Phonological Externalization volume 2 (2017), 25-43.
Hisao Tokizaki (ed.), Sapporo University.

Extending the Application of Merge to Elements in Phonological Representations*

Kuniya Nasukawa
Tohoku Gakuin University

ABSTRACT. Hauser, Chomsky and Fitch (2002) claim that the presence/absence of recursion divides the organisation of the language faculty into two parts: FLB and FLN. The origins of properties in FLB are thought to be shared by non-human species, whereas FLN contains the species-specific operation Merge, which applies repeatedly to syntactic objects to generate recursively-structured expressions. However, this paper claims that Merge applies not only to (morpho)syntactic objects but also to phonological primitives that make up the phonological structure of morphemes. Accordingly, phonological categories are engaged not only in the externalization of internally-constructed expressions but also in internal computation.

Keywords: phonology, recursion, Merge, phono-syntax, elements

1. Introduction

Studying the evolution of language is a challenge because of the lack of fossil records: no physical evidence can be gleaned from spoken languages. To overcome this problem, the last half century has seen efforts to establish an interdisciplinary (cross-field) research platform involving cooperation between various related fields such as linguistics, biology, psychology, neuroscience and mathematics. Recent results show signs of significant progress in the field.

One of the most notable attempts at understanding the evolution of language is an approach which focuses on the relation between the architecture of language and the nature of other capacities of humans. This approach is based on the premise that it is essential to understand the nature of the language faculty before we can begin to identify the origins of human language and establish which properties of human language make it species-specific. This approach is consistent with the Minimalist Program (MP: Chomsky 1993, 1995, 2000a, b, 2001, 2004, 2005, 2007, 2008, 2013) of Generative Grammar, which regards human language as a manifestation of the mental faculty, a biologically-determined organ alongside others such as the visual, respiratory and digestive organs. In principle, MP sets out two key questions: *What is language?* and *Why does it have the properties it has?* It claims that the

---

* As this paper is currently under review, it may be subject to further developments. An earlier version was presented at the 4th Workshop on Phonological Externalization of Morphosyntactic Structure: Universals and Variables (PHEX4) (Niigata University, Japan, 21 February 2017). I thank the participants at the conference for their constructive comments.

faculty of language is a genetically-determined highly-non-redundant system in human beings, and assumes that it is optimally designed in response to our conceptual and physical requirements.

Working within MP, Hauser, Chomsky and Fitch (2002) and Chomsky (2010) suppose that internal language is a system which bridges sound and meaning in a particular way: in more precise terms, it is seen as a computational system (CS) which generates an infinite number of hierarchically-organised (syntactic) objects, each of which underlies an array of instructions to two interfaces: the Sensory-Motor (SM) systems and the Conceptual-Intentional (C-I) systems. This is depicted as follows.

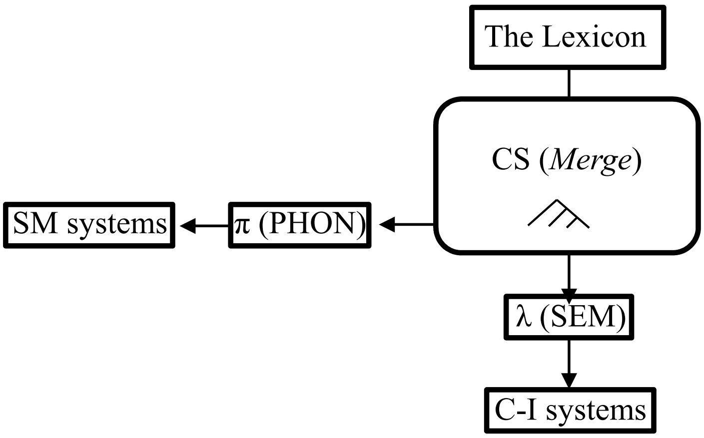

Figure 1 A model of the language faculty in the Minimalist Program

First of all, it is generally agreed that there is a lexicon, a place to store lexical items (i.e. words and morphemes), although there is still controversy over whether the lexicon is actually part of the language faculty or not (Fujita 2009). In order to construct a linguistic expression, lexical items (e.g. α and β) are taken from the lexicon, and are then concatenated by *Merge* to form a new object (e.g. {α, β}). In Chomsky (2010) and Hauser, Chomsky and Fitch (2002), Merge is considered to be the only operation in the CS and it applies not only to lexical objects but also to derived objects which were themselves constructed by Merge (e.g. $\gamma + \{\alpha, \beta\} \rightarrow \{\gamma \ \{\alpha, \beta\}\}$). This repetitive application of Merge to its own derivatives (in the CS of the language faculty) generates an infinite number of recursively structured expressions, which are then thought to be submitted to the interface systems, Sensorimotor (SM) and Conceptual-Intentional (C-I).

Assuming this derivational flow in the language faculty, Chomsky (2010, pp. 59–62) claims that language is perfectly designed to satisfy the conditions that are required at the C-I interface, and that the mapping of recursively structured expressions to the SM interface provides a maximally efficient solution to cognitive problems. In other words, the use of Merge to map constructed expressions into a form interpretable at C-I is the primary process of language. By contrast, processes of externalization are regarded as secondary processes, since they are constrained by UG (Universal Grammar: internal language).

Recursion is the only operation in the CS, and is claimed to be the fundamental property which distinguishes language from other cognitive faculties (Hauser, Chomsky and Fitch 2002). This divides the faculty of language into two parts: FLN (the faculty of languages in the narrow sense) and FLB (the faculty of languages in the broad sense).

(1) The FLN-FLB division

a. FLN (the faculty of language in the narrow sense):
a subset of FLB. FLN is unique to humans and to language.
Recursion found in syntactic computation is the central property of FLN.

b. FLB (the faculty of language in the broad sense):
no part of language is necessarily unique to humans or to language.
FLB consists of the sensory-motor (SM) systems, conceptual-intentional (C-I) systems and possibly other systems.

It is thought that FLN, which makes use of recursion, has no equivalent in non-humans. On the other hand, equivalents to FLB are presumed to be possible also in non-humans.

