

REVIEW ARTICLE

GEERT BOOIJ

Vakgroep Taalkunde, Vrije Universiteit Amsterdam

(Received 18 November 1987)

Marina Nespør & Irene Vogel, *Prosodic Phonology* (Studies in Generative Grammar 28). Dordrecht: Foris Publications, 1986. Pp. xiv + 327.

The theory of Prosodic Phonology is a theory about the representation of phonological structure. Its basic claim is that 'ultimate phonological constituents do not occur in an utterance as the individual bricks occur in a row of bricks. Rather, they occur in clusterings, these occur in still larger clusterings, and so on, up to the level of the whole utterance. That is, the phonological structure of the utterance shows a hierarchic organization...' (Hockett, 1955: 43).¹ In the book under review here, Nespør and Vogel develop a particular version of this theory, which is partly inspired by the work on prosodic phonology by Selkirk (1978, 1980a, b). Its basic claim is that we have to assume the following hierarchy of prosodic constituents:

- (1) syllable (σ),
foot (Σ),
phonological word (W),
clitic group (C),
phonological phrase (Ph),
intonational phrase (I),
utterance (U).

The hierarchical prosodic structure of a sentence is created by a set of mapping rules which derive it from the morphological structure. These principles are assumed to form part of post-lexical phonology. Thus, post-lexical phonological rules can refer to these prosodic categories. Hence, the organization of the phonological component is as follows:

- (2) syntactic surface structure
↓
mapping rules
↓
prosodic representations
↓
phonological rules
↓
phonetic representations

[1] This review was written as part of research programme Lett 83/7, Vrije Universiteit Amsterdam. See Booij (1983) for a short survey of the historical roots of prosodic phonology which include the work of Hockett, Haugen, Firth and Pike.

Nespor and Vogel argue (51) that there are four types of motivation for a particular prosodic category:

- (a) there are rules that need to refer to it in their formulation;
- (b) there are rules that have it as their domain of application;
- (c) it is a domain of phonotactic restrictions;
- (d) it is the bearer of prominence relations.

After two introductory chapters, they devote separate chapters to the motivation of each of the prosodic categories mentioned in (1) except that the syllable and the foot are treated together in one chapter. The book ends with two chapters on related issues, 'Prosodic constituents and disambiguation', in which the role of prosodic structure in speech perception is investigated, and 'Prosodic domains and the meter of the *Commedia*', in which the theory of Prosodic Phonology is applied in the metrical analysis of Dante's *Divina Commedia*.

The assumption of a hierarchy of prosodic categories is intimately related to the insight that often the domains in which phonological rules apply cannot be stated insightfully in terms of syntactic or morphological structure. That is, there is not always isomorphy between the morphological-syntactic structure and the structure required for the proper application of phonological rules. For example, the morphological structure of the English word *eater* (*eat-er*) is not isomorphic to the syllabification pattern of that word (*ea-ter*). The same applies at the level of phrases, as we will see below in the discussion of the individual prosodic categories.

In the preliminary chapters Nespor and Vogel propose the following general principles with respect to prosodic trees (7):

- (3) *Principle 1.* A given non-terminal unit of the prosodic hierarchy, X^P , is composed of one or more units of the immediately lower category, X^{P-1} .
- Principle 2.* A unit of a given level of the hierarchy is exhaustively contained in the superordinate unit of which it is a part.

These two principles are similar to what has been called the Strict Layer Hypothesis by Selkirk (1984). Below, I will point out certain problems with respect to this hypothesis.

Before dealing with the contents of the individual chapters, I would like to remark that this monograph on Prosodic Phonology is an important contribution to the further development and articulation of non-linear Phonology. Although certain details may be disputable, the necessity of an independent level of representation, prosodic structure, is convincingly argued for. The book also derives its quality from the wealth of data that it provides from a large number of languages, among which English, Italian, Dutch and Greek occupy a prominent place. Since it is impossible to

comment on all the specific analyses within the scope of a review article, I will focus below on the general issues raised by Prosodic Phonology.

Chapter 3, 'The syllable and the foot', provides the different types of motivation (listed above) for these two prosodic categories. Given the wealth of literature on the syllable and its rather uncontroversial nature, Nespor and Vogel rightly refrain from giving a very detailed treatment of this prosodic category. The only specific point they make is that the domain of syllabification is not the word in the grammatical sense, but the phonological word. For instance, in many languages compounds which are syntactically one word can be shown to consist of two or more phonological words. Thus, their syllabification patterns may seemingly violate the Maximal Onset Principle, as is illustrated by a compound like *pack ice* with the syllabification pattern (pack)_σ(ice)_σ.

