Two phonologies

Harry van der Hulst

1. Introduction

In this article, I will develop a general perspective on the relationship between phonology and the other modules of grammar. One conclusion of this exploration is that we need two types of phonology (or two phonological levels), one dealing with phonotactic structure (phonology, "grammar phonology" proper) and the other dealing with utterance structure (utterance phonology, phonetics), which, as I will suggest, is distinct from phonetic implementation. Crucially, I will suggest that the constituents of prosodic phonology, as well as the phenomena that both motivate and depend on them, need to be parceled out over these two phonological levels, some belonging to the phonotactic level whose structure closely follows morphotactics (morphology and syntax) and some belonging to the utterance level whose domains are driven by the *metrical grid*. I will also include a brief discussion of intonation and focus.

2. Kinds of phonology

There appear to be numerous "phonologies" all dealing with the sound structure of languages (where "sound structure" ranges from the structure of actual sound sequences to cognitive structures that underlie the storage and retrieval, as well as production and perception of these sound sequences). In some cases, these phonologies are different, competing theories (historically consecutive or co-temporal) that intend to cover the same empirical domain, thus overlapping in many ways, yet differing in numerous details (e.g., Fischer-Jørgensen 1975; Anderson 1985). In other cases we are dealing with subtheories that are meant to be complementary. Subtheories can be complementary either by characterizing different aspects of *one* phonological level (e.g., autosegmental phonology, metrical phonology, lexical phonology), or they can be complementary in the sense of characterizing different levels. The difference between these two types of complementarity is not always obvious because of the many ways in which

the term *level* is being used. In this section, I will argue in favor of recognizing two phonological levels, using this term to refer to any autonomous analysis of linguistic expressions (as in Goldsmith 1993). Levels, in this sense, are not derivationally related, as for example the underlying and the surface *representation* in standard generative phonology (SPE; Chomsky and Halle 1968). We must be aware of the fact that in the phonological literature the term level is not only used as a substitute for representation but also for *layer* or *domain* (in a hierarchical construct, e.g. "word level", "syllable level"). In this article, I will be careful in distinguishing between the notions level (autonomous analysis), representation (formal expression built from a primitive vocabulary in accordance with wellformedness constraints) and layer or domain (hierarchy or grouping, part of a representation); if a distinction between underlying and surface representations is made, I will refer to these as derivational *stages*. A fourth term, *plane*, will be used to refer to different, simultaneous dimensions of a representation (such as the metrical plane and the autosegmental plane) although, at this point, I do not wish to acknowledge that phonological representation are indeed multiplanar.

2.1. The interpretations of the notion "Dual Articulation" – sequential and parallel

Grammar forms the link between sound and meaning. As such, grammars generate structures that mediate between sound structures and meaning structures. However, grammatical structures appear to display a duality in the sense of having a (morpho)syntactic or *morphotactic* part and a phonological, or more properly *phonotactic* part. The idea of *Dual Articulation* (or *Dual Patterning*) can be found in the work of Martinet (1960: 22-24), and Hockett (1960), and can be traced back even further (see Anderson 1992). In what Martinet called the *first* articulation (morphotactics), minimal meaningful units (morphemes) are combined into larger meaningful structures so that an infinite array of situations can be described, while in the *second* articulation (phonotactics), meaningless units, i.e. phoneme sequences are combined into units, that, at the lowest morphotactic layer of morphemes, are (largely arbitrarily) linked to meanings. In this view then, phonotactics and morphotactics are "sequential", the first articulation being built from the units that the second articulation produces.[1] Phonotactics

[1] In this respect, it would have been better to call phonology the first and morphosyntax the second articulation. I suppose that Martinet (1960) called morphosyntax the first articulation because it deals with the more important notion of producing meaningful linguistic expressions, while phonology merely delivers the form aspect of the smallest meaningful units.

and morphotactics together constitute the *skeleton* of language, and grammar is the system of basic units and combination rules that accounts for the organization of that skeleton. Phonetics provides this skeleton with its perceptible bodily appearance, while semantics breathes life and mind into this body. Like phonotactics and morphotactics, phonetics and semantics are complementary as well, since phonetics interprets the structure created by phonology, while semantics interprets the structures created by morphosyntax.[2]

Both phonotactics and morphotactics are combinatorial systems, consisting of finite sets of basic building blocks (*primitives*) and combinatorial rules (*constraints*). At the level of phonotactics, the basic units are *features*[3] which "combine into," or "are properties of" phonemes which in turn combine into morphemes, while the morphotactic basic units are the category labels like Noun, Verb, etc.[4] The kinds of combinatorial constraints in both domains are largely the same which brings me to the notion *Structural Analogy*.

Structural Analogy refers to the structural resemblance between representations in different levels or stages, layers or planes (within levels). We note, for example, analogies between phonotactics and morphotactics as a whole but also between different hierarchical layers within the structures that each of these modules characterizes. In the end, the idea of structural analogy can be taken to mean that the same kind of structure is likely to repeat itself over and over again (see Anderson 2006):

> *The Structural Analogy Assumption*
> Minimize (more strongly, eliminate) differences between levels (modules, etc.) that do not follow from a difference in alphabet or from the nature of the relationship between the levels (etc.) concerned.

Limiting ourselves here to the analogy between phonotactics and morphotactics as a whole, the idea is that both structures are based on identical principles (such as binarity and headedness), and that their differences are due to the fact that phonotactics must primarily accommodate phonetic matter, while morphotactics must primarily accommodate semantic matter (van der Hulst 2000, 2006b,

---

[2] Instead of speaking of interpretation it is also common to use the term *implementation*, especially on the phonetic side. Below, I will suggest using both terms with different meanings.

[3] This term is used in SPE, following the Roman Jakobson's version of Prague School phonology. In other framework the primes are called *components* (Dependency Phonology; Anderson and Ewen 1987; van der Hulst 1989, 2005), *elements* (Government Phonology; Kaye, Lowenstamm, and Vergnaud 1985, 1990), or *particles* (Schane 1984).

[4] These labels can also be construed as features or combinations of features.

2008b,c).[5] Indeed, semantics is largely irrelevant at the level of phonotactic form, at least to the extent that phonotactic form is not semantically motivated, the best known case of motivation being iconicity. Whether and, if so, to what extent morphotactics accommodates phonetics remains to be seen.

The term "phonology" traditionally refers to two different aspects, namely the representational and the derivational aspect. The former regards the characterization of wellformedness (for which I use the term *phonotactics*; essentially, a set of primitives and combinatorial constraints), while the latter largely regards "repairs", i.e. operations that make changes in representations that violate a constraint.[6] Likewise, morphosyntax can be separated into two systems, i.e. constraints (morphotactics) and operations (the latter comprising, perhaps, inflection and transformations).[7] Again, we notice the architectural analogies. Since the main discussion here is about the "tactic" aspect of each module, I have referred thus far to morphological and syntactic (morphosyntactic) structure as *morphotactics*, analogously to the term *phonotactics*. If we were to accept Martinet's (1960) idea that the first articulation is stacked on top of the second articulation (as layers), we can say that these two systems together form the *tactic level*. The connection between the two tactic layers is defined by linking phonotactic constructs to the category labels that are combined in morphotactics; one hierarchy is built on top of the other:

(1)

Phonotactics	Morphotactics

Each of these two tactic subsystems characterizes a series of hierarchical layers that, in this view, form a continuum of feature structures. In phonotactics, we start with features which form phonemes in accordance with the "segment structure conditions" (as they used to be called in generative phonology). These segments then combine into morphemes in accordance with a set of "morpheme structure conditions". In morphotactics, we start with a basic set of category labels (features or feature combinations) which combine into word labels in accordance with morphological constraints and into "larger-than-words" units (phrases, sentences) in accordance with the syntactic constraints.

---

[5] I use the qualifier "primarily" rather than "exclusively" because, as we will see, phonetics and semantics may not be as complementary as suggested thus far.

[6] Singh (1987); Stewart (1983); Paradis (1988); Calabrese (1995, 2005).

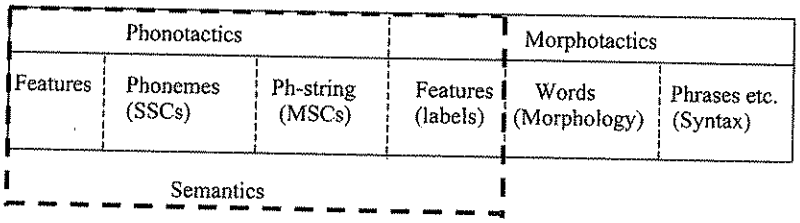
[7] In some approaches, the notion derivation is reduced to a minimum or even eliminated by adopting purely representational means to deal with the variability of linguistic units, such as allomorphy; see van der Hulst (2005, 2008b).

(2)

Phonotactics			Morphotactics		
Features	Phonemes (SSCs)	Morphemes (MSCs)	Features (labels)	Words (Morphology)	Phrases etc. (Syntax)

In this model, morphemes are pivotal units. They connect a phonotactic *form* to a morphotactic *label* and, in fact, to a semantic or *conceptual structure*. In this sense, it is not appropriate, as done in (2), to call the entity that MSCs characterize a morpheme because it is merely the phonotactic side (the phoneme string) of the morpheme, just like the label could be called the *categorial* side of the morpheme. If we include the semantic side, we arrive at the following diagram in which the bold broken lines enclose the three types of information that constitute a morpheme (i.e. its phonotactic structure, its category label and its meaning):

(3)



There are, however, objections to this simple stacking model of phonotactics and morphotactics which is nonetheless very popular in introductory textbooks.