Let us now consider the question of where phonology is located in the faculty of language, although the answer depends on how ‘phonology’ is defined. Assuming the division between FLN and FLB, phonology is generally thought to lie in FLB because phonology as a whole, and morpheme-internal phonology in particular, is widely considered to be free of recursion (Pinker and Jackendoff 2005, Neeleman and van de Koot 2006, Samuels 2009, Scheer 2008, 2011). This view reflects the common belief that the phonological properties of a morpheme are arranged in the form of linearly-ordered segments like beads on a string (Bromberger and Halle 1989, et passim), and not constructed recursively by Merge. In addition, it is assumed that phonology is an interpretive device rather than a structure-building device since phonological regularities are thought to operate on recursively-constructed objects in CS.

It is obvious that the presence/absence of recursion is a key consideration not only in the debate on the organisation of the language faculty but also for the question of the origins of language itself. From an evolutionary perspective it has been suggested that human language in its present form emerged from single-word expressions by employing a recursive merge device (Reuland 2009, Chomsky 2010, van der Hulst 2010). This naturally leads us to assume that recursion is found only in syntax. However, if a syntax-like recursion operation is also found in phonology, then this prompts us to reconsider the organisation of the language faculty—a move which impinges on the study of the origins of language. Below I outline what kinds of structure are relevant to the study of phonology and what roles are played by phonology in the faculty of language. I also address the issue of whether recursion operates in phonology as it does in syntax.

This paper is organized as follows. Section 2 considers whether syntax-like recursive structure is present in phonology. Then section 3 discusses the similarities and differences between sentence (phrase) construction and word (morpheme) construction, and argues for the need to eliminate the linear ordering of segments from phonological representations. Section 4 investigates how the phonological shape of morphemes is obtained through recursive Merge, as found in syntax. Furthermore, by introducing feature-like units called elements as the building blocks of phonological structure, it will be shown that the same repeated application of Merge to elements can be used to construct morpheme-internal phonological structure. Section 5 presents arguments concerning the overall shape the language faculty in the context of the proposed model of phonologically recursive structure. It also reinforces the claim that the operation Merge targets not only morpho-syntactic objects but also phonological primitives. The final section discusses the implications of the proposed model of the language faculty for our understanding of language evolution.

2. Recursion

Chomsky (2010) argues that recursive structure is the result of the repeated application of Merge in CS. The operation Merge simply concatenates two lexical items (e.g. α and β) to create a new object (e.g. {α, β}), where no recursive structure is observed at this newly created level. Then following this, Merge applies again to the derived object (e.g. γ + {α, β}) and the result is a form which exhibits recursive structure (e.g. {γ {α, β}}). Some examples are given below.

(2) a. Relative clause inside a relative clause (Pinker and Jackendoff 2005, p. 211, cf. Nasukawa 2015, p. 215)
a book [that was written by the novelist [you met on the night [that we decided to buy the boat [that you liked so much]]]]
b. Prepositional phrase inside a prepositional phrase (Schreuder, Gilbers and Quené 2009, p. 1244, cf. Nasukawa 2015, p. 215)
the American [in the desert [on a horse [with no name]]]
c. Relative clause inside a relative clause
[[[The thief] the policeman caught] started to cry]
d. Genitive inside a genitive (van der Hulst 2010, p. xxvi)
[[[John's] sister's] dog's] bone was found in the yard

The sentence in (2a) is a typical tail recursion in the left edge of an expression which involves a right-branching structure: a relative clause ([that you liked so much]) embedded inside another relative clause ([that we decided to buy the boat]) which is furthermore embedded inside another relative clause ([you met on the night]) which is also embedded inside another relative clause ([that was written by the novelist]). The same pattern is found in (2b), where a prepositional phrase ([with no name]) is embedded within another prepositional phrase ([on a horse]) which is itself embedded in another prepositional phrase ([in the desert]). On the other hand, there is a left-edge tail recursion which involves a left-branching structure: the subject position in (2d) shows a recursive structure where the relative clause [(whom) the policeman caught] is embedded inside the noun phrase [The thief]. The same form of embedding is also found in the sub-modifying genitives in (2d). Each structure in (2) has one constituent embedded within another; that is, each exhibits recursive structure.

The next question is whether a syntax-like recursion is found in phonology too, where phonology is defined as a set of regularities over the categories that are interpreted by the SM systems. Before discussing the presence/absence of recursion in phonology, we are first required to consider the roles of phonology in the faculty of language. It is generally assumed that phonology has a dual role, because the phonetic realization of a syntactic expression (sentence) refers not only to the phonological structure of individual constituent words, as expressed as (3a), but also to the phonological effects resulting from the concatenation of those words, as described in (3b) (Harris 1994, p. 1–2).

(3) a. to make the phonological attributes of morphemes stored in the lexicon available to the Articulatory-Perceptual systems after syntactic objects are constructed

b. to map syntactic objects constructed by the syntactic computational system (CS) on to forms which are interpretable by the Sensory-Motor (SM) systems

(3a) concerns morpheme-internal phonological patterns such as assimilatory effects, segmental suppression and stress assignment, all of which are often referred to as dynamic alternations in the literature (Harris 1994, p. 3–6). On the other hand, (3b) involves the linearlisation and stress/tonal assignment of morpho-syntactic structure (cf. Kayne 1994). As just noted, both of these roles are carried out after syntactic objects are constructed, in order to submit PHON to the interface. Accordingly, phonology must be located between CS and the interface, as illustrated in Figure 2(b).

