatory and inhibitory, between pieces of structure. These pieces of structure, or *schemata*, license structures, in a manner that is reminiscent of grammar models such as HPSG. There is one crucial difference however: schemata that apply to a certain structure and that are inactive effectively inhibit the structure in question, in the sense that it will not be generated (i.e., that it is ungrammatical, infelicitous, or whatever term one may wish to apply to it). This assumption makes it possible to describe competition as an inhibitory relation between schemata. If two morphs compete for the same slot, one may inhibit the other, or they may both inhibit each other, with the result that as soon as one becomes active, the other one cannot do so anymore.

As stated at the outset, the approach sketched in this chapter raises quite a few questions and it will take much more analysing and describing of actual data before we can begin to form an idea of what such a model can and cannot do. These questions will have to be left open for future research, however.

6
Summary and conclusions

The central argument of the present study is that hierarchical order does not map onto linear order in a direct, one-to-one fashion. The linear order, which under the assumptions made here equates to phonological structure, is multi-tiered and for that reason cannot be predicted on the basis of the hierarchical (i.e., syntactic) structure alone. The exact phonological structure is determined by the hierarchical relations in the syntactic tree together with the phonological forms of the elements involved.

This is best seen in so-called *simultaneity* data from sign languages. Many sign languages have the ability to express certain kinds of elements, negation, adverbials, topic, focus, and *wh* markers, simultaneously with a lexical sign or signs. These simultaneous elements are realised not with the hands but with other parts of the body, mainly the face and head, and are consequently called *non-manuals*. The crucial observation is that although these elements are not part of the string of manual signs, they nonetheless require a position in the syntactic tree.

We are thus faced with the existence of elements that are part of syntax but not part of the linearised segmental string resulting from a syntactic tree. Whether an element behaves in this manner or not is not predictable from its syntactic make-up. It follows that linearisation cannot be a (post-)syntactic process that takes place before phonology. Linearisation can only take place *in* phonology, not *before* it.

Nonetheless, linearisation is dependent on the syntactic structure. Two elements that are sisters in the syntactic tree are adjacent in the linear structure, provided that they are both segmental (or, more precisely, realised on the same autosegmental tier). This is a basic principle of the mapping from syntax to phonology, called *Linear Correspondence*. Elements that mapped onto an autosegmental tier must be associated with an element on the segmental tier in order

to be realised. This association is not arbitrary. It is partially conditioned by phonological factors and partially by syntactic factors. Specifically, an element that maps onto an autosegment must associate in phonology with the head of the structure it selects in syntax. For example, an autosegmental negation in sign language must associated with the verb, because Neg selects a VP in syntax. This second principle is known as Input Correspondence. In essence, it has the same purpose as Linear Correspondence, namely to keep together in phonology what belongs together in syntax. Linear Correspondence does this for elements that are segmental, Input Correspondence does this for elements that are autosegments.

In essence, the argument amounts to this: prosodic morphology does not stop at the morphological level. Its principles apply at the syntactic level as well. In the same way that morphemes can be autosegments and are then subject to specific principles, syntactic elements can be autosegments as well, and are subject to similar principles. Looking at syntax in this way opens up new ways of thinking about certain phenomena. For one, it is no longer necessary to exclude intonational patterns with specific meaning, such as interrogative intonation, from syntax. We can simply say that there is a $C_{[+wh]}$ head in the structure that is realised in phonology as an autosegment. Input Correspondence will then associate it with the entire clause.

The model also allows us to deal with clitics in a more insightful manner. There is no longer any need to assume a separate class of elements called clitics. Clitics are simply elements that do not constitute a prosodic word; syntactically, they may be subject to specific constraints, although they do not have to be. In this way, we can distinguish between various types of elements, affixes, simple and special clitics, and words. It turns out that these terms are merely descriptive, and can at times be highly misleading. By separating the syntactic and the phonological properties of each type of element and specifying which component exactly is subject to which constraint, clitics are no longer 'weird' elements. Rather, they automatically fall out from the system; it would be strange if they did *not* exist.

Clitics are often thought of as elements that stand between syntax and morphology, because they have properties of both. The prosodic syntax model is able to describe them in a way that no longer makes them stand out. The model can in fact go a step further: it offers a way of unifying syntax and morphology that does not need to make a syntactic distinction between heads that are parts of words and heads that are syntactic, contrary to previous attempts at unifying the two modules. The central idea is that whether a head is syntactic or morphological is not a property of the head itself. Rather, it is a perceptual property, in a sense an optical or acoustic illusion. In syntax, a head is a head. If a head maps onto a phonological element that does not constitute a full prosodic word and if it always appears in a fixed position with respect to a certain category of

words, we perceive this head as a morphological element. There is nothing in the *syntactic* structure, however, that marks it as such. The perceived morphological status of the head in question is a result of the interaction of the head's syntactic and phonological properties.

In short, the prosodic syntax model has a number of advantages. It allows us to describe simultaneity effects in sign languages, to classify affixes, clitics and words using a unified set of properties and more generally, to unify syntax and morphology. As with any proposal of this scale, it raises numerous questions that cannot be answered here, but which will hopefully be addressed in future research.

References

Aarons, Debra (1994). "Aspects of the Syntax of American Sign Language". PhD Thesis. Boston University.

Abels, Klaus and Ad Neeleman (2009). "Universal 20 without the LCA". In: *Merging Features: Computation, Interpretation, and Acquisition*. Ed. by José M. Brucart, Anna Gavarro, and Jaume Sola. Oxford: Oxford University Press. Chap. 4, pp. 60–79. URL: http://ling.auf.net/lingBuzz/000279.

Abney, Steven (1987). "The English Noun Phrase in its Sentential Aspect". MIT Working Papers in Linguistics. PhD Thesis. Massachusetts Institute of Technology (MIT).