At the phonotactic level, one could object that it is not obvious that bound morphemes are subject to phonotactic constraints, or, to put it differently, that the phonotactic constraints that are usually formulated truly pertain to units of this size. The grounds to doubt the necessity of MSCs, or their status as *morpheme* structure constraints, is directly related to a second questionable aspect of the diagram, namely its strict sequential (stacking) design. It would seem that not morphemes, but rather words (including both simplex words and a subset of the complex words) are governed by the laws of phonotactics. This would make sense since words, and not morphemes, are meant to be pronounced. We would not notice the difference so much when we focus on free morphemes which are both morphemes and words. However, when we consider complex words – in particular, words that contain affixes – we know that those entities cannot, in many cases, be "as is" concatenations of the phonological strings that make up their morphemes. As every linguist knows, morpheme concatenations call for phonological repairs all the time (assimilations, deletions, insertions, etc.).

It is precisely this particular overlap between phonotactics and morphotactics that gives rise to *allomorphy*.[8] The phonotactic constraints that govern at least a subset of the complex words (i.e. roughly those formed by level I affixes in English) seem to be, at least in part, the same as those that govern simplex words. This means that at least a subset of the complex words behave like simplex words from a phonotactic point of view. It seems clear, and this brings us back to the first objection, that the word-domain phonotactic constraints that govern all simplex words and a subclass of complex words subsume (or at least contain most of) the constraints that were formerly held to be morpheme-domain constraints (including the so-called segment structure conditions). To the extent that constraints on complex words are the same as constraints on simplex words, there is no need to state these constraints twice, i.e. as MSCs and as repair rules that apply in derived environments.[9] Besides the possibility of constraints that are truly morphemic[10], there is another reason for believing that the phonotactics of simplex words and complex words do not need to be *fully* identical, which is that a subclass of the repair rules that apply in complex words seems limited to applying to complex words (*derived environments*) only. Therefore, the more crucial point being made here is that the domains of phonotactics and morphotactics *overlap* in a subset of complex words, namely those whose phonology involves repairs that imply the existence of phonotactic constraints holding above the level of the morpheme.

This finding is incompatible with the strict sequential conception implied in diagram (3) which does seem to imply that, as morphemes are combined into larger constructs, their phonological material is concatenated "as is", i.e. unchanged, which would deny that allomorphy exists. Let us also note here that the stacking model suggests that, to the extent that phonological strings of words and larger units can be said to be hierarchically organized, the hierarchical arrangement is the one that is created by the morphotactics. As just noted, at least the first consequence is invalid; allomorphy does exist and this implies that units larger than morphemes, at least a subclass of the complex words,

---

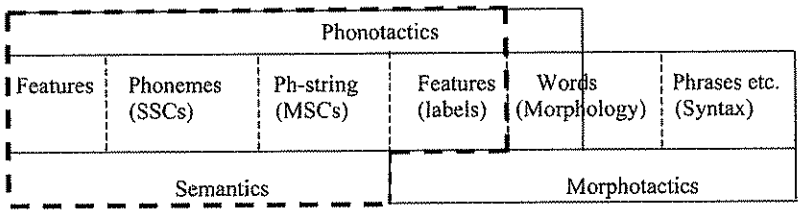
[8] Other adjustments involve predictably adding features which causes (neutralizing or non-neutralizing) allophonic variation. However, these are not dependent on morphological concatenation.

[9] This issue regards the so-called "Duplication Problem"; see Kenstowicz and Kisseberth (1977), van der Hulst (2008b).

[10] It seems likely that a residue of morpheme domain constraints remains (such as constraints that differentiate between bound and free morphemes but perhaps also constraints that are true of simplex words i.e. morphemes but not of complex words); see Paradis and Prunet (1993) and Booij (1999).

form phonotactic domains governed by constraints that drive the allomorphic variation. I will return to the second consequence later on. It would seem, then, that the workspace of phonotactics and morphotactics is overlapping rather than strictly sequential:

(4)



The line going through 'morphology' allows phonotactics to govern a subclass of complex words, notably what used to be called "level I morphology". Traditional level ordered models would have problems explaining why phonotactics skips level II (Borowsky 1993), while being relevant to inflected words (which generally display significant allomorphy) which are mostly assumed to form an even later level (or happen in the syntax). I will need to return to this point below in conjunction with the question whether phonotactics extends all the way up to the sentence domain.

An important aspect of identifying the domain *word* (rather than, or in addition to the morpheme domain) as a relevant phonotactic unit was the recognition of the unit *syllable* as a crucial subword phonotactic domain. Roughly, it was shown in several publications in the post-SPE era that word level phonotactic constraints largely consisted of constraints on the syllable domain with the result that wellformed words could be largely characterized as combinations of wellformed syllables.[11] However, phonotactic constraints that make reference to larger-than-syllable domains remain necessary. For one thing, word-edge syllables often show extra or different phonotactic options. When, for independent reasons, it was then argued that syllables group into units called feet,[12] it became clear that this unit too could be the domain of phonotactic constraints. These developments led to the inclusion into word phonotactics of a whole new set of constraints, namely constraints that characterize a hierarchical (syllabic-and-metrical) organization and, in terms of this hierarchy, constraints on combinations of segments and the location of accents, i.e. stress.

As already suggested earlier, the question whether the overlap only concerns the word domain could now be raised, given that the two tactic systems overlap.

---

[11] Although this point was not new as such; compare Pulgram (1970).

[12] Anderson and Jones (1974); Liberman and Prince (1977).

Since at least Trubetzkoy (2001) made the distinction between word phonology and sentence phonology, it has been known that phonology does *not* stop at the word level. There are two aspects to sentence phonology which became the focus of attention in the 1970s and 1980s due to the work of linguists like Selkirk (e.g. 1978, 1980, 1984), Rotenberg (1978), Pullum and Zwicky (1988), Nespor and Vogel (1986) and Kaisse (1985), among others. On the one hand, when words are strung together into phrasal units, segmental adaptations occur creating varying forms of words, just like repair rules create allomorphy within words. Articles (in languages such as English, French, Spanish, etc.) can take a different form depending on whether the following noun begins with a vowel or consonant, nouns can take different forms in different syntactic contexts (cf. mutation in Welsh), consonantal assimilation may apply across word boundaries, tonal or rhythmic properties of words may change depending on context, and so on. There are, as it would seem, phonotactic constraints that hold over larger domains, as well as repairs that come into action when the concatenation of words leads to violations of these constraints. On the other hand, and this questioned another aspect of the stacking model mentioned above, it would seem that the domains that are relevant for at least some of these phenomena were argued *not* to be syntactic.[13] A typical example of the non-isomorphy of the morphotactic structure and the phonotactic structure involves so-called clitics, i.e. the "little" words that seem to fuse with other words into a non-syntactic (hence presumably phonotactic) unit in a manner that disrespects their syntactic grouping:

(5)

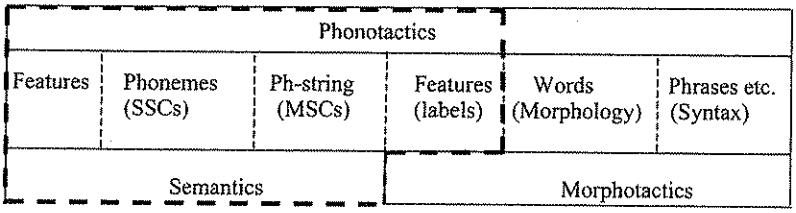
In (5) *he* and *is* merge into a unit which, in most accounts, is not a syntactic unit.[14] The relevant phonotactic structure came to be called the prosodic structure (Selkirk 1978, 1980; Nespor and Vogel 1986).

Thus, instead of the stacking or even partially overlapping model, it was now recognized that a linguistic expression has two largely parallel tactic organizations, one phonotactic and the other morphotactic:

[13] This, of course, is always dependent on the syntactic theory or analysis that one adopts in specific cases.

[14] Categorial models of syntax are more flexible, however; see Steedman (2000).

(6)



As mentioned, at least some of the phenomena that motivated this phonotactic or prosodic structure above the word level had of course been identified long before and proposals had been made for some kind of prosodic grouping. Pulgram (1970) proposes a series of domains including "nexus" and "cursus" domains (both above the word layer) and Abercrombie (1967, section 5; and Lahiri and Plank 2007) pointed to a rhythmic organization which groups syllables (across word boundaries) from one stressed syllable up to the next stressed syllables forming the *Abercrombian* foot (Abercrombie 1967). However, Selkirk (1978) articulated a very specific hierarchy of several domains:

(7) Syllable – Foot – Word – Clitic Group – Phonological phrase – Intonation phrase – Utterance

Nespor and Vogel (1986) then offered a detailed study of various processes that appear to be sensitive to these domains.[15] Not everyone was convinced that *all* these processes motivated a non-isomorphic hierarchy and Kaisse (1985), for example, maintained that many of them could be handled with direct reference to syntactic structure. Those that could not, she termed "fast speech rules".

To avoid the appearance that phonotactics has a *larger* domain than morphotactics, as well as to properly express that semantic structure and morphotactic structure are parallel rather than stacked, I will replace (6) by the following diagram which more properly expresses that all three dimensions of grammar have the same scope.[16]

[15] This is not always so easy to see because syntactic structures differ from theory to theory.

[16] This reminds us of Jackendoff (2002) who proposes a parallel architecture. I will discuss some specific aspects of his model below.

(8)

morphemes[17]	words ⋮ phrases etc.
Phoneme string	Phonotactic structure
Category label	Morphotactic structure
Conceptual structure	Semantic structure

As in the previous diagram, (8) maintains a separation between morphemes, words and larger units. Each morpheme *arbitrarily* (or *conventionally*) combines a chunk of form and a chunk of meaning under a categorial label. Both chunks can be simplex (one primitive) or complex.[18] If complex, wellformedness is not an issue at the morpheme level[19], assuming that no illformed structure will make it pass the constraints that bear on the word domain. All structures at these different levels are subject to constraints and repairs[20]. It seems plausible to me that the division between word and larger-than-word domains runs all the way through the three levels and corresponds to word and sentence subsystems at each level (van der Hulst 2005, 2008b). At the phonological level, this distinction might be thought of, for the moment, as one between word and sentence phonology, or lexical and post-lexical phonology. At the categorial level, the distinction corresponds to the traditional distinction between morphology and syntax (with inflection belonging to one or the other depending on specific views). Finally, in semantics, the distinction is known as lexical semantics and sentence semantics. I will return to these divisions, especially the one in phonology, below.