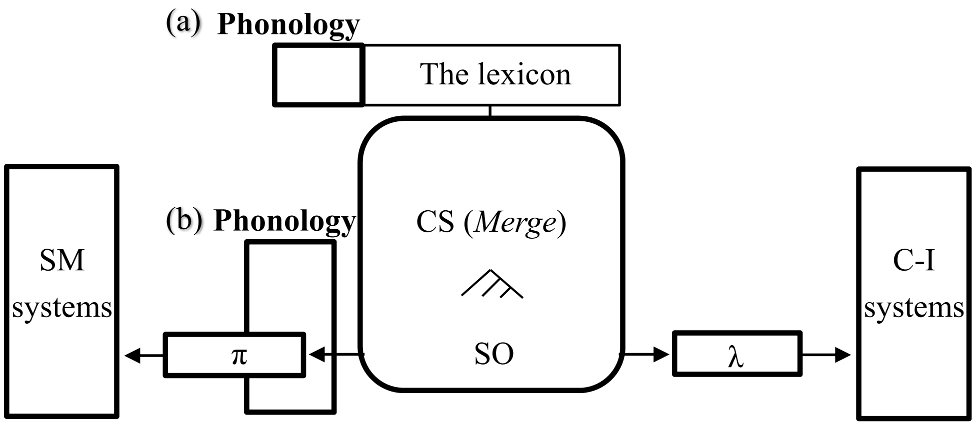

Figure 2 The distribution of phonology in the faculty of language

What about static distributional patterns, such as phonotactics, which also characterize the phonological shape of lexical items? In general, the phonological shape of an item (morpheme or word) listed in the lexicon is expected to conform to the static distributional regularities of a given language. This suggests there is another type of phonology which works when we store morphemes/words in the lexicon, as shown in Figure 2(a). However, the idea that two types of phonology are present at different points in the derivational flow within the faculty of language appears to be out of step with the spirit of minimalism, since it makes the system more complex; moreover, it is difficult to pinpoint any direct relation between the two sound-related components concerned. Since the existence of phonology in Figure 2(b) is widely accepted in the literature (e.g. Scheer 2004, Samuels 2011, Idsadi and Raimy 2013), we should focus on the question of whether the phonological component Figure2(a) actually exists or not. To do so involves establishing the nature of the mechanism responsible for the phonological shape of lexical items stored in the lexicon.

3. Similarities between sentence construction and word construction

The ability to Merge items recursively makes it possible for human language to create any sentence of any length (insofar as it conforms to a grammar). A sentence can be extended indefinitely by embedding syntactic objects inside larger ones, as in (4).

(4) a. [[[The thief] the policeman caught] was gazing at the building] ...
    b. [James assumes [that John believes [that Beth thinks [that Michael means ...

In this way, sentences generated by a grammar are infinite not only in length but also in kind. The fact that we rarely encounter very long sentences in performance is more a reflection of our limited memory capacity than the mechanism of syntactic computation.

Taking a look at the phonological shape of words, we find a similar situation. The words of a language vary greatly in length: some consist of just a single vowel (e.g., the article [ə] ‘a’) while others have 40 or more segments (in cases where internal morphology is involved), as shown in (5a).

(5) a. The longest (artificial and technical) word in the Oxford English Dictionary
    Pneumonoultramicroscopicsilicovolcanoconiosis (generally known as silicosis)
    /njuːˌmɒnoʊˌʌltrəˌmaɪkrəˈskɒpɪkˌsɪlɪkoʊvɒlˌkeɪnoʊˌkɒnaɪˈoʊsɪs/ (44 segments)
    (The name of a lung disease caused by inhaling very fine ash and sand dust)
    b. A long monomorphemic place name
    Massachusetts /ˌmæsəˈtʃuːsɪts/ (10 segments)

Place names sometimes provide examples of long words comprising a single morpheme. The word *Massachusetts* in (5b) is one instance, with 10 segments and 4 syllables.

Like sentences, there is in principle no limit on the length of a phonological word, although long words typically have some form of internal morphology. In contrast, however, words cannot be generated by combining any random sequence of segments. Rather, there are grammatical (so-called 'phonotactic') conditions on the phonological shape of words: for instance, restrictions on the sequence of syllable-initial consonants (e.g., 'sk-' is possible in English whereas 'ks-' is not) and restrictions on the hetero-syllabic sequence of consonants (e.g., '-r.t-' is possible whereas '-t.r-' is impossible). An implication of these examples is that words that are generated by a grammar (e.g. phonotactics) are also infinite both in length and in kind. The fact that we rarely encounter very long words stored in the lexicon must be a consequence of the limitations of our long-term memory rather than the mechanism of word construction.

Thus, word construction and sentence construction appear to be similar in the sense that an infinite number of objects (expressions) can be constructed from a finite repertoire of units (i.e. segments in word construction, syntactic units in sentence construction). On the other hand, a crucial point which makes word construction different from sentence construction has been believed to be the absence/existence of recursion: as summarized in (6), recursive structure is found in sentences that are constructed by Merge at the CS, whereas the sound shape of words has generally been seen as the result of concatenating segments in a linear fashion, where no recursion is involved. This state of affairs partly results from the view that the phonological properties of a lexical item, in some pre-theoretical sense, take the form of a bead-like string of segments (Bromberger and Halle 1989). This view is presumably based on the following premises (Nasukawa 2015, p. 212).).

(6) Two widespread premises concerning the phonological shape of lexical items

a. Morpheme/word-internal phonological structure consists of a set of linearly-ordered segments in the lexicon.

b. Phonology is a module which merely interprets fully concatenated strings of morphemes. Phonology is not responsible for constructing phonological structure in the lexicon.

Given the premises in (6), most phonological theories assume that lexical items have phonological representations that are structurally flat, for which (unlike in syntax) no recursive device is needed (Bromberger and Halle 1989, Raimy 2000, Scheer 2004, Samuels 2009).

Let us return to the issue of how the phonological structure of lexical items is created. The concatenation of segments in a linear fashion involves the simple task of adding segments one after another to a pre-existing sequence, so it seems theoretically trivial to even discuss the mechanism by which the sound shape of words/morphemes is constructed. In linguistics textbooks it is usually stated that a lexical item (word or morpheme) is a set of morpho-syntactic, semantic and phonological properties, yet we rarely find any detailed explanation of how those phonological properties are organized. (However, the analysis of morphologically-driven word formation is often discussed in the scholarly literature (e.g. Lexical Phonology: Kiparsky 1982, 1985, Distributed Morphology: Marvin 2002, Marantz 2007).)