Ackema, Peter (1995). "Syntax below Zero". PhD Thesis. Utrecht Research Institute for Language and Speech OTS.

Ackema, Peter and Ad Neeleman (2004). *Beyond Morphology: Interface Conditions on Word Formation*. Oxford: Oxford University Press.

– (2007). "Morphology ≠ Syntax". In: *The Oxford Handbook of Linguistic Interfaces*. Ed. by Gillian Ramchand and Charles Reiss. Oxford: Oxford University Press. Chap. 10, pp. 325–352.

Agbayani, Brian and Chris Golston (2010). "Second-Position is First-Position: Wackernagel's Law and the Role of Clausal Conjunction". In: *Indogermanische Forschungen* 115, pp. 1–21. URL: http://zimmer.csufresno.edu/~chrisg/Publications%5C_files/SecondIsFirst.pdf.

Anderson, Diane E. and Judy S. Reilly (1998). "Pah! The Acquisition of Adverbials in ASL". In: *Sign Language and Linguistics* 1–2, pp. 117–142.

Anderson, John R. (2007). *How Can the Human Mind Occur in the Physical Universe?* Oxford: Oxford University Press.

Anderson, John R., Daniel Bothell, et al. (2004). "An Integrated Theory of the Mind". In: *Psychological Review* 111.4, pp. 1036–1060. DOI: 10.1037/0033-295X.111.4.1036.

Anderson, Stephen R. (1992). *A-morphous Morphology*. Cambridge: Cambridge University Press.

Aronoff, Mark (1994). *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: The MIT Press.

Ayoun, Delila (1999). "Verb Movement In French L2 Acquisition". In: *Bilingualism: Language and Cognition* 2.2, pp. 103–125.

– (2000). "Verb Movement In French Revisited: Syntactic Theory And Experimental Data". In: *Arizona Working Papers in Second Language Acquisition and Teaching* 7, pp. 70–91.
Baker, Mark (1985). "The Mirror Principle and Morphosyntactic Explanation". In: *Linguistic Inquiry* 16.3, pp. 373–415.
Beard, Robert (1988). "On the Separation of Derivation from Morphology: Toward a Lexeme / Morpheme-Based Morphology". In: *Quaderni di Semantica* 9, pp. 3–59.
Belletti, Adriana and Luigi Rizzi (1988). "Psych-Verbs and θ-Theory". In: *Natural Language and Linguistic Theory* 6.3. Last updated - 2013-02-23, p. 291. URL: http://search.proquest.com/docview/1298072336?accountid=11144.
Bildhauer, Felix (2007). "Representing Information Structure in an HPSG Grammar of Spanish". PhD Thesis. University of Bremen.
Bird, Steven and Ewan Klein (1994). "Phonological Analysis in Typed Feature Systems". In: *Computational Linguistics* 20.3, pp. 455–491.
Bod, Rens (2009). "From Exemplar to Grammar: Integrating Analogy and Probability in Language Learning". In: *Cognitive Science* 33.4, pp. 752–793. URL: http://staff.science.uva.nl/~rens/analogy.pdf.
Boeckx, Cedric and Sandra Stjepanović (2001). "Head-ing toward PF". In: *Linguistic Inquiry* 32, pp. 345–355.
Booij, Geert (1995). *The Phonology of Dutch*. Oxford: Oxford University Press.
– (1996). "Cliticization as Prosodic Integration: The Case of Dutch". In: *TLR* 13.3-4, pp. 219–242.
– (2009). "Lexical Integrity as a Morphological Universal: A Constructionist View". In: *Universals of Language Today*. Ed. by Sergio Scalise, Elisabetta Magni, and Antonietta Bisetto. Berlin: Springer, pp. 83–100.
Borer, Hagit (1999). "Deconstructing the Construct". In: *Beyond Principles and Parameters*. Ed. by Kyle Johnson and Ian Roberts. Dordrecht: Kluwer Academic Publishers, pp. 43–90.
– (2005). *In Name Only*. Structuring Sense, Volume I. Oxford: Oxford University Press.
– (2009). "Afro-Asiatic, Semitic: Hebrew". In: *The Oxford Handbook of Compounding*. Ed. by Rochelle Lieber and Pavol Stekauer. Oxford: Oxford University Press. URL: http://www-bcf.usc.edu/~borer/borer%5C%202008.pdf.
Bošković, Željko (1995). "Participle Movement and Second Position Cliticization in Serbo-Croatian". In: *Lingua* 96.4, pp. 245–266.
– (2001). *On the Nature of the Syntax-Phonology Interface: Cliticization and Related Phenomena*. Amsterdam: Elsevier Science.

Bošković, Željko (2012). "Clitic Placement in South Slavic". In: *Journal of Slavic Linguistics* 12.1-2, pp. 39–93.

Box, George E. P. and Norman R. Draper (1987). *Empirical Model-Building and Response Surfaces*. John Wiley and Sons.

Browne, Wayles (1974). "On the Problem of Enclitic Placement in Serbo-Croatian". In: *Slavic Transformational Syntax*. Ed. by Richard D. Brecht and Catherine V. Chvany. Michigan Slavic Materials 10. Ann Arbour: Department of Slavic Languages and Literatures, University of Michigan, pp. 36–52.

– (1975). "Serbo-Croatian Enclitics for English-Speaking Learners". In: *Contrastive Analysis of English and Serbo-Croatian*. Ed. by Rudolf Filipović. Vol. 1. Zagreb: Institute of Linguistics, Faculty of Philosophy, University of Zagreb, Yugoslavia, pp. 105–134.

Bury, Dirk (2003). "Phrase Structure and Derived Heads". PhD Thesis. University College London.