As is well known, the foot is a much more debated category. The proponents of the grid-only theory (Prince, 1983; Selkirk, 1984) have proposed abolishing this prosodic category because its main function, the description of rhythmic alternations in strings of syllables, can be better accounted for by means of metrical grids. However, Nespor and Vogel argue that the foot also functions as a domain of phonological rules, and as a domain of phonotactic restrictions. For instance, the English rule of *t*-aspiration is formulated as follows: 'a *t* is aspirated if and only if it is the first segment of a foot' (91). Crucial evidence for not formulating this process in terms of stress (for instance, a *t* is aspirated at the beginning of a stressed syllable) are words like *terráin* and *typhóon* which consist of two feet and in which, according to Nespor and Vogel, the initial *t* is aspirated, even though the first syllable is not stressed. Evidence for the phonotactic role of feet is provided from Žu'hōasi, a Namibian tone language. For instance, 'although any single syllable may bear one of four basic tones, in bisyllabic words we actually find fewer than half of the possible combinations that would result from freely combining the different syllable types (i.e. 6 out of 16)' (101).

Chapter 4 deals with 'The phonological word'. It provides evidence for this prosodic category as a domain of phonological rules, and deals extensively with the mapping rules which derive phonological words in those classes where one grammatical word consists of more than one phonological word. These mapping rules vary from language to language. For instance, in Greek each compound corresponds to one phonological word, whereas in Dutch the constituent parts of compounds, certain (diacritically marked) suffixes like *-achtig* '-like', and prefixes form phonological words of their own.

A number of comments are in order here. First, note that Nespor and Vogel do not deal explicitly with the question whether there is also phonotactic motivation for this prosodic category. If there were no such motivation, Selkirk would be right in claiming that 'constituents like prosodic word or higher seem to have no role in phonology other than the

providing of domains' (Selkirk, 1986: 385). However, I think, unlike Selkirk, that the phonological (or prosodic) word has a phonotactic role to fulfil. First of all, we know that in many languages a well-formed word cannot simply be defined as a concatenation of one or more well-formed syllables, since words may have extra restrictions or extra combinatorial possibilities at their edges (cf. Booij, 1983, for a survey). For instance, in Dutch – and similar phenomena occur in other Germanic languages – a word can have an extra *s*, a *t* or a combination thereof after the last syllable, the so-called appendix. This appendix is a phonotactic possibility for the phonological word, not for the syntactic word, because in compounds such appendices also occur word-internally, as is illustrated by the Dutch compound [[herfst][aster]] 'Michaelmas daisy' with the syllabification pattern (herfst)_σ (as)_σ(ter)_σ. A second example can be taken from Yidip. Nespor and Vogel cite (135) Dixon's observation that a word in this language must contain at least two syllables. This should be interpreted as referring to the phonological word, because in that interpretation it follows that monosyllabic suffixes cannot form a phonological word of their own, whereas bisyllabic suffixes can (and do).

A second comment concerns the derivational mechanism for phonological words. On the one hand, we must first know how phonological strings are divided into phonological words before syllabification can take place, since it is assumed that the phonological word is the domain of syllabification. On the other hand, the syllables must already be there before the construction of phonological words, in accordance with the following principle of universal grammar proposed by Nespor and Vogel (7):

(4) *Prosodic Constituent Construction.*

Join into an N-ary branching X^P all X^{P-1} included in a string delimited by the definition of the domain of X^P .

This paradox only occurs with respect to the sequence of prosodic categories syllable, foot and phonological word. The construction of phonological phrases etc. is not problematic in this respect, because it only concerns the putting together of prosodic categories into higher categories within a certain, syntactically defined domain.

This paradox has to do with another problematic assumption made by Nespor and Vogel, namely that prosodic structure is completely derived in the postlexical phonological component. They do allow for phonological rules in the lexicon, but their claim is that lexical phonological rules do not refer to prosodic structure. In Booij (1987) I have shown that this hypothesis is incorrect. First, there are clearly cyclic, i.e. lexical, rules that have to refer to the syllable or the phonological word as their domain. Secondly, word formation rules may be sensitive to the stress properties of their base words. Since the representation of stress presupposes the existence of prosodic structure, the prosodic structure of words must be derived in the lexicon.