At this point, the notion "dual articulation" can be seen in a new light. Instead of understanding this notion in the sense of phonotactics and morphotactics being stacked (the sequential or layered view), we now have a complete parallelism between the two articulations, each constituting a different *level* of analysis.

---

[17] Depending on different theories, it is or is not possible that a morpheme lacks one of those levels.

[18] As mentioned, categorial labels, too, can be simplex or complex if one assumes a system of features that characterize the part of speech labels and other morphosyntactic labels.

[19] This is what is now often called "Freedom of the Base"; see Prince and Smolensky (1993).

[20] Some theories prevent the use of repairs by using disjunctive specification in the lexicon, essentially prelisting the allomorphs and different words, accompanied by appropriate insertion frames (see van der Hulst 2008c for discussion).

2.2. Two phonologies?

The set of phonotactic constraints covers both segmental constraints (segment structure constraints) and constraints on phoneme or tonal combinations within and across word boundaries as well as constraints on accents and rhythm. In addition, there are constraints on the hierarchical structure itself which, in metrical phonology, subsume matters of accent and rhythm.

McMahon (2007) sees the constraints on segments and their combinations as fundamentally different from constraints on tonal, accentual and rhythmic properties, i.e. those that are traditionally called suprasegmental or prosodic. According to her, segmental phonology is just the segments and local processes or rules that affect them. Prosody, she says, includes syllable organization and anything higher, as well as processes such as vowel harmony and tonal phonology. She also includes intonation in prosody. She supports the separation between segmental and suprasegmental phonology with considerations from acquisition, language change and neurology. If I were to make the distinction between segmental and suprasegmental aspects of phonology, I would regard some form of syllable structure as belonging to the segmental layer, precisely because, as stated earlier, a significant subset of the constraints on the wellformedness of the segmental strings crucially are syllable structure constraints.[21] Also, I think that processes like vowel harmony fall within the segmental phonology. In short, even though I like the idea of recognizing two phonologies, I would separate the class of phonological phenomena in a rather different way. Luckily, some of the evidence that McMahon uses to support her separation will apply to my separation as well. More crucially to this article, however, without denying that, what I would call a "planar" distinction can be made between segmental and suprasegmental properties, I believe that there is another distinction that can be made which is even more fundamental, namely a distinction between two phonological *levels*.

Before I elaborate this point, it is important to avoid confusion with another way of talking about different phonologies, namely one that is based on differences in domains. I have suggested that a division between word and sentence phonology may be useful. In its modern guise, this distinction is known as the distinction between *lexical* and *post-lexical phonology*. McMahon's (2007) distinction between segmental and suprasegmental phonology is orthogonal to this domain distinction since within both types of domains, we could distinguish between segmental and suprasegmental (or prosodic) plane.

---

[21] This claim is not undisputed. See Steriade (1999), and Blevins (2003).

It would seem that the distinction between lexical and post-lexical domain is one that refers to different layers (like the distinction between word and sentence phonology) within the – or a – phonological level. However, in van der Hulst (2003), I argued that the stacking view of lexical and post-lexical is *not* correct. Anticipating the distinction between two phonological levels that I will advocate below, I suggested that post-lexical phonology, as it is commonly understood does *not* regard layers of structure above the word but rather regards the whole phonological spectrum from feature to utterance. Understood in this way, lexical phonology and post-lexical phonology regard two different *levels* of representation, at least with reference to word domains (since lexical phonology does not bear on phrasal and larger units). This view is different from the common understanding that the lexical system produces wellformed *phonological words* (including syllable and foot structure) which are then organized into phonological phrases and so on, the post-lexical phonology picking up from where the lexical phonology has left off. This serial view of the lexical and post-lexical phonological systems has been around in generative grammar ever since Selkirk and after her Nespor and Vogel developed and motivated a prosodic organization (despite the fact that everyone acknowledged that post-lexical phonology could reach back to the lowest layers inside words as well).

What reasons did I have for saying that at least up to the word domain there are two distinct organizations? To begin with, it is evident that there are two distinct notions of syllabic organization. This "duality of structure" (giving rise to what I called *structure paradoxes*) leads to terminology such as "phonological" syllables (relevant to phonotactics) as opposed to "phonetic" syllables (relevant to details of phonetics or implementation). Most striking becomes the discrepancy if one embraces theories of phonotactic structure that employ the principles of Government Phonology due to which words like *strike* easily end up as trisyllabic whereas there seems to be no motivation at all to assume that at the phonetic level. As mentioned in the article, I also referred to structure paradoxes regarding foot structure. The notion "foot" that is relevant to the so-called *call chant* in Dutch (Gussenhoven 1993) is not the same as the one that is needed at the lexical level for the purpose of calculating primary accent (van der Hulst 1984; Kager 1989).

If the prosodic structure of sentences is not built *on top of* a word structure, the alternative is that a complete post-lexical structure is erected from the ground up. This phonological structure may not even start out with phonemes (or elements) but rather might be gesture-based.[22] This means that we need to replace (8) by (9):

[22] *Gestures* in the sense of Browman and Goldstein (1986).

(9)

morphemes	word	phrases etc.
"Gestural score"	Utterance structure	
Phoneme string	Phonotactic structure	
Category label	Morphotactic structure	
Conceptual structure	Semantic structure	

Here, I am using the term "utterance-structure" instead of "post-lexical" structure because it is now clear that the terms lexical and post-lexical suggest a false dichotomy. Whereas "lexical phonology" can still be used for phonotactic phonology at the word layer at *both* phonological levels, the *whole* utterance structure as well as the larger-than-word structure at the phonotactic level are in a sense post-lexical.[23]

Interestingly, the idea that there are indeed two complete phonologies, as expressed in (9), was taken to be self-evident in Nespor and Vogel (1986) who show that both at the word and phrasal level, we encounter phonological phenomena that are sensitive to lexical or morpho-syntactic information and structure as well as phenomena that are blind to such information and indeed sensitive to domains of a different sort (Nespor and Vogel 1986: 33). I thus propose to abandon the lexical/post-lexical dichotomy and replace it by the phonotactic and utterance dichotomy. Much of what has been called post-lexical is located at the utterance level. As for prosodic phonology in the sense of Selkirk as well as Nespor and Vogel, it remains to be seen how much of it is part of the utterance level. I will turn to that question below.

We thus end up with a parallel view of two phonologies, but not the one that is suggested by McMahon (2007), because both phonologies have segmental and suprasegmental (or prosodic, hierarchical) aspects, however one defines these notions precisely. The distinction drawn here is also different from the lexical

[23] I realize that the distinction between lexical and post-lexical phonology was never meant to match the word – sentence distinction. Yes, lexical phonology would be confined to units that are contained in the lexicon, and this would mostly be words (and little work has ever been done on the lexical phonology of larger units that arguably are lexical units), but post-lexical phonology would take scope over the whole sentence and thus could include processes operating on smaller units of the word.

and post-lexical distinction, as explained above. We can still maintain a division between word and sentence phonology which could be relevant at either level although I would be inclined to emphasize it more at the phonotactic level.[24]

These two phonologies should, as said, be understood as different *levels* (and, as argued, not as layers in one hierarchy, nor as planes within one level), one being close to the traditional notion of the underlying or phonemic level, the other being close to the surface of phonetic level. However, rather than being derivationally related as in standard Generative Phonology, I adopt the view that each level is characterized by its own set of constraints (and repairs).[25] The relationship between the two levels is such that the output of the phonotactic level (i.e. after repair) can be taken as the input for the utterance level. The phonotactic level in turn can be taken to be the output of the morphotactic level, if we take this level to be responsible for combining the phonological material of morphemes and words, and not just of the categorical labels; I will return to this point below.

This notion of two phonologies is different in one further respect from McMahon's division. I do not take either phonotactic or utterance phonology to *include* intonation. Below, I will adopt the point made in Gussenhoven (1984a) saying that there are two *grammars*, one for the "text" and one for the "tune" (i.e. intonation). The latter like the former has several levels of analysis, perhaps the same as distinguished for the text grammar. This point will be taken up in section 5.

2.3. Extending the scope of "Utterance"

The idea of structural analogy demands that we ask the following question. If we have two phonological levels, differing in the degree of concreteness (since, indeed, the utterance level is closer to actual realizations of linguistic expressions), do we also find two semantic levels then? I believe that the answer to that question is yes. So let me render the diagram in (9) more complete by including an extra semantic level:

---

[24] The distinction made here is also made in Goldsmith (1993), who speaks of the phonological and the phonetic level respectively, and in Lakoff (1993).

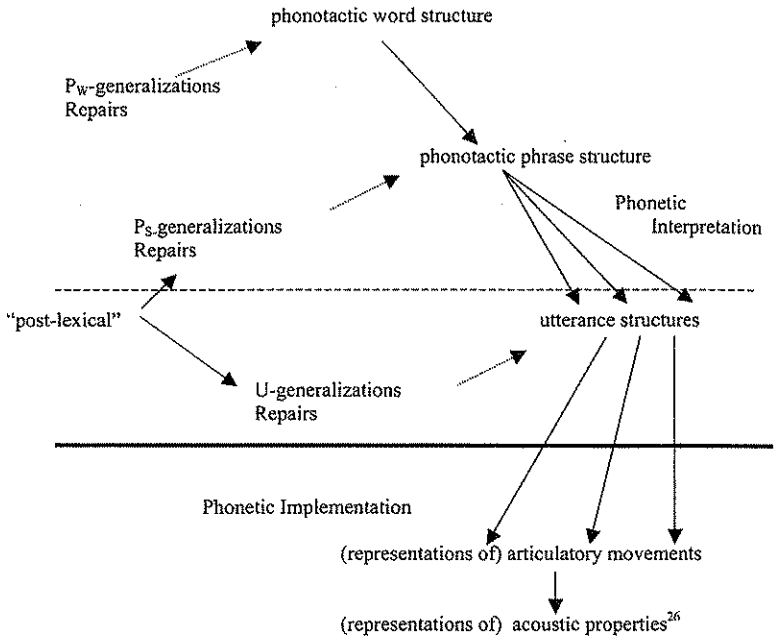
[25] Again, this is like the proposals in Goldsmith (1993) and Lakoff (1993). A difference is that I would like to reject rules that perform changes going from one level to the next; see van der Hulst (2008c) for further discussion and illustrations.