Carstairs-McCarthy, Andrew (2004). "What Morphology Can Tell Us About Grammar". In: *Proceedings of the LFG04 Conference*. Ed. by Miriam Butt and Tracy Holloway King. Stanford, CA: CSLI Publications, pp. 141–148.

Chomsky, Noam (1955). "Transformational Analysis". PhD Thesis. University of Pennsylvania.

– (1957). *Syntactic Structures*. The Hague: Mouton.

– (1959). "A Review of B. F. Skinner's *Verbal Behavior*". In: *Language* 35.1, pp. 26–58. ISSN: 00978507, 15350665. URL: http://www.jstor.org/stable/411334.

– (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press.

– (1993). "The Minimalist Program". In: *The View from Building 20: Essays in Linguistics in Honor of Sylvain Brom-berger*. Ed. by Kenneth Hale and Samuel Jay Keyser. Vol. 24. Current Studies in Linguistics. Cambridge, MA: The MIT Press, pp. 1–52.

– (1995a). "Bare Phrase Structure". In: *Government and Binding Theory and the Minimalist Program*. Ed. by Gert Webelhuth. Oxford: Blackwell, pp. 383–439.

– (1995b). *The Minimalist Program*. Cambridge, MA: The MIT Press.

– (1999). "Derivation by Phase". MIT Occasional Papers in Linguistics 18.

– (2001). "Derivation by Phase". In: *Ken Hale: A Life in Language*. Ed. by Michael Kenstowicz. Cambridge, MA: The MIT Press.

– (2005). "Three Factors in Language Design". In: *Linguistic Inquiry* 36.1, pp. 1–22.

– (2008). "On Phases". In: *Foundational Issues in Linguistic Theory*. Ed. by Robert Freidin, Carlos Otero, and Maria-Luisa Zubizaretta. Cambridge, MA: The MIT Press, pp. 133–166.

Christiansen, Morten H. and Simon Kirby, eds. (2003). *Language Evolution*. Oxford: Oxford University Press.

Christophe, Anne et al. (2008). "Bootstrapping Lexical and Syntactic Acquisition". In: *Language and Speech* 51.1, pp. 61–75.

Cinque, Guglielmo (1993). "A Null Theory of Phrase and Compound Stress". In: *Linguistic Inquiry* 24.2, pp. 239–297.

– (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York: Oxford University Press.

Cinque, Guglielmo and Luigi Rizzi (2010). "The Cartography of Syntactic Structures". In: *The Oxford Handbook of Linguistic Analysis*. Ed. by Bernd Heine and Heiko Narrog. Oxford: Oxford University Press, pp. 51–65.

Clark, Eve V. (1992). *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.

Corina, David, Ursula Bellugi, and Judy Reilly (1999). "Neuropsychological Studies of Linguistic and Affective Facial Expressions in Deaf Signers". In: *Language and Speech* 42.2-3, pp. 307–331.

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

Dachkovsky, Svetlana and Wendy Sandler (2009). "Visual Intonation in the Prosody of a Sign Language". In: *Language and Speech* 52.2–3, pp. 287–314.

Deacon, Terrence W. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: W.W. Norton & Co. URL: `http://www.lel.ed.ac.uk/~jim/deaconreview.tls.html%20(review)`.

– (2003). "Universal Grammar and Semiotic Constraints". In: *Language Evolution*. Ed. by Morten H. Christiansen and Simon Kirby. Oxford: Oxford University Press, pp. 111–139.

Desouvrey, Louis-H. (2005). "Romance Clitic Clusters: The Case Connection". In: *Clitic and Affix Combinations: Theoretical Perspectives*. Ed. by Lorie Heggie and Francisco Ordóñez, pp. 33–79.

Diesing, Molly, Dušica Filipović Đurđević, and Draga Zec (2009). "Clitic Placement in Serbian: Corpus and Experimental Evidence". In: *The Fruits of Empirical Linguistics. Volume 1: Process*. Ed. by Sam Featherston and Susanne Winkler. Berlin: Mouton de Gruyter. URL: `http://conf.ling.cornell.edu/interface/pdfs/LingEvidence.pdf`.

Eisenberg, Peter (2013). *Grundriss der deutschen Grammatik. Band 2: Der Satz*. 4th ed. Stuttgart: J.B. Metzler.

Elman, Jeffrey L. (1990). "Finding Structure in Time". In: *Cognitive Science* 14.2, pp. 179–211.

– (1993). "Learning and Development in Neural Networks: The Importance of Starting Small". In: *Cognition* 48.1, pp. 71–99.

Elman, Jeffrey L. et al. (1997). *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: The MIT Press.

Embick, David and Rolf Noyer (2001). "Movement operations after syntax". In: *Linguistic Inquiry* 32.4, pp. 555–595.

Engberg-Pedersen, Elisabeth (1994). "Some Simultaneous Constructions in Danish Sign Language". In: *Word-Order Issues in Sign Language*. Ed. by Mary Brennan and Graham H. Turner. Papers presented at a workshop in Durham, England, 18–22 September 1991. University of Durham, The Deaf Studies Research Unit, pp. 71–88.

Epstein, Samuel David and Daniel Seely (2002). "Rule Applications as Cycles in a Level-free Syntax". In: *Derivation and Explanation in the Minimalist Program*. Ed. by Samuel David Epstein and Daniel Seely. Malden, Mass.: Blackwell Publishers, pp. 65–89.

Fanselow, Gisbert (2007). "The Restricted Access of Information Structure to Syntax – A Minority Report". In: *The Notions of Information Structure*. Ed. by Caroline Féry, Gisbert Fanselow, and Manfred Krifka. Vol. 6. Interdisciplinary Studies on Information Structure. Potsdam: Universitätsverlag Potsdam.