While there is a general lack of explanation for the phonological structure of lexical items, it must be admitted that, by constructing a linear string of ordered segments for lexical items, we are forced to assume the existence of the module in Figure 2(a), which is independent of the phonological module located between CS and PHON in Figure 2(b). But as noted above, there are problems in assuming that the grammar contains two different components responsible for the organization of sounds, each in a different location within the faculty of language. Specifically, this situation should be avoided because it goes against the spirit of minimalism, and moreover, it is difficult to explain how these two phonological components are related without introducing another path of derivation between them.

However, it is possible to take an alternative view of the phonological structure of lexical items. Below I will discuss the idea developed in Nasukawa (2015, p. 212–213), in which it is shown that whether or not a given domain is characterized as having recursive structure depends on how one characterizes the basic units of phonological structure and their organization. In the Precedence-free model of phonological representation, the following two premises are proposed in place of those given in (6).

(7) a. Morpheme-internal phonological structure consists of no segment-based precedence information, but of a set of primitives which are hierarchically concatenated by Merge.

b. Phonology is responsible for generating expressions in PHON, which are readable by the SM systems.

Although the details differ between one theory and another, the premise in (7b) is widely shared by different models of phonological representation. On the other hand, the premise in (7a) plays a core role in Precedence-free Phonology but is rarely employed in other theoretical approaches. (7a) derives from the assumption that information which specifies the linear order of segments in a lexical item is redundant; instead, dependency relations holding between phonological units are sufficient to account for linear ordering and also phonological processes. A precedence-free structure relies on the presence of embedded units in morpheme-internal phonological structure. In the following section, therefore, it is claimed that recursion exists in the phonological structure of morphemes.

4. An alternative model of components relevant to phonological properties

4.1 Recursive Merge and phonological primitives

Recall the model of the faculty of language (Figure1). In order to construct a linguistic expression, lexical items are taken from the lexicon, and are then concatenated by *Merge* to form a new object. The repetitive application of Merge to its own derivatives generates an infinite number of *recursively structured* expressions, which are then submitted to the interface systems, SM and C-I. Taking the premises in (7) in place of those in (6), I claim that we employ the same mechanism for constructing the phonological shape of linguistic forms which are stored in the brain. But in the case of phonological forms—unlike syntactic objects, where the basic units for building structures are words—the units for building representations are phonological primitives (features). As illustrated below, in order to construct a linguistic expression, primitives (e.g. α and β) are taken from the inventory of primitives and are then concatenated by *Merge* to form a new object (e.g. {α, β}).

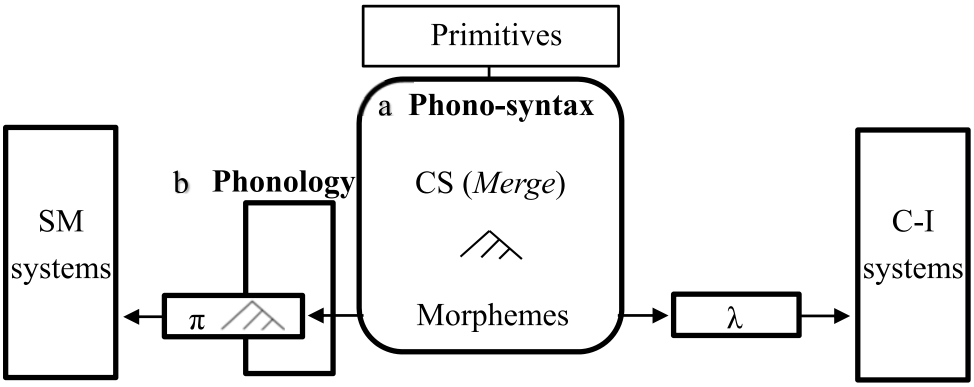

Figure 3 **Phonology in the faculty of language**

The same operation also applies to derived objects which were themselves constructed by Merge (e.g. $\gamma + \{\alpha, \beta\} \rightarrow \{\gamma\ \{\alpha, \beta\}\}$). The repeated application of Merge to its own derivatives generates recursively structured expressions (‘syllable’ and ‘foot’ may be viewed as particular levels to which primitives are ‘projected’: for a detailed discussion, refer to Nasukawa 2016a), which are submitted to the two interface systems.

At this stage the constructed objects, so-called ‘morphemes’, contain phonological primitives but no other type of information such as semantic or syntactic properties. In the sound portion of Figure 3, therefore, all properties (primitives and hierarchical structure consisting of inter-prime relations) are visible: they are all subject to a set of rules or constraints (e.g. assimilatory effects, segmental suppression, stress assignment, and possibly linearization in a broad sense) forming the phonology which is located between CS and the interface. So the outcome (π), which is submitted to the interface, must only contain properties that are readable by the SM systems.

On the other hand, the content of the object constructed at the CS is invisible to the C-I, since non-phonological properties are absent within the object. The role of the C-I is to associate the entire constructed object with a particular conceptual (physical) referent.

4.2 Elements as basic units for structure building in phono-syntax

At this point, we turn to the question of which units are involved in building the phonological structure of morphemes. In order for Merge to operate in phono-syntax, the model in Figure 3 must employ phonological primitives of a type which are not structurally fixed and which may concatenate freely. Furthermore, it must be possible for each primitive to be present without the need for support from the other primitives. The units which best fit these criteria are those developed in the Element Theory approach (Kaye, Lowenstamm and Vergnaud 1985, Harris 1994, et passim), and in particular, in the version of Element Theory developed in Nasukawa (2015, 2016a). The units in question are called *elements*, and they have a central role in building phonological structure.

The model presented in Nasukawa (2015, 2016a) follows other versions of Element Theory in viewing elements as being single-valued (monovalent: presence vs. absence) for contrastive purposes. Also, each element can be phonetically realised on its own; that is, it does not require support from the presence of other elements. Therefore, Element Theory does not require any template-like organization of primitives or any universally-fixed matrix of primitives.