(10)

Utterance (Phonological) Structure
Phonotactic Structure
Morphotactic Structure
Conceptual Structure
Utterance (Semantic) Structure

Just like utterance (phonological) structure emerges when a linguistic expression is put to, or rather prepared for *actual use* in a specific context, the conceptual (semantic) side of expressions can also be distinguished from what is sometimes called the utterance meaning (Löbner 2002). This utterance meaning is richer and less schematic than the conceptual structure, being bound to a specific context of use. The step away from conceptual meaning (called [descriptive] expression meaning in Löbner 2002), which is often called semantic interpretation, links conceptual structure to actual situations (in some world). In a broader sense, we could call the utterance meaning *pragmatic* meaning. Likewise, we could refer to the step that links phonotactic structure to utterance phonetic structure as phonetic interpretation, although it is more common to use the phrase *phonetic implementation*. In fact, Liberman and Pierrehumbert (1984) argue that what is called post-lexical phonology (roughly my utterance phonology) is in fact nothing less than (the result of) phonetic implementation, suggesting that all properties and processes that pertain to this level are gradient and non-continuous rather than categorial. I am hesitant, however, about equating the utterance structure with actual phonetic (i.e. articulatory or acoustic) events. Rather, as argued in van der Hulst (2008c), we should regard utterance structures as still being a *cognitive* level. I suggest, therefore, to attach different meaning to the notions interpretation and implementation. The diagram below is, with same adaptations, taken from van der Hulst (2008b):

(11)



It could likewise be argued that semantic conceptual and semantic utterance (pragmatic) level are both internal, cognitive levels. The pragmatic level, after all, should not be equated with the actual mind-external referents, just like the utterance level is not identical to the actual motor activities or resulting sounds. An exploration of that analogy lies clearly beyond the scope of this article. In any event, interpretation is the linking of phonotactic and semantic structure to contexts of use, whereas implementation is used for the linking of utterances structures to real world objects.[27]

---

[26] On the phonological side of grammar we face the question of how to account for the duality of production and perception. The question is whether U-structure compromises between production and perception or is determined primarily by only one of the two, and if so which one is secondary. We find defenders of both possible views. In the so-called "Motor Theory of Speech Perception", articulation is primary, whereas the idea that form is psycho-acoustically- based is found in "Government and Dependency Phonology".

[27] This makes more sense on the P-side since the real world objects (physical articulations and sounds) are literally caused by the phonological utterance structures. This cannot be said on the C-side, where the referents of semantic structures exist independently.

Note, that the model in (11) is meant to express that for each phonotactic representation there is a multitude (or infinite number) of utterance level representations, each suited for specific conditions of use.

2.4. Dependencies

The models in (9) or (10) with their parallel levels look very Jackendovian (cf. Jackendoff 2002). This model seems uncontroversial in its essential claim that the semiotic coupling of conceptual meaning and phonetic form is mediated by a grammatical system. The controversial aspect of Jackendoffs' model regards the independence that he attributes to each of the grammatical modules or levels. With this view, Jackendoff distances himself from what he calls *syntactocentrism* according to which the structures in the phonotactic and conceptual domain are largely dictated by morphotactics.

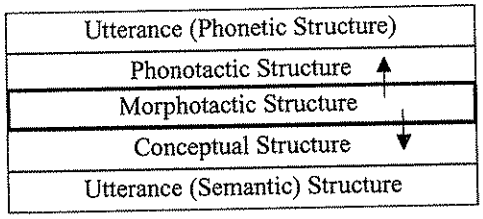
Jackendoff (2002) argues that each of the three types of structures are autonomous, i.e. subject to their own set of constraints. I will not dispute this. He then concludes from this that the grammar should be seen not as driven by a central morphotactic engine, but as a system of three independent structure-producing engines (or structure-admitting constraint sets) *plus* a system of linking or correspondence rules which characterize wellformed combinations of wellformed P-, M-, and C-structures. Marantz (2005) points out, correctly I believe, that this conception of grammar can only derive the high degree of structural isomorphism between P-, M- and C-structure (which Jackendoff does not dispute) in a stipulative fashion. This isomorphism has to be stipulated in the system of correspondence rules. Note that isomorphism cannot be explained as a form of structural analogy. Structural analogy leads us to expect that P-, M- and C-structures use similar (identical) structural principles, not that they are structural isomorphic. Departing from Jackendoff's idea that the grammar characterizes three autonomous structures and conditions on uniting these into linguistic expressions, it strikes me as more realistic to say that the actual linguistic expressions are formed by combining morphemes, words and phrases which, at the lowest layer (morphemes), are combinations of phonological form and conceptual structure.[28]

In van der Hulst (2008a), I have argued against the strict parallel view, ending up with the more traditional view in which we see morphotactics as combing

[28] Jackendoff would probably say that I confuse the notion *competence* and *performance* but it is not at all clear how that distinction can be made and, if it is made, what is accounted for where.

morpheme and word labels, dragging along the phonological and semantic properties of these units:[29]

(12)



In this conception, M-structure truly mediates between P- and C-structure, as such imposing a constituent structure on chunks of phonological and conceptual material (i.e. morphemes), which is indicated by the arrows in figure (12). Since both form and meaning have their own wellformedness requirements, one might expect that M-structure aims to achieve a "compromise" between the truly autonomous P- and C-structures, thus minimizing the amount of adjustment or repair that is necessary at these levels. This makes M-structure central in a way, but hardly in the sense that the other modules must bow to its superiority. Rather, as just stated, M-structure is a hybrid system that serves the autonomous and distinct demands of P- and C-structure, and in this sense, morphotactics is the one that bows. However, being a separate level, we expect M-structure to have its own demands on wellformedness as well, which, presumably, limit its serving capacity.

M-structure, existing by virtue of its central, mediating role, then decides on the way that linguistic expressions are built out of morphemes and words. If the combinatorial system would only pay attention to the *phonotactic* properties of morphemes and words, it is likely that, unintentionally, a structure would arise that lumps together chunks of meaning that would not reflect a compositional, i.e. transparent semantic organization. Accordingly, the resulting structure would require a lot of semantic computation at the C-level. If, on the other hand, the combinatorial system would only pay attention to the semantic properties of morphemes and words, an organization would arise that presumably would not be acceptable from the point of view of phonotactic form. Thus, again, a categorial system has emerged (M-structure) that, I would think, mostly serves semantic compositionality but departs from it where the demands from phonotactics are just too strong. If this is indeed true, it would have to be the case that all M-structure that compromises semantic compositionality is, in origin at

[29] By assuming this view, I part company with "Distributed Morphology" views or any other view which regards the categorial engine (morphotactics) as bearing on the category labels only, form and meaning being attached later on.

least, either phonotactically driven, or rooted in autonomous constraints on M-structure. (A third reason for structure paradoxes of this sort presumably lies in M-structure trying to serve yet another master, namely "information structure" which causes it to split up semantic units. To show that all this is indeed the case would require more than a few paragraphs.)

2.5. Is word order (linearization) a phonotactic matter?

We have decided to regard phonotactics and morphotactics as different levels. Together, these levels are often regarded as pertaining to the form of language and it that sense, they jointly differ from semantics. This unity of morphotactics and phonotactics is not indicated in figure (12) and I will make no attempt to do so. From the model in (12) we cannot derive precisely what the division of labor is between phonotactics and morphotactics, and that this question arises, again, indicates the functional affinity between them. According to one view, certain aspects of what would often be considered morphotactic matters can also be placed within the realm of phonotactics. This, specifically, regards the issue of *linearization*. Several syntactic frameworks make a formal distinction between constituency-as-grouping and linear order within constituents. With this distinction in place, it could be argued that syntax only regards grouping, while linearization is located in the phonology. Additionally, within generative grammar, other, formerly held to be syntactic phenomena, have been assigned to the phonology (to PF) such as various kinds of movement of syntactic constituents. It seems to me that if the phonological component is burdened with accounting for all the aspects of what used to be syntax that differ from one language to the next, i.e. that are not universal, what remains to the syntax proper ("narrow syntax") is little more than an account of relationships that could just be called semantic as well. Indeed, this seems to be implied in a view of minimalist syntax as an invariant, universal system, presented in Burton-Roberts (2000), who regards phonology as everything that functions to externalize the universal syntactic representations, thus including into phonology what would constitute a lot, if not most of what in other models might be called syntax. It is interesting to note (as pointed out in van der Hulst 2002) that what Burton-Roberts (2000) thus takes to be the core of grammar, i.e. this narrow universal syntax, is hard to distinguish from the grammar of thought, which is taken by Jackendoff (2002) to be *outside* grammar in the strict sense. Jackendoff takes the grammar of conceptual structure to exist in its own right, with the linguistic grammar characterizing the linguistic expressions (i.e. M- and P-structure) that are used to linguistically express a subset of the conceptual structures. Of course, the difference between these two views is purely terminological and both linguists

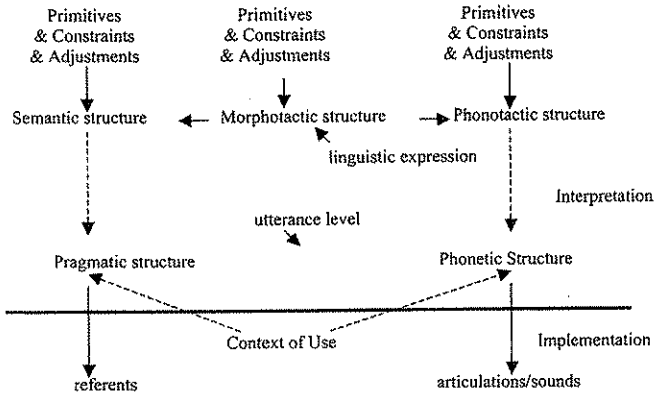
agree that phonotactics-cum-morphotactics is a system that humans use to externalize (some of their) thoughts (which does not deny that such externalizations can be kept internal in the form of silent speech).