Thirdly, languages may need information about the prosodic structure of words for the correct application of morphological processes. For instance, in Yidjñ the first two syllables of a stem are reduplicated regardless of the make up of these syllables (Dixon, 1977: 156).²

My conclusion therefore is that prosodic structure up to the level of the phonological word has to be derived in the lexicon. To be fair, Nespor and Vogel themselves already point out that their position is not unproblematical because the mapping rules for the creation of phonological words must have access to the internal morphological structure of words (e.g. whether a word is a compound). Such information will not be available in the post-syntactic phonological component due to the Bracketing Erasure Convention which erases the internal morphological structure of words before they are inserted into syntactic structure, and thus accounts for the fact that the internal structure of complex words is opaque for syntactic rules.

Selkirk (1986: 385) also observes this discontinuity in the hierarchy of prosodic categories. For her, the prosodic word is a derived domain in sentence phonology, and not on a par with the syllable and the foot. Here, I disagree with Selkirk. For instance, stress properties of words must be derived in the lexicon, and we need the prosodic category label W for a proper representation of word stress in terms of strong/weak relations between constituents in a prosodic tree.

A nice consequence of the division of the derivation of prosodic categories across two components, the lexicon and the post-lexical component, is that another observation by Nespor and Vogel can also be accounted for, namely that in those languages where syllabification may take place across word boundaries, like e.g. French, it is still necessary first to syllabify the individual words. For instance (cf. 71; cf. also Booij, 1984), the phrase *première année* 'first year' is syllabified as follows:

(prə)_σ(mjɛ)_σ(ra)_σ(ne)_σ.

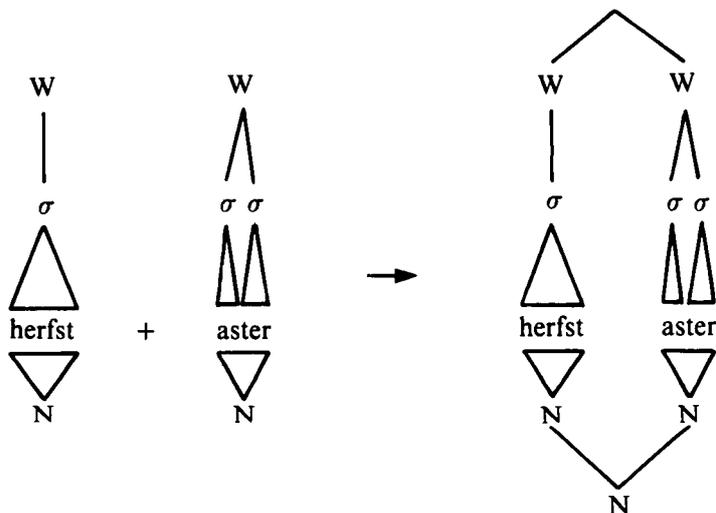
Yet, the second syllable must have been a closed one at some previous stage in the phonological derivation because the vowel [ɛ] occurs in closed syllables, in accordance with the rule of Closed Syllable Adjustment (CSA). If we assume that CSA is a lexical rule, the occurrence of [ɛ] is explained: the conditioning environment 'closed syllable' has been removed by resyllabification across word boundaries in sentence phonology, the so-called Enchaînement.

The solution to the problem of the derivational paradox that I would like to propose is that, for instance, compounds are not 'prosodified' after the formal operation of compounding, but rather that the compounding rule concatenates two words which have already been prosodified independently.

[2] In Booij (1987) I provide a number of examples of lexical phonological rules that refer to prosodic structure, and of word formation rules which are sensitive to the prosodic properties of their potential base words.

For instance, the compound *herfstaster* 'Michaelmas daisy' is derived as follows:

(5)



A similar procedure can be worked out for affixed words (cf. Booij, 1987). This procedure is possible, once we assign prosodic structure in the lexicon.