In Element Theory, the phonetic exponence of elements is based on perceptual properties (Nasukawa and Backley 2008, Backley 2011, Nasukawa 2015). Phonological structure is built using the six monovalent elements |A I U ʔ H N|, which are thought to be species-specific and all active in human languages. The six elements are given in Table 1 along with their principal phonetic attributes (Nasukawa 2014, p. 3).

elements	*label based on acoustic pattern*	*manifestation as a consonant*	*Manifestation as a vowel*
|A|	'mAss'	uvular, pharyngeal POA	non-high vowels
|I|	'dIp'	dental, palatal POA	front vowels
|U|	'rUmp'	labial, velar POA	rounded vowels
|ʔ|	'edge'	oral or glottal occlusion	laryngealised vowels
|H|	'noise'	aspiration, voicelessness	high tone
|N|	'murmur'	nasality, obstruent voicing	nasality, low tone

Table 1 Elements

As already mentioned, each element can be pronounced independently. In the syllable nucleus, for example, where a given element is realised as a vowel, the first three elements |A I U| correspond to the phonetic values [ə ɨ ɯ] respectively. These vowels typically behave as default vowels in phenomena such as vowel epenthesis, vowel-zero alternation and vowel reduction. This model assumes that one of these elements |A|, |I| and |U|, which are interpreted as default vowels when they appear in isolation, is selected by parameter as the head of a phonological structure: for example, |A| ([ə]) is the foundation for English, |I| ([ɨ]) for Cilungu and |U| ([ɯ]) for Japanese. As Nasukawa (2014, 2016b) discusses, the reason why these single-element structures are realized as the default vowels [ə], [ɨ] and [ɯ], rather than as full vowels such as [a], [i] and [u], is that a structurally simplex object is assumed to provide only minimal and non-contrastive information.

By comparison, full vowels possess much more lexical information. This means that, in acoustic terms, they have phonetically salient segmental and suprasegmental attributes, and in representational terms, they have greater structural complexity (Nasukawa and Backley 2015, Nasukawa 2016b). This additional complexity comes in the form of Head-Dependency relations, as illustrated in (4). First compare the representations of [ə] in Figure 4(a) and [a] in Figure 4(b).

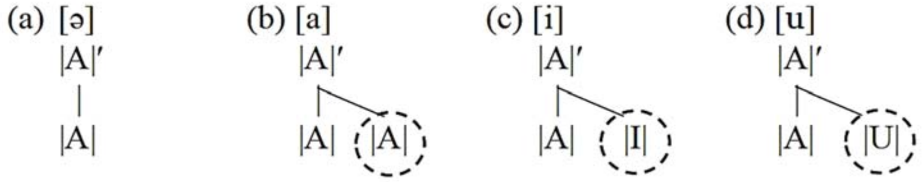

Figure 4 Default vowels vs. full vowels in English

When a second |A| is added to the base element |A|, it is the base element which functions as the head while the additional element takes the dependent role. In the structure in Figure 4(b), the acoustic characteristics of the dependent |A| (an exaggerated form of the baseline vowel [ə]) are superimposed onto the schwa-type baseline (the head |A|). As a result, the whole structural expression is phonetically realised as [a]. This follows the claim in Nasukawa and Backley (2015) that heads, just like heads in syntax, are structurally important but lexically/informationally unimportant whereas dependents are structurally unimportant but lexically/informationally important. And this difference is reflected in the phonetic realization of organizing units: dependents show a greater acoustic prominence than heads in terms of the size of their modulated carrier signal (Nasukawa 2016b, cf. Harris 2006). Likewise, examples (c) and (d) in Figure 4 also show that a dependent element is crucial in providing additional lexical information through its phonetically prominent attributes.

Since most languages have vowel systems which contain vowels other than those in Figure 4, further concatenation of elements is needed in order to generate other vowel contrasts. For example, many languages have a five-vowel system which has, in addition to [a], [i] and [u], the mid vowels [e] and [o]. To represent these mid vowels, theories which refer to the primitives |A|, |I| and |U| (e.g. Element Theory, Dependency Phonology) employ the structures |A|+|I| (for [e]) and |A|+|U| (for [o]). In the present model of element concatenation, the further addition of |A| to the structures in Figure 4(c) and Figure 4(d) generates the structures for [e] and [o] respectively.

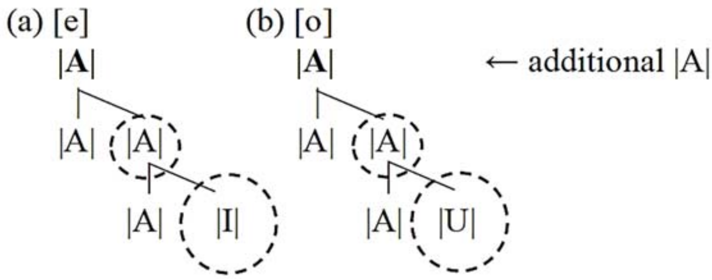

Figure 5 Full vowels [e] and [o] in English

To repeat the point made earlier, the recursive model being presented here gives dependents greater prominence than heads in terms of the size of their modulated carrier signal (Nasukawa and Backley 2015). In Figure 5(a) and Figure 5(b), therefore, the most deeply embedded elements |I| and |U| are acoustically the most prominent in their respective structures: the phonetic pattern of the lowest (dependent) element makes the greatest contribution to the whole expression. As for the |A| heads, which are acoustically weak in both cases, their acoustic signatures are masked by those of the dependent elements. The degree of phonetic saliency of the intermediate |A|s also reflects their relative position in the structure, in that an intermediate position produces an intermediate level of prominence. What these simple illustrations show is that phonetic interpretation depends on which elements are present and also on the headedness of their concatenated structures. Further successive levels of embedding can be introduced recursively until all the required vowel categories are uniquely represented. Furthermore, the structure of consonant sounds, as well as that of higher-level constituents equivalent to 'syllable', 'foot' and 'morpheme' are also represented following the same principle of phonetic realisation (i.e. that dependents are phonetically more prominent than heads—see Nasukawa (2016a) for a detailed account of these embedded, recursive representations). Constructed forms such as those described above are thought to be stored in memory.