Terminology and quibbles about what is and what is not grammar aside, I will maintain the picture in (12) which implies that the system that allows us to externalize thoughts is divided into a phonotactic system and a morphotactic system, assuming that the latter system is responsible for the *surface linear order* of words and phrases; thus, I take syntax to be broad. It is after all unlikely that all aspects of linearization (morpheme and word order) can be attributed to the kinds of factors that drive phonotactic structure. The necessity to produce morphemes and words linearly leaves room for different solutions which do not seem to depend on phonotactic principles or parameters. As mentioned, specific placement of phrases can reflect information structure (which I take to be part of or being close to conceptual structure, perhaps a distinct semantic plane) or, in an iconic fashion, the course of events. Phonotactics or information structure, while perhaps being potential original motivators for specific M-structures, cannot be taken to do the work for M-structure, which is in principle independent from both and tries as best as it can to compromise between the demands of form and meaning. By spelling out linear organization, M-structure is serving the demands of P-structure and by spelling it out in specific ways it serves the demands of information structure.

2.6. The interaction between grammatical components

The following diagram taken from van der Hulst (2008c), with small changes, summarizes the interaction of the various grammatical components and the relationship between the different levels, and between those levels and "the external world":

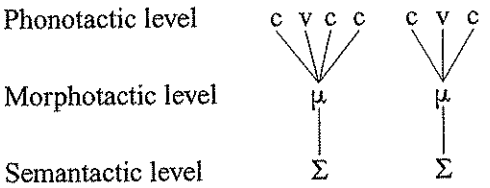
(13)



Here the tactic levels and semantics jointly characterize what I call linguistic expressions, utterances being a different level at which the expressions are enriched with contextual properties (i.e. pragmatic and phonetic detail). Everything below the bold line is "the world", above it "the mind".

Phonotactic structures are arbitrarily linked to conceptual structures, let us say *via* or *under* a category label. The labels, in a sense, function as a skeleton that coordinates bits of structure in the phonotactic and semantic plane:

(14)



Contrary to traditional views which regard syllable structure as derivable from the linear sequence of phonemes, I adopt the government phonology claim that morphemes are in fact phonologically structured expressions, consisting of onsets and rhyme constituents, augmented with *lateral licensing relationships* (cf. van der Hulst 2006a). Morphemes that would appear to be phonotactically incomplete contain empty structural material. For example, a consonantal affix would be a syllable with an empty rhyme. In this view, repairs or adjustments that need to be made when morphemes are combined are minimal but nonetheless necessary (cf. van der Hulst 2008c). The symbol Σ (capital sigma) stands for a conceptual structure. When a morpheme enters into a combination with other morphemes, as directed by the morphotactic system, *each* of its three levels enters into combinations that are subject to wellformedness conditions and, if necessary, repairs. In this view then, the same structural organization is initially present on all three planes. This organization is licensed by the morphotactic constraints. Wherever this structure does not meet the wellformedness constraints at the P- or C-level, structural adjustments (restructurings) are required *at these levels*. M-structure that does not reflect semantic compositionality, when projected at the C-level, requires semantic rules that either create compositional structure or compute semantic structure non-compositionally (appealing to function application and other mechanisms). At the P-level, structural adjustments are made wherever word and phrasal groupings result that violate the phonotactic wellformedness structural constraints; this leads to incorporation of weak elements. In addition, segmental repairs are made when, within phonotactic domains, constraints on the segmental content are violated. This is what we mean when we say that P- and C-structure have to be computed from the morphotactic structure. These steps are often called phonological and semantic interpretation but

it would be more proper to call them phonological and semantic *reconstruction*, in particular because I have reserved the term interpretation for something else.

2.7. Questions of evolutionary precedence

Let us briefly ask what language is for, or for what reason it developed in our species. Was the reason to *facilitate* thought or to *externalize* thought? Note that having evolved for thought does not exclude *axaptation* for communicative purposes. The following example serves as a comparison: Most people have forgotten why the internet and WWW were invented in the first place, assuming there was one reason to begin with. Now it is used for many reasons, never foreseen by the original inventors, and most likely more crucial than whatever the original reasons were. In the (near) future, new uses will undoubtedly emerge. Answering questions regarding the function of language depends on what we call grammar. In my opinion, it is highly likely that what Jackendoff (2002) calls grammar (phonotactics and morphotactics) emerged as a system for externalizing thoughts. However, this is fully compatible with the claim that prior to developing a system for externalization, a grammar of thought had to develop (which for Burton-Roberts [2000] is grammar), i.e. a conceptual system which no longer was dependent on sensory stimuli or episodic memories and which allowed humans to structure the same proposition in a variety of ways. Once a conceptual system of this sort had emerged and thoughts that were not linked to the here and now could be juggled in the minds of our ancestors, a system of externalization could develop that would be different from the call systems that we find in many animal species. Thus, syntax in Burton-Roberts's (2000) sense had to develop first. It is plausible that once the system of externalization based on words and combinations of words was in place, the possibility of exchanging thoughts made the grammars of thought increasingly more sophisticated, which in turn put higher demands on the system of externalization.

The emergence of this system of externalization could have "piggy-backed" on an already available lexicon of calls, i.e. the notion of linking chunks of information to perceptible, visual or auditory signals. From this point on, compositionality would have developed in two directions, downwards into the utterance form of elementary expressions, leading to phonotactics (the second articulation), and upwards, leading to morphotactics (the first articulation). Perhaps, Martinet (1960) was of the opinion that the development of morphotactics came first. Others, such as Carstairs-McCarthy (1999), believe that phonotactics came first. Novak and Krakauer (e.g. 1999) claim that the development of both can be understood as resulting from a pressure caused by the need for increasingly complex communicative descriptions.

It is often noted that a critical difference between P-structure and M-structure is that the latter is recursive while the former is not. It seems obvious to me that the recursivity of M-structure is one of the ways in which M-structure pays respect to C-structure. Conceptual structure must have developed recursion first, as part of representational apparatus that was needed to construe representation of the outside world and of other minds. There is recursion in the world in the sense that things can be inside things which in turn contain other things that may be of the same type as the things they are contained in, and so on.[30] Likewise, as has often been pointed out, social relationships (combined with a theory of mind) invite, if not require, recursive thinking. The recursion of M-structure was simply *serving* the recursive organization of C-structure. P-structure is intrinsically unlike C-structure; its structure is linear and repetitive at each layer. Recently, it has been questioned whether P-structure, given its apparently layered organization needs to be constituency-based (see Neeleman and van der Koot 2006). I will return to this point below where I will discuss the properties of P-structure and (phonological) utterance structure.

3. What is the division of labor between the two phonologies?

First, I have argued that we cannot place the phonotactic and morphotactic structure in sequence, as layers within one level. Second, we discussed the way in which morphotactics mediates between phonotactic and semantic structure, serving their interests the best it can. (Whether it does so perfectly is not clear.) Then I argued that there is not one phonological level but rather that two levels can or need to be distinguished – the phonotactic level and the utterance level. Now we have to ask what the division of labor is between these two phonologies. The upshot of my discussion will be that the information that is encoded in the so-called *prosodic hierarchy*, including the metrical grid, and the processes that motivate these structures, must be parceled out over the P-level and the U-level. It seems to me that the literature on prosodic constituency has conflated phenomena and arguments for two different hierarchies. How this parceling out is done in detail remains to be worked out. Here, I will only be able to make some programmatic suggestions.

The distinction between lexical and post-lexical phonology has often been seen in the light of the obligatoriness or optionality of the processes involved.[31]

---

[30] Of course, with real things, we run into size issues because the containees need to be smaller than the containers.

[31] See Kiparsky (1995), Kaisse and Shaw (1995).

Post-lexical phenomena would be seen as subject to speech rate and speech style and *could* be gradient rather than categorical. I suggest that precisely these style-dependent phenomena are to be located at the U-level. Thus, whereas phonotactic structure is fixed, invariant and discrete-categorical, the utterance structure is variable and potentially gradient, dependent as it is on the intended context of use. Indeed, as stated earlier, for any given phonotactic structure there is an infinite number of corresponding utterance structures. When we consider the structures and processes that are surveyed in Nespor and Vogel (1986), it seems clear that in many cases the formation of domains, and application of processes that motivate them, are sensitive to speech rate and speech style, suggesting that these phenomena are within the post-lexical and thus utterance sphere. These are the processes that Kaisse (1985) refers to as "fast speech rules".

However, this is not necessarily the case for *all* phenomena that have been said to be sensitive to prosodic structure. Kaisse (1985) analyzes various phenomena that Nespor and Vogel place in their prosodic phonology as direct syntax rules, i.e. rules (or constraints plus repairs) that operate with reference to syntactic domains.[32] In more recent work by Truckenbrodt (1999, 2007), building on others, it becomes very clear that one specific prosodic constituent, the *phonological phrase*, plays a pivotal role in many apparently categorial and obligatory phonological phenomena, and I suggest that these are more properly located at the P-level, assuming that at this level the domains are largely determined by the morphotactic structure, as argued earlier. Thus, even though the phonological phrases are not *strictly* isomorphic with syntactic phrases in all cases, they are so close that Truckenbrodt (1999) can state his "Wrap-XP constraint" which says that each syntactic phrase is contained in a phonological phrase. Phonological phrases, in other words, do not cut across syntactic phrase boundaries.