In (5) the prosodic top node of *herfstaster* does not bear a label. The Strict Layer Hypothesis implies, that – leaving the category C aside for the moment – the compound *herfstaster* forms a phonological phrase, just like NP's like *een rode aster* 'a red daisy'. Thus, Nespors and Vogel make the prediction that compounds behave like certain word groups with respect to phonological rules that apply in the domain of the phonological phrase. This consequence does not receive explicit attention in the book, and it would be important to see positive evidence for it. Given the fact that the rule of compound stress in Dutch is a lexical rule (it also has lexical exceptions), this would lead to the conclusion that even the category 'phonological phrase' would sometimes have to be derived in the lexicon.

Chapter 5, on 'The Clitic Group', makes the proposal that we should assume a prosodic category in between the phonological word and the phonological phrase, the Clitic Group. This is an innovation compared to Nespors & Vogel (1982), and the papers by Selkirk referred to above. The reason for assuming this category is that there are rules which apply (a) in word + clitic combinations but not in words, or (b) in words and word + clitic combinations, but not in phonological phrases. An example is the Latin stress rule which assigns penultimate stress to word + clitic combinations, e.g. *rósa* – *rosáque* 'rose/and the rose'. Nespors and Vogel argue that the penultimate stress pattern of *rosáque* cannot be accounted for by assuming

that *-que* has become part of the preceding phonological word: the Main Stress Rule of Latin only assigns penultimate stress to heavy syllables. As a consequence of the Strict Layer Hypothesis both *rosa* and *rosa^{que}* will be qualified as Clitic Groups, in accordance with the proposed mapping rule (154):

(6) *Clitic Group Formation.*

I. The domain of C consists of a W containing an independent (i.e. nonclitic) word plus any adjacent W's containing (a) a Directional Clitic, or (b) a Clitic such that there is no possible host with which it shares more category memberships.³

II. Join into an n-ary branching C all W's included in a string delimited by the definition of the domain of C.

In this way Nespor and Vogel try to account for the intermediate status of clitics, in between words and affixes. Technically, the fact that a word like *rosa* is qualified as a C is no problem since the special Latin stress rule for word + clitic combinations is defined as only applying to branching C's, and thus the clitic group *rosa* will not be affected. Conceptually, however, it seems rather awkward to accept the consequence that most clitic groups as defined by (6) do not contain a clitic at all.

Selkirk (1986: 387, 402, note 9) suggests the alternative of considering the category 'prosodic word' as a prosodic category of sentence phonology. This category is meant to be equivalent to the Clitic Group as proposed by Nespor and Vogel. In Selkirk's view, a prosodic word consists of a word of a major lexical category, or such a word plus an adjacent function word. Selkirk's presupposition is that the notion 'prosodic (or phonological) word' is not required in the lexicon, because it has no phonotactic role, and only plays a role as a domain of phonological rules. However, as we saw above, there are a number of reasons why we do need the prosodic word in the lexicon: phonotactic constraints, stress rules etc.

Nespor and Vogel's proposal presupposes that (a) all clitics are phonological words, and (b) all word + clitic combinations are derived post-lexically, in the phonological component. Both implications seem to me to be rather doubtful. As to (a), Dutch has clitics which only consist of a consonant, the first person singular pronoun /k/ and the third person singular pronoun /t/, and which thus cannot form phonological words of their own, because there are no vowels. Secondly, for certain languages (e.g. Dutch and Polish, cf. Booij & Rubach, 1987) the assumption of lexical clitics is well motivated. Note also that certain clitics induce obligatory resyllabification and fuse with an adjacent phonological word. For instance, the Dutch clitic pronoun *-ie* 'he' fuses obligatorily with a preceding verb or

[3] Nespor and Vogel do not define the notion category membership, which they borrow from a paper by Hayes.

complementizer, as in *komt-ie* 'lit. comes he' and *dat-ie* 'that he'. This obligatory resyllabification is not accounted for by calling *komt-ie* and *dat-ie* a clitic group, because in Nespor and Vogel's theory the domain of syllabification is the phonological word, not the clitic group.

A more general methodological question is whether the domain of application of one rule is sufficient evidence for the assumption of a specific prosodic category. For instance, the Latin stress rule for word+clitic combinations could also be formulated as a rule that shifts the stress of words before the relevant clitics to the final syllable of that word. A related question is whether all languages possess all prosodic categories.

In conclusion, I think that the whole issue of how to treat clitics in prosodic phonology should be investigated in much more detail, before we are able to evaluate properly Nespor and Vogel's postulation of a prosodic category Clitic Group.