It is not feasible to apply this kind of approach to models of morpheme-internal phonological representation such as Feature Geometry (FG), which employ a template-based segment structure. This is because FG-based models do not incorporate any structure-building process. As discussed in Nasukawa (2015, p. 220), in FG and other approaches which refer to traditional distinctive features, those features function as minimal contrastive units yet they are not the basic units for building segmental structure. In feature-based theories, the basic units of segmental structure are the segments themselves, as these are the minimal units of phonetic interpretation (i.e. a single feature such as [+high] cannot be phonetically realized by itself). This situation stems in part from the fact that most phonological theories see morpheme-internal phonological structure as comprising a set of linearly-ordered segments.

Finally, it was noted above that elements are primarily associated with acoustic properties since the language faculty employs the auditory channel of transmission facilities (the SM systems) in the externalization process. In the case of sign languages, however, it appears that secondary elements could also be externalized via the visual channel if the auditory channel is unavailable. Although there is a clear preference in terms of the relation between elements and (acoustic) modality, the primitives are by nature modality-neutral. (For a more radical discussion of modality-neutral primitives, see Hale and Reiss 2008).

5. The overall shape of the language faculty

It has been acknowledged that syntax-like concatenation is involved in word formation: that is, morphemes are also concatenated by Merge for constructing words (cf. Distributed Morphology: Marvin 2002, Marantz 2007). To construct a linguistic expression, morphemes (e.g. α and β) are concatenated by Merge to form a new object (e.g. {α, β}). Then, the subsequent and repeated application of Merge to its own output form generates a recursively-structured expression at the word level. When this mechanism of morphology is combined with those of syntactic computation in (2) and the phono-syntax in (13), the following view emerges of the architecture of the language faculty as a whole (π = PHON, λ = SEM, SO = syntactic object).

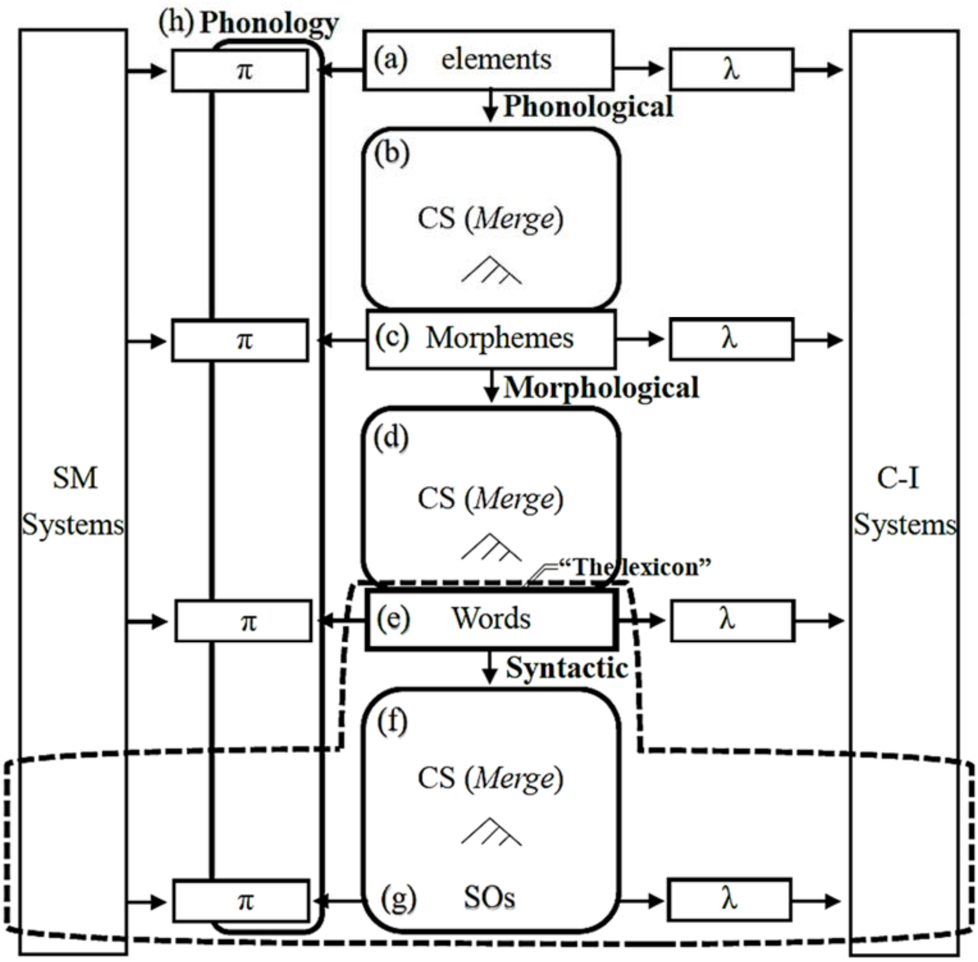

Figure 6 **A proposed model of the language faculty**

The phono-syntax is located in Figure 6(b) where, as illustrated in Figure 3, elements which have been selected from Figure 6(a) are concatenated recursively by Merge in the creation of structures like Figure 4 and Figure 5; these form the phonological structure of morphemes. They are then stored in the brain Figure 6(c) and can be used as building blocks for constructing words. This may involve at least two distinct levels such as the Root-level (cf. Level 1) and the Word-level (cf. Level 2, post-lexical) (as discussed in Lexical Phonology: Kiparsky 1982, 1985 and Government Phonology: Kaye 1995). Most of the constructs are stored in the brain Figure 6(e) and are then available for sentence (phrase) construction. Words are selected from Figure 6(e)—an entity which the linguistics literature typically refers to as the lexicon—and these are then concatenated by Merge to form a new linguistic expression. The repeated application of Merge to constructed objects generates a hierarchically organised sentential/phrasal structure, which is submitted to the two interfaces (SM and C-I) as PHON (π) and SEM (λ). Assuming that the derivational flow proposed above is correct, then it turns out that the entity traditionally labelled 'the lexicon' (if such an entity exists) in the syntax literature is actually located at Figure 6(e); this is the module from which lexical items are selected and then concatenated to form syntactic structure. In the model proposed here, however, I claim that there are at least three locations for the storage of categories/objects: storage for primitives (elements) Figure 6(a), storage for morphemes Figure 6(c) and storage for words Figure 6(d). It is an open question as to whether they are located inside the 'lexicon' or whether they form part of the faculty for short-term memory.