Fassi Fehri, Abdelkader (1993). *Issues in the Structure of Arabic Clauses and Words*. Vol. 29. Studies in Natural Language and Linguistic Theory. Dordrecht: Kluwer Academic Publishers. ISBN: 0-7923-2082-4.

Féry, Caroline (2010). "Informationsstruktur: Begriffe und grammatische Korrelate". In: *Linguisten-Band 9 der Japanischen Gesellschaft für Germanisten*. Ed. by Shin Tanaka. Japanische Gesellschaft für Germanisten. URL: `http://web.uni-frankfurt.de/fb10/fery/publications.html`.

– (2013). "Focus as Prosodic Alignment". In: *Natural Language and Linguistic Theory* 31, pp. 683–734.

Féry, Caroline and Frank Kügler (2006). "German as a Tone Language". Ms. University of Potsdam.

Fox, Danny and David Pesetsky (2005). "Cyclic Linearization of Syntactic Structure". In: *Theoretical Linguistics* 31.1–2, pp. 1–46.

Franks, Steven (1997). "South Slavic Clitic Placement is *still* Syntactic". In: *Penn Working Papers in Linguistics* 4. Proceedings of the 21st annual Penn Linguistics Colloquium, pp. 111–126.

– (1998). "Clitics in Slavic". Paper presented at the Comparative Slavic Morphosyntax Workshop, Bloomington, Indiana.

– (2000). "Clitics at the Interface: An introduction to *Clitic Phenomena in European languages*". In: *Clitic Phenomena in European Languages*. Ed. by Frits Beukema and Marcel den Dikken. Amsterdam / Philadelphia: John Benjamins, pp. 1–46.

Gazdar, Gerald et al. (1985). *Generalized Phrase Structure Grammar*. Oxford: Blackwell Publishers.

Georgi, Doreen and Géréon Müller (2010). "Noun-Phrase Structure by Reprojection". In: *Syntax* 13.1, pp. 1–36.

Gold, E. Mark (1967). "Language Identification in the Limit". In: *Information and Control* 16, pp. 447–474.

Goldsmith, John (1976). "Autosegmental Phonology". PhD Thesis. Cambridge, MA: The MIT Press.

– (1990). *Autosegmental and Metrical Phonology*. Oxford: Blackwell.

– ed. (1995). *The Handbook of Phonological Theory*. Cambridge, MA: Blackwell.

Goldstein, Louis, Dani Byrd, and Elliot Saltzman (2006). "The Role of Vocal Tract Gestural Action Units in Understanding the Evolution of Phonology". In: *From Action to Language via the Mirror Neuron System*. Ed. by Michael A. Arbib. Cambridge: Cambridge University Press, pp. 215–249.

Golosio, Bruno et al. (2015). "A Cognitive Neural Architecture Able to Learn and Communicate through Natural Language". In: *PLoS ONE* 10.11, e0140866. DOI: 10.1371/journal.pone.0140866. URL: http://dx.doi.org/10.1371%2Fjournal.pone.0140866.

Grimshaw, Jane (1990). *Argument Structure*. Cambridge, MA: The MIT Press.

Haider, Hubert (1986). "Fehlende Argumente: vom Passiv zu kohärenten Infinitiven". In: *Linguistische Berichte* 101, pp. 3–33.

Hale, Kenneth and Samuel Jay Keyser, eds. (1993). *The View from Building 20: Essays in Linguistics in Honor of Sylvain Brom-berger*. Vol. 24. Current Studies in Linguistics. Cambridge, MA: The MIT Press.

Halle, Morris and Alec Marantz (1993). "Distributed Morphology". In: *The View from Building 20: Essays in Linguistics in Honor of Sylvain Brom-berger*. Ed. by Kenneth Hale and Samuel Jay Keyser. Vol. 24. Current Studies in Linguistics. Cambridge, MA: The MIT Press. Chap. 3, pp. 111–176.

Halpern, Aaron Lars (1992). "Topics in the Placement and Morphology of Clitics". PhD Thesis. Stanford University.

Hanke, Thomas (2004). "HamNoSys – Representing Sign Language Data in Language Resources and Language Processing Contexts". In: *LREC 2004, Workshop Proceedings: Representation and Processing of Sign Languages*. Ed. by Oliver Streiter and Chiara Vettori. Paris: ELRA, pp. 1–6. URL: http://www.sign-lang.uni-hamburg.de/dgs-korpus/index.php/hamnosys.html.

Happ, Daniela and Marc-Oliver Vorköper (2006). *Deutsche Gebärdensprache: ein Lehr- und Arbeitsbuch*. Frankfurt am Main: Fachhochschulverlag.

Harley, Heidi and Rolf Noyer (1999). "Distributed Morphology". In: *Glot International* 4.4, pp. 3–9.

Harris, James W. (1985). "Spanish Diphthongization and Stress: A Paradox Resolved". In: *Phonological Yearbook* 2, pp. 31–45.

Haspelmath, Martin (2011). "The Indeterminacy of Word Segmentation and the Nature of Morphology and Syntax". In: *Folia Linguistica* 45.1, pp. 31–80.

Hauser, Marc D., Noam Chomsky, and W. Tecumseh Fitch (2002). "The Faculty of Language, What Is It, Who Has It, and How Did It Evolve?" In: *Science* 298, pp. 1569–1579.

Hauser, Marc D. and W. Tecumseh Fitch (2003). "What Are the Uniquely Human Components of the Language Faculty?" In: *Language Evolution*. Ed. by Morten H. Christiansen and Simon Kirby. Oxford: Oxford University Press, pp. 158–181.

Hayes, Bruce and Aditi Lahiri (1991). "Bengali Intonational Phonology". In: *Natural Language and Linguistic Theory* 9.1, pp. 47–96.