The chapter on 'The Phonological Phrase' is perhaps the most central chapter of this book. It argues for the prosodic category Phonological Phrase (Ph), primarily by analysing three rules of Italian, Raddoppiamento Sintattico, Stress Retraction and Final Lengthening. Other rules discussed here are Iambic Reversal (English), Liaison (French) and Vowel Lengthening (ChiMwi:ni). The basic claim is that the domain of such rules, which apply in a number of syntactic configurations, can uniformly be defined in terms of a general definition of Ph, which reads as follows (168):

(7) *Phonological Phrase Formation.*

(I) Ph-domain: The domain of Ph consists of a C which contains a lexical head (X) and all C's on its non-recursive side up to the C that contains another head outside of the maximal projection of X.

(II) Ph-construction: Join into an n-ary branching Ph all C's included in a string delimited by the definition of the domain of Ph.

For instance, the following prosodic structure is assigned to the sentence *Ho visto tre colibri molto scuri* 'I saw three very dark hummingbirds':

(8) (Ho visto)_{Ph} (tre colibri)_{Ph} (molto scuri)_{Ph}.

The rule of Raddoppiamento Sintattico, which lengthens word-initial consonants if the preceding word ends in a stressed vowel and both words belong to the same Ph, will affect the underlined consonants in (8).

Each language is supposed to have a recursive side. In Italian and French, this is the right side. Note, however, that it may depend on the lexical category which side is recursive. For instance, in Dutch the recursive side for A and V is left, for N and P it is right (cf. Hoekstra, 1984). Nespor and Vogel (186, note 2) seem to be aware of this problem, and remark that for such languages the definition of Ph has to be more complicated. Note that this problem cannot be solved by introducing a variable for 'recursive side' the

value of which is determined by the lexical category. For instance, in the Dutch VP [[*geeft*]_V [*het boek*]_{NP}]_{VP} 'gives the book' the determiner *het* is both on the non-recursive side of the V (right) and on the non-recursive side of the N (left). Hence, it remains indeterminate whether it forms a Ph with the preceding V or the following N.

The well-known French rule of Liaison is also claimed to apply within Ph, at least in the casual style of speech. Since the data with respect to liaison are so complicated, it is hard to evaluate this claim within this review. I refer the reader to Booij & de Jong (1987) for an evaluation of the different theories on the domain of Liaison. Alternative proposals can be found in Kaisse (1985) and Selkirk (1986).

The last two prosodic categories proposed by Nespor and Vogel are the Intonational Phrase and the Utterance. The Intonational Phrase consists of a number of Ph's. Its size is variable, and depends on syntactic structure, length of the constituents, style of speech and pragmatic factors. Therefore, it differs from Ph which is defined by reference to syntactic properties only. I is the domain of spreading of intonational contours and the domain of application of certain segmental rules.

The Utterance consists of one or more I's within the domain X^n , usually, but not always a sentence. An Utterance may be smaller than a full sentence in the syntactic sense, e.g. elliptical sentences like *At five* and *Near by*. It can also be larger than a sentence: in case two consecutive sentences have some syntactic or semantic connexion and are uttered by the same speaker without intervening pause, certain segmental rules that apply in U, appear to apply also across sentence boundaries, as in *What a nice sofa [r] Is it new?* These two sentences are related by the anaphoric relation between *sofa* and *it*. Again, we see that there is no isomorphy between syntactic and prosodic structure.

In sum, we have seen that Nespor and Vogel propose a hierarchy of prosodic categories which can be divided into three layers:

- (9) (a) syllable, foot, phonological word;
- (b) clitic group, phonological phrase;
- (c) intonational phrase, phonological utterance.

The prosodic structure under (a) depends on, as we have seen, morphological information, directly (the phonological word) or indirectly (syllable and foot). The prosodic structure under (b) is dependent on syntactic structure and the difference between lexical categories (including clitics). The prosodic structure in terms of (c) also depends on other types of information: phonetic, semantic and pragmatic properties.

Nespor and Vogel's theory is not only a theory about prosodic structure, it is also a theory about the organization of the grammar. It is assumed (cf. 302) that the first rules to apply in the post-lexical phonological component are the mapping rules that transform syntactic structure into prosodic

structure. Leaving out of consideration the fact that, as we saw above, certain prosodic categories must already be derived in the lexicon, this proposal claims that there are no rules of sentence phonology that can refer directly to syntactic information, and that all apparently syntax-sensitive rules can be translated into prosodic rules. Whether this claim can be maintained deserves further investigation. For instance, Selkirk (1986: 374) argues that such syntax-sensitive rules do exist.