Regarding PHON, since its function is to externalize the internal objects that have been constructed by Merge, the properties of PHON must readable by the SM systems. These include the phonological primitives (elements) and their relational properties (hierarchical structure). In addition, the relational properties in words Figure 6(e) and phrases/sentences Figure 6(g) are also required to form part of PHON since prosodic phenomena such as word-level stress-assignment, phrasal stress patterns and pitch assignment are dependent on these relational properties. In this mode, there is no translation between syntactic structure and prosodic structure (cf. Selkirk 1984). In the pursuit of minimalism, this paper assumes that there is no structural reorganization of objects which have already been constructed. Rather, properties are subject to visibility requirements. As discussed in the literature (Chomsky 1995), visibility is determined by linguistic category: phonological categories are visible while morpho-syntactic and semantic categories are invisible on the left-hand side of Figure 6; on the other hand, phonological categories are invisible while non-phonological categories are visible on the right-hand of Figure 6. In contrast, I claim that *all structural properties created by Merge are readable not only at the C-I interface, but also at the SM interface*. Phono-syntactic structure created at Figure 6(b) is necessary for determining morpheme-internal stress and tonal assignment where segments are weakened or strengthened. To carry out the same processes at the word level, morphological structure established at

Figure 6(d) is also needed at the SM interface. Furthermore, any syntactic structures (spelled-out by phase) are assumed to be essential in determining phrasal and sentential stress patterns. Likewise, all relational properties are also thought to be necessary for the C-I side.

It should be noted that it is not only syntactic objects constructed by Merge at Figure 6(g) but also words in Figure 6(e) and morphemes in Figure 6(c) which may be submitted to the SM and C-I interfaces. This is because any object from 6(c), 6(e) or 6(f) may have both a sound (phonological) association and a meaning (morpho-syntactic or semantic) association. In fact, words and morphemes which have not been incorporated into syntactic expressions are often referred to in the description of linguistic behaviour. Additionally, elements (i.e. the basic units of phonological structure) may be associated with semantic properties. For example, there is a cross-linguistic tendency for the element |A| (the mAss pattern) in adjectives to be linked to semantic properties such as large, broad and heavy, while |I| (the dIp pattern) may be related to properties such as small, narrow and light (cf. sound symbolism: Ohala 1983, 1994). As for |H| (aspiration), it may denote properties such as strong, big, clear in aspiration languages like Korean. The image associated with a particular element is explained in terms of its relation with the two interfaces.

In the proposed model it is assumed that constructed objects are subject to a phonology—that is, a set of rules or constraints which may include assimilatory effects, segmental suppression, stress assignment, and possibly linearization in a broad sense—which is located between CS Figure 6(b, d, f) and the SM interface. It follows, then, that there are two different places where phonological categories are referred to: Figure 6(b) and Figure 6(h). Figure 6(h) refers to phonology, which is traditionally assumed in the literature, while Figure 6(b) refers to a version of syntax which deals with phonological elements as opposed to morpho-syntactic categories/objects. In this sense, all forms constructed by Merge at Figure 6(b, d, f) are syntactic objects—they differ only in terms of the types of constituents involved, as indicated in Figure 7.

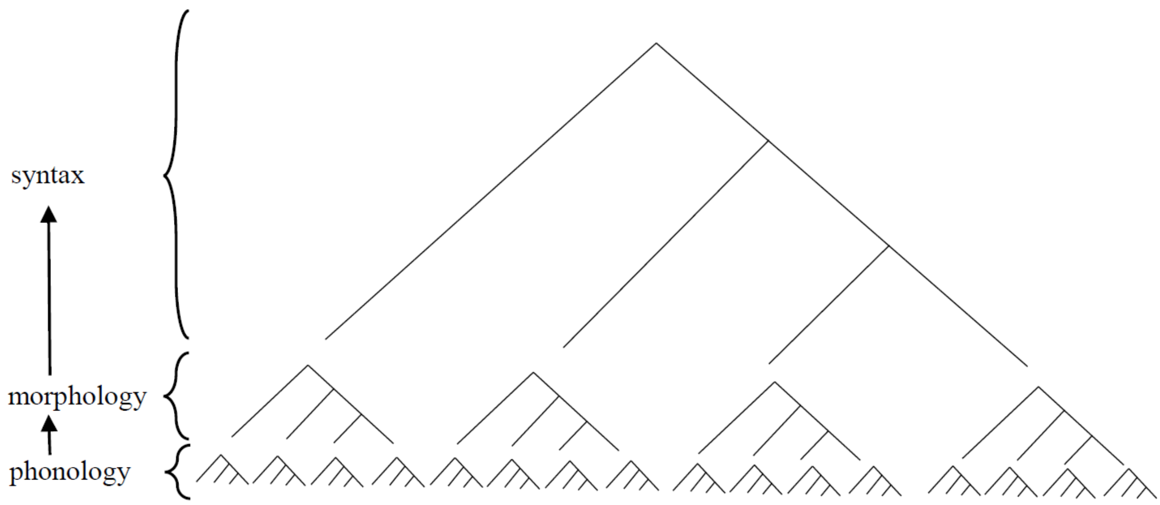

Figure 7 **The overall shape of recursive structure constructed by Merge**

6. Evolutionary perspectives

As mentioned in section 1, in the discussion of language evolution Chomsky (2010) claims that the conceptual interface is primary. Guided by third factor principles, the operation of unbounded Merge emerged and it started to apply to concepts with intricate properties (emergence of the CS for thought). Once its genetically-basic properties had spread to a sufficient number of humans (in a small community), the SM systems—which appear to have been intact for hundreds of thousands of years—are assumed to have been exploited for externalizing expressions constructed by the CS.