Heine, Bernd and Tania Kuteva (2002). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Hicks, Jessica Peterson (2006). "The Impact of Function Words on the Processing and Acquisition of Syntax". PhD Thesis. Northwestern University, Evanston, Illinois.

Höhle, Tilman (1999). "An Architecture for Phonology". In: *Slavic in Head-Driven Phrase Structure Grammar*. Ed. by Robert D. Borsley and Adam Przepiórkowski. Stanford, CA: CSLI Publications, pp. 61–90.

Hurford, James R. (1988). *Review of* The Symbolic Species*:* The Co-Evolution of Language and the Human Brain*, by Terrence Deacon*. The Times Literary Supplement, p. 34. URL: http://www.lel.ed.ac.uk/~jim/deaconreview.tls.html.

– (2002). "Expression/Induction Models of Language Evolution: Dimensions and Issues". In: *Linguistic Evolution through Language Acquisition*. Ed. by Ted Briscoe. Cambridge: Cambridge University Press.

Iatridou, Sabine (1990). "About Agr(P)". In: *Linguistic Inquiry* 21.4, pp. 551–577.

Jackendoff, Ray (1997). *The Architecture of the Language Faculty*. Vol. 28. Linguistic Inquiry Monograph. Cambridge, MA: The MIT Press.

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

– (2007). "A Parallel Architecture Perspective on Language Processing". In: *Brain Research* 1145.2-22. DOI: 10.1016/j.brainres.2006.08.111.

Joanisse, Marc F. and Mark S. Seidenberg (1999). "Impairments in Verb Morphology after Brain Injury: A Connectionist Model". In: *Proceedings of the National Academy of Sciences of the United States of America* 96, pp. 7592–7597. DOI: 10.1073/pnas.96.13.7592.

Johnson, Kent (2004). "Gold's Theorem and Cognitive Science". In: *Philosophy of Science* 71.4, pp. 571–592.

Johnston, Trevor A. and Adam Schembri (2007). *Australian Sign Language (Auslan): An Introduction to Sign Language Linguistics*. Cambridge: Cambridge University Press.

Jordens, Peter (2012). *Language Acquisition and the Functional Category System.* Vol. 39. Studies on Language Acquisition. Berlin: Mouton de Gruyter.

Julien, Marit (2002). *Syntactic Heads and Word Formation.* Oxford: Oxford University Press.

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

Klavans, Judith L. (1985). "The Independence of Syntax and Phonology in Cliticization". In: *Language* 61.1, pp. 95–120. URL: http://www.jstor.org/stable/413422.

Klein, Ewan (2000). "Prosodic Constituency in HPSG". In: *Grammatical Interfaces in HPSG.* Ed. by Ronnie Cann, Claire Grover, and Philip Miller. Stanford, CA: CSLI Publications, pp. 169–200.

Koopman, Hilda and Anna Szabolcsi (2000). *Verbal Complexes.* Cambridge, MA: The MIT Press.

Kremers, Joost (2003). "The Noun Phrase in Arabic: A Minimalist Approach". LOT Dissertation Series 79. PhD Thesis. University of Nijmegen.

– (2005). "Adjectival Constructs in Arabic". In: *Linguistische Berichte* 203, pp. 331–348.

– (2009a). "A prosodic approach to particle verbs". In: *Linguistische Berichte* 217, pp. 81–96.

– (2009b). "Recursive Linearization". In: *The Linguistic Review* 26.1, pp. 135–166.

– (2012a). "Arabic Verbal Nouns as Phonological Head Movement". In: *Proceedings of JeNom 4.* Ed. by Gianina Iordăchioaia. Vol. 9. Working Papers of the SFB 732 *Incremental Specification in Context.* Stuttgart University, pp. 73–96. URL: http://elib.uni-stuttgart.de/opus/volltexte/2012/7639/.

– (2012b). "The Syntax of Simultaneity". In: *Lingua* 122.9, pp. 979–1003. DOI: 10.1016/j.lingua.2012.03.008.

– (2013). "Linearization as Repair". In: *Repairs: The Added Value of Being Wrong.* Ed. by Patrick Brandt and Eric Fuß. Berlin: Mouton de Gruyter, pp. 207–236.

– (2014). "The Selection of Marked Exponents in a Parallel Grammar". In: *Lingue e Linguaggio* 2014.1, pp. 103–125. ISSN: 1720-9331. DOI: 10.1418/77002. URL: http://www.rivisteweb.it/doi/10.1418/77002.

– (2015). "Morphology is in the Eye of the Beholder". In: *Linguistische Berichte* 243, pp. 246–294.

Kyle, James G. and Bencie Woll (1985). *Sign Language: The Study of Deaf People and Their Language.* Cambridge: Cambridge University Press.

Ladd, D. Robert (1986). "Intonational Phrasing: The Case for Recursive Prosodic Structure". In: *Yearbook of Morphology* 3, pp. 311–340. URL: http://www.jstor.org/stable/4615402.

Lakoff, George (1972). "The Arbitrary Basis of Transformational Grammar". In: *Language* 48.1, pp. 76–87.

Lapointe, Steven Guy (1980). "A Theory of Grammatical Agreement". PhD Thesis. University of Massachusetts.

Larson, Richard (1988). "On the Double Object Construction". In: *Linguistic Inquiry* 19.3, pp. 335–391.

Lasnik, Howard and Juan Uriagereka (2002). "On the Poverty of the Challenge". In: *The Linguistic Review* 19, pp. 147–150.

Lenerz, Jürgen (1977). *Zur Abfolge nominaler Satzglieder im Deutschen.* Tübingen: Narr.

Leuninger, Helen (2005). "Sign Languages: Representation, Processing, and Interface Conditions". ms. University of Frankfurt.