The relationship between syntactic phrases and phonological phrases is rather similar to that between morphological words and phonological words. The generalization that Truckenbrodt observes at the phrasal level also applies at the word level in the sense that we can say that each morphotactic word is contained in a phonological word. Each simplex word is a phonological word, but

[32] Candidates for "direct syntax" segmental phonology have been discussed and then critically evaluated in Hayes (1990) who argues that the alternations involved reflect what he calls "precompiled phonology", i.e. the lexical storage of alternate forms of words and selection frames which guide selection of the right word form in specific syntactic contexts. However, the option of precompilation (a mechanism that can also be applied to deal with allomorphy) does not take away that phonotactic constraints can operate in larger domains than words. Pullum and Zwicky (1988) discuss many cases of syntax-dependent phonology.

phonological words can also correlate with a subclass of the complex words. Following Kabak and Revithiadou (this volume), let us say that the phonotactic organization is primarily drapped around the morphotactic units *word* and *phrase*. Phonotactic words and phrases are either isomorphic to morphotactic words and phrases, or deviate from those in incorporating "weak words" (inflectional affixes, clitics, second halves of compounds, as proposed in Vogel this volume) or an adjacent small phrase, thus allowing a limited (perhaps only one-time) recursion.

Space limitations prevent me from discussing even a small sample of processes that motivate this high degree of isomorphism of P-structure to M-structure at both the word and phrase level. However, I will specifically draw attention to one class of phenomena that, in my view, must be accounted for at the phonotactic rather than the utterance level, namely *accent*.[33] Here too, we find analogies between the word and the phrasal layer. I suggest that the location of *primary* accents at both the word and phrasal level are handled at the P-level with reference to domains that are largely morphotactic plus or minus weak elements. In this way we expect primary accent at both levels to be sensitive to grammatical information of a variety of sorts. Primary word accent, for example, can be sensitive to word classes (as it is in English) and to details of the phonotactic organization, specifically with reference to the complexity of rhymes (phonotactic weight). In addition, accents can be encoded lexically in so-called lexical accent systems, or by way of exceptions to general rules. Phrasal accent can also be sensitive to grammatical information such as argument structure (as in various Germanic languages), if, indeed, this is considered morphotactic information. Perhaps, phrasal accent can be sensitive to phrasal categories (I do not know an example of that), or to lexical information, which would mean that certain words require to be accented (see Oehrle 1981 who discusses words that have specific accentual requirements). Of course, both word and phrasal accents can be merely sensitive to M-edges which we see in so-called weight-insensitive word accentual systems (Icelandic, Polish, etc.) and in the phrasal accentual systems of the Romance languages (see Ladd 1996) and Bengali (Hayes and Lahiri 1991).

With respect to word accent, I have argued at length (van der Hulst 1984, 1996, 1999, 2002, 2008a) that unlike the primary accent location, the overall rhythmic organization is post-lexical, which, in the current proposal, means that it belongs to the utterance level. I suggest that we explore the same idea for the overall rhythmic structure of entire utterances, meaning that all matters

[33] Another phenomenon that needs to be accounted for at the sentential P-level is inflectional allomorphy.

of rhythm (eurhythmic principles and rhythmic adjustments; see Hayes 1984; Prince 1983; Visch 1989, 1999) are dealt with at the utterance level in terms of a metrical grid structure. This idea then naturally feeds into a further idea, namely that prosodic structure at the utterance level is derived by or determined by this grid structure. I will furthermore suggest that at this level the prosodic structure has two principal domains, the (Abercrombian) foot and the Intonation phrase.[34]

4. Similar proposals

The broad aspects of the proposal that I am making here are not entirely original. I already referred to Goldsmith (1993) and Lakoff (1993) who propose different phonological levels (even a third, namely the morphemic level). However, my proposal regards specific ideas about the notion of hierarchical organization at both levels, and therefore I wish to point to other proposals that are very similar in this respect.

In his work on phrasal accentuation on Danish, Rischel (1982) refers to "an abstract prosodic structure which is directly coupled to syntax", which is what I here call the phonotactic structure. Rischel then also postulates a "surface phonology" which, from my perspective, equals U-structure. In other words, I suggest that Rischel's distinction between his "deep" and "surface" phonology is exactly the distinction between P-structure (which reflects morphotactic organization) and U-structure (which is independent from morphotactic information). Rischel also argues that phrasal accents are assigned with reference to the syntactic structure, which, in the view presented here, means that they are assigned at the phonotactic level, as I suggested earlier.[35]

Another author who has suggested the need for two prosodic hierarchies is Eric Fudge (manuscript) who states his case as follows:

These two hierarchies relate to different tasks: among other things, the first [hierarchy] assigns accents (potential stresses) to certain parts of the utter-

---

[34] Liberman and Pierrehumbert (1984) argue in the context of discussing the notion phonetic implementation that above the phonological phrase, no further constituency is needed. This squares with what I am proposing here since the phonological phrase is the largest constituent at the P-level. Other, larger domains are expressed in terms of their notion of phonetic implementation which subsumes my notion of utterance structure.

[35] The idea that phrasal accent is sensitive to syntactic domains is also defended in works by Cinque (1993) and Zubizarreta (1998).

ance, while the second determines which potential stresses are made into actual stresses when it comes to realising the utterance in speech.

He then compares this to a distinction made by Bolinger (1981) between "syllabic rhythm" (first hierarchy) and "accentual rhythm" (second hierarchy). Fudge places the formation of feet in the second hierarchy, making an explicit reference to the notion of foot discussed in Abercrombie (1964).

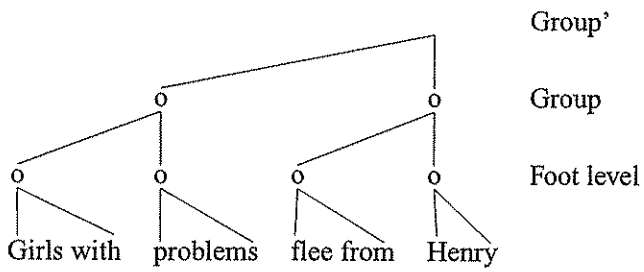
Anderson and Ewen (1987) also make a distinction between what they call "word structure" versus "utterance structure", the second structure matching the utterance level that I have proposed here, as well as Rischel's (1982) "surface phonology" and Fudge's second hierarchy. Both Fudge as well as Anderson and Ewen limit their first hierarchy to words, seemingly excluding here the assignment of phrasal accents. Rischel, however, explicitly includes phrasal accents in his first hierarchy, and I agree with him in this respect.

In a less specific sense, my proposal also agrees with other proposals. In retrospect, it could be said that Liberman and Prince (1977), as well as Hayes (1984) had it right when they suggested two metrical organizations, one closely following morphotactics (the metrical tree) and the other being independent from that structure merely carrying over the primary accents expressed in the grid. They were wrong in including rhythmic organization also in the arboreal structure and this is why the tree and the grid seemed to overlap to the extent that some tried to eliminate the former (Prince 1983), while others argued against the latter (Kiparsky 1979; Giegerich 1985).

Finally, Lahiri and Plank (2007) suggest that the kind of prosodic structure that has been proposed in the Selkirk/Nespor and Vogel tradition ignores a simpler structure that, as they point out, has a long tradition in the literature on sentential rhythm. In essence, like Rischel, Anderson and Ewen as well as Fudge, they argue in favor of the Abercrombian foot, which, as they point out was around long before Abercrombie drew attention to it. As is clear, my proposal is fully compatible with that suggestion if we agree that they are referring to the utterance level. Since, at this level and in terms of the Abercrombian foot, we cannot account for many of the phenomena that have been analyzed in the Selkirk/Nespor and Vogel tradition however, we *also* need the P-level organization which, while still being phonological, is only marginally different from surface syntactic structure.

Turning back to the linguists who have argued for a dual phonological structure, let us take a brief look at how the utterance structure is created. All authors cited propose to make "foot formation" the crucial step in the formation of the utterance structure. The following example is taken from Anderson and Ewen (1987: 103):

(15)



Anderson and Ewen adopt a group level in addition to their foot level. The Group seems to be directed by phrasal accents while feet are directed by word accents. Since these authors do not assume a P-level that comprises larger-than-words units, they express matters like phrasal accent *only* in their utterance structures. However, as pointed out, phrasal *primary* accent location is *not* dependent on speech rate or style. The phrasal primary accent location is invariant for any given type of syntactic phrase. It is noteworthy that a dependency style arboreal structure comprises the metrical grid in that each head (indicated by a vertical line ending in a 'o') is a grid mark (as pointed out in Ewen 1986) and can be seen as deriving from it (although Ewen, as did Liberman and Prince 1977, saw the relation in the opposite direction).

Fudge (manuscript) also forms feet in his utterance structure which stretch from stressed syllable to stressed syllable. He states, as do Anderson and Ewen (1987), that the relevant notion of foot was long ago recognized by the phonetician Abercrombie (e.g. 1967). Clearly, these Abercrombian feet are quite different from the metrical feet proposed by Liberman and Prince (1977) which encode rhythmic ripples in between word accents that only appear to be confined to word boundaries when words are treated as complete utterances.