Nespor and Vogel's model of sentence phonology also competes with the model of sentence phonology defended in Kaisse (1985). Kaisse proposed a post-lexical phonological component divided into two blocks: a block of external sandhi rules which are directly sensitive to syntactic structure, and – following pause insertion – a block of fast speech rules which are not sensitive to structural information.⁴ Thus we have at least three competing organizational models of sentence phonology, those of Kaisse (1985), Nespor & Vogel (1986) and Selkirk (1986). Whatever model we may finally end up with, I think that Nespor and Vogel have shown convincingly that there must be a block of rules with domain specifications in terms of prosodic categories.

Finally, some short remarks on the more peripheral chapters 9 and 10 which are meant to show the relevance of prosodic theory for other descriptive problems than rule domains. Chapter 9 argues that the possibility of disambiguating sentences depends on the prosodic structure rather than the syntactic structure of sentences. Chapter 10 explains how the verse structure of Dante's *Inferno* can be described in terms of types of matching between the prosodic structure of a verse and its abstract metrical structure.

In conclusion, Nespor and Vogel's monograph is an inspiring study in that it opens new research perspectives in the area of Prosodic Phonology. It proposes a number of well-articulated hypotheses with respect to prosodic structure, some of which may, of course, turn out to be wrong in subsequent research, but all of which help us to acquire a deeper insight into the issues concerning the structure of phonological representations.

*Author's address: Vakgroep Taalkunde,
Vrije Universiteit,
P.O. Box 7161,
1007 MC Amsterdam.*

REFERENCES

- Berendsen, E. & van den Os, E. (1985). Review of Kaisse (1985). *Linguistics* 23, 993–998.
 Booij, G. E. (1983). Principles and parameters in prosodic phonology. *Linguistics* 21, 249–280.
 Booij, G. E. (1984). French C/Ø-alternations, extrasyllabicity and Lexical Phonology. *The Linguistic Review* 3, 181–207.
 Booij, G. E. (1987). On the relation between Lexical Phonology and Prosodic Phonology. To appear in Bertinetto, P. M. & Loporcaro, M. (eds), *Proceedings of the Cortona Phonology Workshop*. Pisa: Scuola Normale Superiore.

[4] See Berendsen and van den Os (1985) for a review of Kaisse (1985) that compares it to the Nespor & Vogel-theory of sentence phonology.

REVIEW ARTICLE

- Booij, G. E. & de Jong, D. (1987). The domain of Liaison. *Linguistics* 25, 1005-1025.
- Booij, G. E. & Rubach, J. (1987). Postcyclic versus postlexical rules in Lexical Phonology. *LIn* 18, 1-44.
- Dixon, R. M. W. (1977). *A grammar of Yidjip*. Cambridge: Cambridge University Press.
- Hockett, Ch. F. (1955). *A manual of phonology. IJAL Memoir 11*. Baltimore: Waverley.
- Hoekstra, T. (1984). *Transitivity*. Dordrecht: Foris.
- Kaisse, E. (1985). *Connected speech. The interaction of syntax and phonology*. Orlando, FL: Academic Press.
- Nespor, M. & Vogel, I. (1982). Prosodic domains of external sandhi rules. In van der Hulst, H. & Smith, N. (eds), *The structure of phonological representations*. Dordrecht: Foris. 225-255.
- Prince, A. (1983). Relating to the grid. *LIn* 14, 19-100.
- Selkirk, E. O. (1978). On prosodic structure and its relation to syntactic structure. Published in Fretheim, T. (ed.), *Nordic prosody II*. Trondheim: Tapir, 1981. 111-140.
- Selkirk, E. O. (1980a). Prosodic domains in phonology: Sanskrit revisited. In Aronoff, M. & Kean, M.-L. (eds), *Juncture*. Saratoga, Calif.: Anma Libri. 107-129.
- Selkirk, E. O. (1980b). The role of prosodic categories in English word stress. *LIn* 11, 563-605.
- Selkirk, E. O. (1984). *Phonology and syntax. The relation between sound and structure*. Cambridge, MA: MIT Press.
- Selkirk, E. O. (1986). On derived domains in sentence phonology. *Phonology Yearbook* 3, 371-405.