Leuninger, Helen, Annette Hohenberger, and Elke Menges (2005). "Zur Verarbeitung morphologischer Informationen in der Deutschen Gebärdensprache (DGS)". In: *Gebärdensprachen: Struktur, Erwerb, Verwendung.* Ed. by Helen Leuninger and Daniela Happ. Hamburg: Helmut Buske Verlag, pp. 325–358.

Liddell, Scott K. (1980). *American Sign Language Syntax.* Berlin: Mouton de Gruyter.

Liddell, Scott K. and Robert Johnson (1989). "American Sign Language: The Phonological Base". In: *Sign Language Studies* 64, pp. 197–277.

Lieber, Rochelle (1992). *Deconstructing Morphology: Word Formation in a Syntactic Theory.* Chicago: University of Chicago Press.

Lieber, Rochelle and Sergio Scalise (2006). "The Lexical Integrity Hypothesis in a New Theoretical Universe". In: *Lingue e Linguaggio* 6, pp. 7–32.

Longobardi, Giuseppe (1994). "Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form". In: *Linguistic Inquiry* 25.4, pp. 609–665.

Lüdeling, Anke (2001). *On Particle Verbs and Similiar Constructions in German.* Stanford, CA: CSLI Publications.

Maiden, Martin (2004). "When Lexemes Become Allomorphs – On the Genesis of Suppletion". In: *Folia Linguistica* 38.3–4, pp. 227–256. DOI: 10.1515/flin.2004.38.3-4.227.

– (2005). "Morphological Autonomy and Diachrony". English. In: *Yearbook of Morphology 2004.* Ed. by Geert Booij and Jaap van Marle. Yearbook of Morphology. Springer Netherlands, pp. 137–175. ISBN: 978-1-4020-2899-1. DOI: 10.1007/1-4020-2900-4_6.

Malouf, Robert P. (2000). *Mixed Categories in the Hierarchical Lexicon.* Stanford, CA: CSLI Publications.

Marantz, Alec (2005). "Generative Linguistics within the Cognitive Neuroscience of Language". In: *The Linguistic Review* 22, pp. 429–445.

Matushansky, Ora (2006). "Head Movement in Linguistic Theory". In: *Linguistic Inquiry* 37.1, pp. 69–109.

McCarthy, John (1981). "A Prosodic Theory of Nonconcatenative Morphology". In: *Linguistic Inquiry* 12.3, pp. 373–418.
– (1986). "OCP Effects: Gemination and Antigemination". In: *Linguistic Inquiry* 17.2, pp. 207–263.
McCarthy, John and Alan Prince (1990). "Prosodic Morphology and Templatic Morphology". In: *Perspectives on Arabic Linguistics II*. Ed. by Mushira Eid and John McCarthy. Vol. 72. Current Issues in Linguistic Theory. Amsterdam / Philadelphia: John Benjamins, pp. 1–54.
– (1993). "Generalized Alignment". In: *Yearbook of Morphology*, pp. 79–154. URL: `http://scholarworks.umass.edu/linguist_faculty_pubs/12/`.
– (1995). "Prosodic Morphology". In: *The Handbook of Phonological Theory*. Ed. by John Goldsmith. Cambridge, MA: Blackwell. Chap. 9, pp. 318–366.
– (1996). "Prosodic Morphology 1986". ms. University of Massachusetts, Rutgers University.
Meir, Irit and Wendy Sandler (2008). *A language in Space: The Story of Israeli Sign Language*. New York: Lawrence Erlbaum Associates.
Morgan, James L. and Katherine Demuth, eds. (1996). *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
Müller, Géréon (1999). "Optimality, Markedness, and Word Order in German". In: *Linguistics* 37.5, pp. 777–818.
– (2000). *Elemente der optimalitätstheoretischen Syntax*. Tübingen: Stauffenburg Verlag.
Müller, Stefan (2002). "Syntax or Morphology: German Particle Verbs Revisited". In: *Verb-Particle Explorations*. Ed. by Nicole Dehé et al. Berlin: Mouton de Gruyter, pp. 119–139.
– (2013). "Unifying Everything". In: *Language* 89.4, pp. 920–950. URL: `http://hpsg.fu-berlin.de/~stefan/Pub/unifying.html`.
Nam, Hosung (2007). "Syllable-level Intergestural Timing Model: Split-gesture Dynamics Focusing on Positional Asymmetry and Moraic Structure". In: *Laboratory Phonology 9*. Ed. by Jennifer Cole and José Ignacio Hualde. New York: Mouton de Gruyter, pp. 483–506.
Neeleman, Ad and Fred Weerman (1993). "The Balance Between Syntax and Morphology: Dutch Particles and Resultatives". In: *Natural Language and Linguistic Theory* 11, pp. 433–475.
– (1999). *Flexible Syntax: A Theory of Case and Arguments*. Dordrecht: Kluwer Academic Publishers.
Neidle, Carol et al. (2000). *The Syntax of American Sign Language: Functional Categories and Hierarchical Structure*. Cambridge, MA: The MIT Press.

Nespor, Marina, Mohinish Shukla, et al. (2008). "Different phrasal prominence realizations in VO and OV languages". In: *Lingue e Linguaggio* 7.2, pp. 139–168. ISSN: 1720-9331. URL: http://hdl.handle.net/10281/6919.

Nespor, Marina and Irene Vogel (1986). *Prosodic Phonology*. Dordrecht: Foris.

Nunes, Jairo (2001). "Sideward Movement". In: *Linguistic Inquiry* 32.2, pp. 303–344.

– (2004). *Linearization of Chains and Sideward Movement*. Cambridge, MA: The MIT Press.

O'Reilly, Randall C. et al. (2012). *Computational Cognitive Neuroscience*. Wiki Book, 1st Edition. URL: http://ccnbook.colorado.edu.