Rischel's (1982) surface phonology feet also comprise a syllable with primary word stress followed by syllables with lower degree of stress. Interestingly, he explicitly states that this notion of foot is relevant as a domain in the analysis of intonation (referring to work by Nina Thorsen, e.g. Thorsen 1983). Whether this means that his feet can be regarded as "intonational groups" or subunits of such groups remains to be seen. Anderson and Ewen suggest that what they call their Group unit (which is superordinate to the feet) is relevant to the spreading out of intonational melodies. Perhaps then both the Foot and the Group are relevant domains for intonation, similar to the notions Minor Phrase and Major Phrase in works such as Beckman and Pierrehumbert (1986).

All these proposals seem to embody an approach to utterance structure that follows a simple principle:

(16) U-structure is constructed consistent with word and phrasal accents in P-structure

Earlier, I referred to Neeleman and van der Koot (2006) who argue that prosodic structure is a non-arboreal structure. We have now established that the structure at the U-level is much less rich than the original prosodic hierarchy, and it may thus be true that a proper formal expression of structure at the U-level does not have to be arboreal, as in (15), but could appeal to a more impoverished mechanism that involves a grid structure enriched with boundary markers (see e.g. Reiss 2008; Scheer, in press).

5. Intonation and focus

A complete utterance consists of a text and a tune that must somehow be lined up. In this section, I wish to endorse a view of intonation which indeed regards intonational melodies as independent expressions, generated by their own grammar which in principle has the same organization as the text grammar (see Pierrehumbert 1980, and in particular Gussenhoven 1984a). In this compositional view of intonation melodies, intonation is thus not just a matter of phonology or phonetics. Several types of intonational units have been distinguished: Pitch accents, phrasal accents and boundary tones (Pierrehumbert 1980; Ladd 1996; Gussenhoven 2004). Pitch accents are intonational words that have a phonological form (including an accent), and a meaning, as well as a category (namely *Pitch Accent*). Pitch accent may occur as such or modified by means of a morphological operation (cf. Gussenhoven 1984a). Phrasal accents, linked to smaller units, and boundary tones, linked to larger units, are simpler entities; they have no morphology but they do have a phonology, meaning and category. The phonotactic structure of an intonation melody is a sequence of high and low tones, grouped into units such as pitch accents (one of which is accented) and other tones and provided with an overall hierarchical structure.[36]

The question now arises how the *text* phonotactic structure and the *tune* phonotactic structure are fused. This is a very difficult subject. The simple part of the problem is that pitch accents (i.e. accented tones) line up with certain types of phrasal accents, while phrasal tones and boundary tones line up with edges of domains. In accordance with the authors sited in the previous section, it would seem reasonable to locate details of the alignment process at the utterance

[36] Since it is often remarked that boundary tones are optional elements, which, however, does not include tones that mark question intonation, it is perhaps also possible to relegate the appearance of these optional boundary tones to U-structure.

level since this level seems to supply the domains that have been claimed to be relevant to the spreading of intonational tones. However, we can locate the linking of pitch accents to phrasal primary accents already at the P-level where such accents are assigned.

At this point, we need to bring in the notion *focus*. It is well-known that pitch accents occur on phrasal heads that belong to domains that are placed in focus. But at which level does focus-marking take place? Even though focus seems to be an inherently semantic notion, most frameworks assume that focus-marking occurs at the morphotactic level. In that sense, focus-marking can be seen as one way in which the morphotactic structure serves semantics. A morphotactic constituent that has been marked as being in focus calls for an expression of this status which can come in three forms: A specific morphotactic position (up front in the sentence), a specific focus particle or a specific pitch accent (see de Swart and de Hoop 2002). Each of these three mechanisms can be seen as subroutines that are activated by a focus feature, and this then suggests that, when a pitch accent is the chosen mechanism, the head of the pitch accent i.e. its accented tone is directly linked (affixed) to the primary accent of the focus domain, leaving the precise alignment of the entire intonational expression to the utterance level.

Assuming that linguistic expressions can contain more than one focus domain, the tune grammar produces an intonation expression, containing one pitch accent and several boundary tones for each focus domain.

Even though focus-marking can occur at the syntactic level where it triggers exponents that are strictly morphotactic (fronting, particles), there are at least two reasons for getting the phonotactic domain involved in the generation of pitch accents. First, strictly phonotactic units such as syllables can form focus domains (*I didn't say Record, I said Accord*). Secondly, depending on the length of focus domains or specific syntactic configuration, a focus domain can be split up into two domains, thus calling for two pitch accents (cf. Gussenhoven 1984b; Baart 1987).

6. Conclusions

In this article, I have defended the idea that we need two phonological levels. Even though this view is not all that new and echoes various other dichotomies including that between phonology and phonetics, it seemed to me that this point is worth (re)stating in view of the fact that there has been a tendency to concentrate accounts of phenomena that regard domains that are larger than words within one specific level (with specific reference to the "prosodic structure"

at this level) which many linguists take to be *the* phonological level. Another reason for trying to flesh out the nature of the two phonologies is that there are various other dichotomies around in the literature that, in one way or another, are versions of the proposal made here (without necessarily agreeing with it in all details). This concerns the *lexical – postlexical* dichotomy, the *metrical tree – metrical grid* dichotomy as well as the proposals discussed in section 5 of which Rischel's (1982) ideas most closely resemble the one I have intended to elaborate here. The distinction between two phonologies also makes sense of the apparent discrepancies between the constituents of Nespor and Vogel's (1986) prosodic hierarchy and the more traditional prosodic divisions that are discussed in the works discussed in section 5, particularly by Lahiri and Plank (2007). By parceling out the various prosodic constituents over two levels, we explain the dependency of some of them on morphosyntactic structure (as pointed out in Revithiadou and Kabak, this volume, as well as Truckenbrodt 1999) and the total independence of others (such as the Abercrombian foot and the Intonation phrase). In short, I believe that the idea of two phonologies with the division of labor suggested here is worth exploring further.

References

Abercrombie, David
1967 *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Anderson, John
1992 *Linguistic Representation: Structural Analogy and Stratification*. Berlin: Mouton de Gruyter.
2006 Structural analogy and universal grammar. *Lingua* 116: 601–633.

Anderson, John and Colin Ewen
1987 *Principles of Dependency Phonology*. Cambridge: Cambridge University Press

Anderson, John and Charles Jones
1974 Three theses concerning phonological representations. *Journal of Linguistics* 10: 1–26.

Anderson, Stephen
1985 *Phonology in the Twentieth Century. Theories of Rules and Theories of Representations*. Chicago: The University of Chicago Press.

Baart, Joan
1987 *Focus, syntax and accent placement*. Dissertation, Leiden University.

Beckman, Mary and Janet Pierrehumbert
1986 Intonational structure in English and Japanese. *Phonology Yearbook* 3: 255–309.

Blevins, Juliette
2003 The independent nature of phonotactic constraints: An alternative to syllable-based approaches. In: Caroline Féry, and Ruben van de Vijver (eds.), *The Syllable in Optimality Theory,* 375–403. Cambridge: Cambridge University Press.

Bolinger, Dwight
1981 Two kinds of vowels, two kinds of rhythm. Included as appendix A in Dwight Bolinger, *Intonation and its Parts: The Melody of Language*. Stanford, CA: Stanford University Press.

Booij, Geert
1999 Morpheme structure constraints and the phonotactics of Dutch. In: Harry van der Hulst, and Nancy A. Ritter (eds.), *The Syllable: Views and Facts*, 53–68. Berlin: Mouton de Gruyter.

Borowsky, Toni
1993 On the word level. In: Ellen Kaisse and Sharon Hargus (eds.), *Studies in Lexical Phonology,* 199–234. [Phonetics and Phonology. Volume 4]. New York: Academic Press.

Browman, Catherine and Louis Goldstein
1986 Towards an articulatory phonology. *Phonology Yearbook* 3: 219–252

Burton-Roberts, Noel.
2000 Where and what is phonology? A representational view. In: Noel Burton-Roberts, Philip Carr, and Gerard Docherty (eds.), *Phonological Knowledge: Its Nature and Status,* 39–66. Oxford: Oxford University Press.

Calabrese, Andrea
1995 A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry* 26: 373–463.
2005 *Markedness and Economy in a Derivational Model of Phonology*. Berlin/New York: Mouton de Gruyter.

Carstairs-McCarthy, Andrew
1999 *The Origins of Complex Language. An Inquiry into the Evolutionary Beginnings of Sentences, Syllables and Truth*. Oxford: Oxford University Press.

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

Cinque, Guglielmo
1993 A null theory of phrase and compound stress. *Linguistic Inquiry* 24: 239–297.

de Swart, Henriëtte and Helen de Hoop
2002 Topic and focus. In: Lisa Cheng, and Rint Sybesma (eds.), *The Second Glot International State-of-the-Article Book,* 79-106. Berlin: Mouton. [Also in Glot International 1.7, 3–7.]

Ewen, Colin
1986 Segmental and suprasegmental structure. In: Jacques Durand (ed.), *Dependency and Non-Linear Phonology,* 203–222. London: Croom Helm.
Fischer-Jørgensen, Eli
1975 *Trends in Phonological Theory*. Copenhagen: Academisk Forlag.
Fudge, Eric
Words and feet. Manuscript. University of Reading, England.
Giegerich, Heinz J.
1985 *Metrical Phonology and Phonological Structure*. Cambridge University Press, Cambridge.
Goldsmith, John
1993 Harmonic phonology. In: John Goldsmith (ed.), *The Last Phonological Rule,* 21–60. Chicago and London: The University of Chicago press.
Gussenhoven, Carlos
1984a Intonation: A whole autosegmental language. In: Harry van der Hulst, and Norval Smith (eds.), *Advances in Nonlinear Phonology,* 117–132. Dordrecht: Foris Publications.
1984b *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Foris.
1993 The Dutch foot and the chanted call. *Journal of Linguistics* 29: 37–63.
2004 *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
Hayes, Bruce
1984 The phonology of rhythm in English. *Linguistic Inquiry* 15: 33–74.
1990 Precompiled phrasal phonology. In: Sharon Inkelas, and Draga Zec (eds.), *The Phonology-Syntax Connection,* 85–108. Chicago and London: The university of Chicago Press.
Hayes, Bruce and Aditi Lahiri
1991 Bengali intonational phonology. *Natural Language and Linguistic Theory* 9: 47–96.
Hockett, Charles
1960 The origin of speech. *Scientific American* 203: 89–96.
Hulst, Harry van der
1978 Recent developments in phonological theory. *Lingua* 49: 207–38.
1984 *Syllable Structure and Stress in Dutch*. Dordrecht: Foris.
1989 Atoms of segmental structure: Components, gestures and dependency. *Phonology* 6: 253–284.
1996 Separating primary accent and secondary accent. In: Rob Goedemans, Harry van der Hulst, and Ellis Visch (eds.), *Stress Patterns of the World,* 1–26. HIL Publications 2. The Hague: Holland Academic Graphics.
1999 Word accent. In: Harry van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe,* 3–116. Berlin and New York: Mouton de Gruyter.