Partee, Barbara (1999). "Focus, Quantification, and Semantics-Pragmatics Issues". In: *Focus: Linguistic, Cognitive, and Computational Perspectives*. Ed. by Peter Bosch and Rob van der Sandt. Cambridge: Cambridge University Press, pp. 213–231.

Partee, Barbara, Alice ter Meulen, and Robert Wall (1993). *Mathematical Methods in Linguistics*. Dordrecht: Kluwer.

Peirce, Charles Sanders (1897, 1903). "Logic as Semiotic: The Theory of Signs". In: *The Philosophical Writings of Peirce*. Ed. by Justus Buchler. (1955). New York: Dover, pp. 98–119.

Pereltsvaig, Asya (2008). "Split Phrases In Colloquial Russian". In: *Studia Linguistica* 62.1, pp. 5–38.

Pfau, Roland (2001). "Typologische und Strukturelle Aspekte der Negation in Deutscher Gebärdensprache". In: *Gebärdensprachlinguistik 2000: Theorie und Anwendung*. Ed. by Helen Leuninger and Karin Wempe. Vol. 37. Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser. Hamburg: Signum, pp. 13–31.

– (2002). "Applying Morphosyntactic and Phonological Readjustment Rules in Natural Language Negation". In: *Modality and Structure in Signed and Spoken Languages*. Ed. by Richard Meier, Kearsy Cormier, and David Quinto-Pozos. Cambridge: Cambridge University Press. Chap. 11, pp. 263–295.

– (2008). "The Grammar of Headshake: A Typological Perspective on German Sign Language Negation". In: *Linguistics in Amsterdam* 1, pp. 34–71.

Pfau, Roland and Josep Quer (2002). "V-to-Neg raising and negative concord in three sign languages". In: *Rivista di Grammatica Generativa* 27, pp. 73–86.

– (2007). "On the syntax of negation and modals in Catalan Sign Language and German Sign Language". In: *Visible Variation. Comparative Studies on Sign Language Structure*. Ed. by Pamela M. Perniss, Roland Pfau, and Markus Steinbach. Trends in Linguistics. Mouton de Gruyter, pp. 129–161.

Pinker, Steven (1999). *Words and Rules*. New York: Basic Books.

Poeppel, David and David Embick (2005). "Defining the Relation between Linguistics and Neuroscience". In: *Twenty-First Century Psycholinguistics: Four*

Cornerstones. Ed. by Anne Cutler. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 103–118.
Pollard, Carl J. and Ivan A. Sag (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
Pollock, Jean-Yves (1989). "Verb Movement, Universal Grammar and the Structure of IP". In: *Linguistic Inquiry* 20.3, pp. 365–424. URL: http://www.jstor.org/stable/4178634.
Prince, Alan and Paul Smolensky (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, Mass.: Blackwell.
Pullum, Geoffrey K. and Barbara C. Scholz (2002). "Empirical Assessment of Stimulus Poverty Arguments." In: *The Linguistic Review* 19.1-2, pp. 9–50.
Pulvermüller, Friedemann (2003). *The Neuroscience of Language: On Brain Circuits of Words and Serial Order*. Cambridge: Cambridge University Press.
– (2010). "Brain Embodiment of Syntax and Grammar: Discrete Combinatorial Mechanisms Spelt out in Neuronal Circuits". In: *Brain & Language* 112, pp. 167–179.
Richards, Marc (2004). "Object Shift and Scrambling in North and West Germanic: A Case Study in Symmetrical Syntax". PhD Thesis. University of Cambridge.
Rizzi, Luigi (1997). "On the fine structure of the left-periphery". In: *Elements of Grammar*. Ed. by Liliane Haegeman. Dordrecht: Kluwer, pp. 281–337.
Ross, John Robert (1967). "Constraints on Variables in Syntax". PhD Thesis. Cambridge, MA: The MIT Press.
Sadock, Jerrold M. (1992). *Autolexical Syntax*. Chicago: University of Chicago Press.
– (2012). *The Modular Architecture of Grammar*. Cambridge: Cambridge University Press.
Samek-Lodovici, Vieri (2005). "Prosody-syntax interaction in the expression of focus". In: *Natural Language and Linguistic Theory* 23, pp. 687–755.
Sandler, Wendy (1986). "The Spreading Hand Autosegment of American Sign Language". In: *Sign Language Studies* 50, pp. 1–28.
– (1989). *Phonological Representation of the Sign: Linearity and Nonlinearity in American Sign Language*. Dordrecht: Foris.
– (1990). "Temporal Aspects and ASL Phonology". In: *Theoretical Issues in Sign Language Research*. Ed. by Susan D Fischer and Patricia Siple. Vol. I. Chicago: University of Chicago Press, pp. 7–36.
– (2011). "Prosody and Syntax in Sign Languages". In: *Transactions of the Philological Society* 108.3, pp. 298–328.
Sandler, Wendy and Diane Lillo-Martin (2006). *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press.
Scheer, Tobias (2008). "Why the Prosodic Hierarchy Is a Diacritic and Why the Interface Must Be Direct". In: *Sounds of Silence: Empty Elements in Syntax and*

Phonology. Ed. by Jutta Hartmann, Veronika Hegedűs, and Henk van Riemsdijk. Amsterdam: Elsevier, pp. 145–192.

Schütze, Carson T. (1994). "Serbo-Croatian Second Position Clitic Placement and the Phonology-Syntax Interface". In: *Papers on Phonology and Morphology*. Ed. by Andrew Carnie and Heidi Harley. Vol. 21. MIT Working Papers in Linguistics. Cambridge, MA: The MIT Press, pp. 373–473. URL: http://www.ai.mit.edu/projects/dm/schutze-secondposition.pdf.