2000 Modularity and modality in phonology. In:. Noel Burton-Roberts, Philip Carr, and Gerard Docherty (eds.), *Phonological Knowledge: Its Nature and Status,* 207–244. Oxford: Oxford University Press.

2002 Stress. In: Lynn Nadel (ed.), *Encyclopedia of Cognitive Science*. Vol. 4, 246–254. London: Nature Publishing Group.

2003 Structure paradoxes in phonology. In: Stefan Ploch (ed.), *Living on the Edge. A Festschrift for Jonathan Kaye,* 75–94. Berlin: Mouton de Gruyter.

2005 Molecular structure of phonological segments. In: Philip Carr, Jacques Durand, and Colin Ewen (eds.), *Headhood, Elements, and Contrastivity*, 193–234. Amsterdam: John Benjamins Publishing Company.

2006a Licensing constraints in phonology. In: Harry van der Hulst, and Nancy A. Ritter (2008) (ed.), *Developments in Government Phonology*. Special issue of *The Linguistic Review* 23 (4): 383–428.

2006b On the parallel organization of linguistic components. *Lingua* 116: 657–688.

2008a Brackets and grid marks or theories of primary accent and rhythm. In: Eric Raimy, and Charles Cairns (eds.), *Contemporary Views on Architecture and Representations in Phonological Theory,* 225–246. MIT press.

2008b *Linguistic Structures*. Kendall/Hunt.

2008c Constraint-based phonologies. To appear in *UCWLP*.

Jackendoff, Ray

2002 *Foundations of Language*. Oxford: Oxford University Press.

Kabak, Bariş and Anthi Revithiadou

this volume An interface approach to prosodic word recursion.

Kager, René

1989 *A Metrical Theory of Stress and Destressing in English and Dutch*. Dordrecht: Foris Publications.

Kaisse, Ellen

1985 *Connected Speech. The Interaction of Syntax and Phonology*. Orlando: Academic Press, Inc.

Kaisse, Ellen and Patricia Shaw

1995 On the theory of lexical phonology. *Phonology* 2: 1–30.

Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud

1985 The internal structure of phonological elements: A theory of charm and government. *Phonology Yearbook* 2: 305–328.

1990 Constituent structure and government in phonology. *Phonology* 7: 193–232.

Kenstowicz, Michael, and Charles Kisseberth

1977 *Topics in Phonological Theory*. New York: Academic Press.

Kiparsky, Paul
1979 Metrical structure assignment is cyclic. *Linguistic Inquiry* 10: 421–441.
1985 Some consequences of lexical phonology. *Phonology Yearbook* 2: 89–136.

Ladd, D. Robert
1996 *Intonational phonology*. Cambridge: Cambridge University Press.

Lakoff, George
1993 Cognitive phonology. In: John Goldsmith (ed.), *The Last Phonological Rule*, 117–145. Chicago and London: The University of Chicago Press.

Lahiri, Aditi and Frans Plank
2007 On phonological phrasing in relation to morphosyntactic grouping. Ms., University of Konstanz.

Liberman, Mark and Alan Prince
1977 On stress and linguistic rhythm. *Linguistic Inquiry* 8: 249–336.

Liberman, Mark and Janet Pierrehumbert
1984 Intonational invariance under changes in pitch range and length. In: Mark Aronoff, and Richard Oehrle (eds.), *Language and Sound Structure*, 157–233. Cambridge, Mass.: MIT Press.

Löbner, Sebastian
2002 *Understanding Semantics*. London: Arnold.

Marantz, Alec
2005 Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review* 22(2): 429–446.

Martinet, André
1960 *Elements of General Linguistics*. London: Faber and Faber Ltd.

McMahon, April
2007 Sounds, brain, and evolution: or, why phonology is plural. In: Martha C. Pennington (ed.). *Phonology in Context*, 159–185. Palgrave, MacMillan.

Neeleman, Ad and Hans van der Koot
2006 On syntactic and phonological representations. *Lingua* 116 (10), 1524–1552.

Nespor, Marina and Irene Vogel
1986 *Prosodic Phonology*. Dordrecht: Foris Publications.

Novak, Martin and David Krakauer
1999 The evolution of language. *Proceedings of the National Academy of Science*. Volume 96: 8028–8033.

Oehrle, Richard
1981 Lexical justification In: Michael Moortgat, Harry van der Hulst, and Teun Hoekstra, (eds.), *The Scope of Lexical Rules*, 201–228. Dordrecht: Foris Publications.

Paradis, Carole
1988 On constraints and repair strategies. *The Linguistic Review* 6: 71–97.

Paradis, Carole and Jean-Francois Prunet
1993 On the validity of morpheme structure constraints. In: Carole Paradis. and Darlene LaCharité (eds.), *Constraint-Based Theories in Multilinear Phonology,* 235–256. [Special issue of *The Canadian Journal of Linguistics* 38].

Pierrehumbert, Janet
1980 The Phonetics and Phonology of English Intonation. Ph. D. diss., MIT.

Pierrehumbert, Janet and Mary Beckman
1988 *Japanese Tone Structure*. Cambridge, MA.: MIT Press.

Prince, Alan
1983 Relating to the grid. *Linguistic Inquiry* 14: 19–100.

Prince, Alan and Paul Smolensky
1993 *Optimality Theory and Constraint Interaction in Generative Grammar*. Technical Report #2 of the Rutgers Center for Cognitive Science. Rutgers University, New Jersey.

Pulgram, Ernst
1970 *Syllable, Word, Nexus, Cursus, Series Minor* 81. The Hague, Paris: Mouton.

Pullum, Geoffrey and Arnold Zwicky
1988 The syntax-phonology interface. In: Frederick J: Newmeyer (ed.), Volume 1: *Linguistic theory: Foundations,* 255–280. Cambridge: Cambridge University Press, 255–280.

Reiss, Charles
2008 'Long-distance dependencies and other formal issues in phonology'. In: Charles Cairns, and Eric Raimy (eds.), *Contemporary Views on Architecture and Representations in Phonological Theory*, 247–257. Cambridge, MA: MIT Press.

Rischel, Jorgen
1982 On unit accentuation in Danish – and the distinction between deep and surface phonology. *ARIPUC* 16: 191–239.

Rotenberg, Joel
1978 The syntax of phonology. Ph. D. diss., MIT.

Schane, Sanford
1984 The fundamentals of particle phonology. *Phonological Yearbook* 1: 129–55.

Scheer, Tobias
in press Why the prosodic hierarchy is a diacritic and why the interface must be direct. To appear in The Proceedings of the Tilburg Sounds of Silence Conference held in Tilburg in October 2005.

Selkirk, Elisabeth
1978 On prosodic structure and its relation to syntactic structure. In: Thorstein Fretheim (ed.), *Nordic Prosody 11*, 268–271. Trondheim: Tapir.
1980 Prosodic domains in phonology: Sanskrit revisited. In: Mark Aronoff, and Mary-Louise Kean (eds.), *Juncture,* 107–129. Saratoga: C.A. Anma Libri.
1984 *Phonology and Syntax. The Relation between Sound and Structure*. Cambridge, MA.: MIT Press.

Singh, Rajendra
1987 Well-formedness conditions and phonological theory. In: Wolfgang Dressler (ed.), *Phonologica 1984,* 273–286. Cambridge: Cambridge University Press.

Steedman, Mark
2000 Information structure and the phonology-syntax interface. *Linguistic Inquiry* 31 (4): 649–689.

Steriade, Donca
1999 Alternatives to syllabic analysis of consonantal phonotactics. In Osamu Fujimura, and Bohumil Palek (eds.), *Proceedings of the 2nd Linguistics and Phonetics Conference,* 205–245. Prague: Charles University Press.

Stewart, John M.
1983 Akan vowel harmony: The word structure conditions and the floating vowels. *Studies in African linguistics* 14 (2): 111–139.

Thorsen, Nina
1983 Standard Danish sentence intonation: Phonetic data and their representation. *Folia linguistica* 17: 187–220.

Trubetzkoy, Nikolaj S.
2001. *Studies in General Linguistics and Language Structure*. Anatoly Liberman (ed.); Translated by Marvin Taylor, and Anatoly Liberman. Durham and London: Duke University Press.

Truckenbrodt, Hubert
1999 On the relation between syntactic phrases and phonological phrases. *LI* 30: 219–255.
2007 The syntax-phonology interface. In: Paul de Lacy (ed.), *The Cambridge Handbook of Phonology,* 434–456. Cambridge: Cambridge University Press.

Visch, Ellis
1989 *The rhythm rule in English and Dutch*. Ph. D. diss., University of Utrecht.
1999 The rhythmic organization of compounds and phrases. In: Harry van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe,* 161–232. Berlin and New York: Mouton de Gruyter.

Vogel, Irene
this volume The status of the clitic group.
Zubizarreta, Maria Luisa
1998 *Prosody, Focus, and Word Order*. Cambridge: MIT Press.