Selkirk, Elisabeth (1981). "On Prosodic Structure and its Relation to Syntactic Structure". In: *Nordic Prosody II*. Ed. by Thorstein Fretheim, pp. 111–140.

– (1982). *The Syntax of Words*. Cambridge, MA: The MIT Press.

– (1984). "Phonology and Syntax: The Relation between Sound and Structure". PhD Thesis. Cambridge, MA: The MIT Press.

– (1995). "Sentence Prosody: Intonation, Stress and Phrasing". In: *The Handbook of Phonological Theory*. Ed. by John Goldsmith. Cambridge, MA: Blackwell, pp. 550–569.

– (1996). "The Prosodic Structure of Function Words". In: *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Ed. by James L. Morgan and Katherine Demuth. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 187–214.

– (2005). "Comments on Intonational Phrasing in English". In: *Prosodies*. Ed. by Sónia Frota, Marina Cláudia Vigário, and Maria João Freitas. Berlin: Mouton de Gruyter, pp. 11–58.

Siebert, Susann (1999). *Wortbildung und Grammatik: Syntaktische Restriktionen in der Struktur komplexer Wörter*. ("Word Formation and Grammar: Syntactic Restrictions in the Structure of Complex Words"). Tübingen: Max Niemeyer Verlag.

Siloni, Tal (2001). "Construct States at the PF Interface". In: *Linguistic Variation Yearbook* 1, pp. 229–266.

Skinner, Burrhus Frederic (1957). *Verbal Behavior*. Copley Publishing Group.

Skwarski, Filip (2009). "Accounting for Underlying Forms in HPSG Phonology". In: *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar*. Ed. by Stefan Müller. Stanford, CA: CSLI Publications, pp. 318–337. URL: http://www.stanford.edu/group/cslipublications/cslipublications/HPSG/2009/.

Soderstrom, Melanie et al. (2003). "The prosodic bootstrapping of phrases: Evidence from prelinguistic infants". In: *Journal of Memory and Language* 49, pp. 249–267.

Struckmeier, Volker (2011). "Syntax und Prosodie von Scramblingbewegungen im Deutschen: Zur komplexen Architektur grammatischer Beschreibungen". Habilitationsschrift Universität zu Köln.

Struckmeier, Volker and Joost Kremers (2014). "On the properties of attributive phrases in Germanic (and beyond)". In: *Adjectives in Germanic and Romance*. Ed. by Petra Sleeman, Freek Van de Velde, and Harry Perridon. Amsterdam / Philadelphia: John Benjamins, pp. 149–169. ISBN: 9789027255952. DOI: 10.1075/la.212.

Surányi, Balász (2005). "Head Movement and Reprojection". In: *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae* Tomus XXVI. Budapest: ELTE, pp. 313–342.

Thorpe, Simon J. (2003). "Localized versus Distributed Representations". In: *The Handbook of Brain Theory and Neural Networks*. Ed. by Michael A. Arbib. 2nd edition. Cambridge, MA: The MIT Press, pp. 643–646.

Tomasello, Michael (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.

Truckenbrodt, Hubert (1995). "Phonological Phrases: Their Relation to Syntax, Focus and Prominence". PhD Thesis. Cambridge, MA: The MIT Press.

– (2006a). "On the Semantic Motivation of Syntactic Verb Movement to C in German". In: *Theoretical Linguistics* 32.3, pp. 257–306.

– (2006b). *Phrasal Stress*. In: *Encyclopedia of Language & Linguistics*. Vol. 9. Ed. by Keith Brown. Oxford: Elsevier, pp. 572–579.

– (2007). "The Syntax-Phonology Interface". In: *The Cambridge Handbook of Phonology*. Ed. by Paul De Lacy. Cambridge: Cambridge University Press, pp. 435–456.

Tseng, Jesse (2008). "The representation of syllable structure in HPSG". In: *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*. Ed. by Stefan Müller. CSLI Publications, pp. 234–252. URL: http://www.stanford.edu/group/cslipublications/cslipublications/HPSG/2008/.

Vermeerbergen, Myriam, Lorraine Leeson, and Onno Crasborn, eds. (2007). *Simultaneity in Signed Languages: Form and Function*. Amsterdam / Philadelphia: John Benjamins.

Westermann, Gert and Nicolas Ruh (2012). "A Neuroconstructivist Model of Past Tense Development and Processing". In: *Psychological Review*. DOI: 10.1037/a0028258.

Wilbur, Ronnie B. (1994). "Foregrounding Structures in American Sign Language". In: *Journal of Pragmatics* 22.6, pp. 647–672.

– (2000). "Phonological and Prosodic Layering of Nonmanuals in American Sign Language". In: *The Signs or Language Revisited*. Ed. by Karen Emmorey and Harlan Lane. Mahwah, NJ: Lawrence Erlbaum Publishers, pp. 215–244.

Wilbur, Ronnie B. and Cynthia Patschke (1998). "Body Leans and the Marking of Contrast in American Sign Language". In: *Journal of Pragmatics* 30, pp. 275–303.

Yoon, James (1996). "Nominal Gerund Phrases in English as Phrasal Zero Derivations". In: *Linguistics* 34, pp. 329–356.

Zeller, Jochen (1999). "Particle Verbs, Local Domains, and a Theory of Lexical Licensing". PhD Thesis. University of Frankfurt, Frankfurt/Main.

– (2001). *Particle Verbs and Local Domains*. Amsterdam / Philadelphia: John Benjamins.

Zeshan, Ulrike, ed. (2006). *Interrogative and Negative Constructions in Sign Languages*. Nijmegen: Ishara Press.

Zwart, Jan-Wouter (2001). "Syntactic and Phonological Verb Movement". In: *Syntax* 4.1, pp. 34–62.