If this line of argument is on the right track, then in conjunction with the proposed model of the language faculty in Figure 6 it is possible to claim that objects (categories) of any kind which are concatenated by Merge (elements, morphemes and words) are concepts not originally associated with the SM interface: even elements in Figure 6(a) must be phonology-free objects. This is consistent with the view discussed in section 4.2 that elements are modality-neutral in nature (cf. Hale and Reiss 2008). On this basis, one of the fundamental tasks of externalizing conceptual objects must be to learn how to map elements on to phonetic exponents. Thus it may be assumed that humans started to acquire the ability to produce more intricate and complex phonetic patterns involving, for instance, linearisation and differing degrees of acoustic saliency, following the derivational flow Figure 6(a) to Figure 6(g). I assume that this intricate externalization process related to the SM systems is the source of the variation and diversity within human languages.

Interestingly, this appears to be similar to the phonological path of language acquisition: from cooing to the stage of producing simple sentences via vocal play, babbling, the one-word stage and the two-word stage (Sachs 1985, Clark 2003). Although this is an open question, there could be a connection between the process of externalization and the developmental path in phonological acquisition.

7. Concluding remarks

This paper has claimed that the species-specific operation Merge manipulates not only (morpho)syntactic objects but also phonological primitives (elements) for the purposes of constructing the phonological shape of morphemes. This implies that phonological categories are not only employed for externalization of internally-constructed expressions but are also objects for internal computation. Assuming this to be the case, phonological elements must be viewed as conceptual material that is modality-neutral in externalization. The reason why elements are primarily associated with phonetic (especially acoustic) attributes is attributed to perceptual efficiency (auditory externalization is more efficient than visual externalization for sign languages since the former functions in spite of darkness) and articulatory economy (intricate and delicate motor skills for externalization are more easily executed by the vocal organs than by other parts of body).

In cooperation with related fields, further investigation will be of course needed to validate the proposals made in this paper.

References

Backley, Phillip. 2011. *An introduction to Element Theory.* Edinburgh: Edinburgh University Press.

Bromberger, Sylvain and Morris Halle. 1989. Why phonology is different. *Linguistic Inquiry* 20. 51–70.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel J. Keyser, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000a. Minimalist inquiries: The framework. *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000b. *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.

Chomsky, Noam. 2001. Derivation by Phase. *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. *Structures and beyond: The cartography of syntactic structures*, ed. by Adriana Belletti, 104–131. Oxford: Oxford University Press.

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

Chomsky, Noam. 2007. Approaching UG from below. *Interfaces + recursion = language?*, ed. by Uli Sauerland and Hans-Martin Gärtner, 1–29. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2008. On Phases. *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero and Maria Luisa Zubizarreta, 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2010. Some simple evo devo theses: How true might they be for language?” *The evolution of human language: Biolinguistic perspectives*, ed. by Richard K. Larson, Viviane Déprez and Hiroko Yamakido, 45–62. Cambridge: Cambridge University Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.

Clark, Eve V. 2003. *First language acquisition*. Cambridge: Cambridge University Press.

Hale, Mark and Charles Reiss. 2008. *The phonological enterprise*. Oxford: Oxford University Press.

Harris, John. 1994. *English sound structure*. Oxford: Blackwell.

Harris, John. 2006. The phonology of being understood: Further arguments against sonority. *Lingua* 116. 1483–1494.

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

Hulst, Harry van der. 2010. Re recursion. *Recursion and human language*, ed. by Harry van der Hulst, xvi-liii. Berlin: Mouton de Gruyter.

Idsadi, William J. and Eric Raimy. 2013. Three types of linearization and the temporal aspects of speech. *Challenges to linearization*, ed. by Theresa Biberauer and Ian Roberts, 31–56. Berlin: Mouton de Gruyter.

Kaye, Jonathan. D. 1995. Derivation and interfaces. *Frontiers of phonology: Atoms, structures, derivations*, ed. by Jacques Durand and Francis Katamba, 289–332. Harlow, Essex: Longman.

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.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kiparsky, Paul. 1982. From cyclic to Lexical Phonology. *The structure of phonological representations, vol. I*, ed. by Harry van der Hulst and Norval Smith, 131–75. Dordrecht: Foris Publications.

Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook* 2. 83–138.

Marantz, Alec. 2007. Phases and words. *Phases in the theory of grammar*, ed. by Sook-Hee Choe, 191–222. Seoul: Dong In.

Marvin, Tatjana. 2002. Topics in the stress and syntax of words. Doctoral thesis, Massachusetts Institute of Technology.

Nasukawa Kuniya. 2014. Features and recursive structure. *Nordlyd* 41(1). 1–19.

Nasukawa Kuniya. 2015. Recursion in the lexical structure of morphemes. *Representing structure in phonology and syntax*, ed. by Marc van Oostendorp and Henk van Riemsdijk, 211–238. Berlin: Mouton de Gruyter.

Nasukawa Kuniya. 2016a. A precedence-free approach to (de-)palatalisation in Japanese. *Glossa: a journal of general linguistics* 1(1): 9. 1–21.

Nasukawa, Kuniya. 2016b. The phonetic salience of phonological head-dependent structure in a modulated-carrier model of speech. Ms. Tohoku Gakuin University.

Nasukawa, Kuniya and Phillip Backley. 2008. Affrication as a performance device. *Phonological Studies* 11. 35–46.

Nasukawa Kuniya and Phillip Backley. 2015. Heads and complements in phonology: A case of role reversal? *Phonological Studies* 18. 67–74.

Neeleman, Ad and Hans van de Koot. 2006. On syntactic and phonological representations. *Lingua* 116. 1524–1552.

Ohala, John J. 1983. The phonological end justifies any means. *Proceedings of the XIIIth International Congress of Linguists, Tokyo, 29 August – 4 September 1982*, ed. by Shirō Hattori and Kazuko Inoue, 232–243. Tokyo: Proceedings Publishing Committee. [Distributed by Sanseido Shoten.]

Ohala, John J. 1994. The frequency code underlies the sound symbolic use of voice pitch. *Sound symbolism*, ed. by Leanne Hinton, Johanna Nichols and John J. Ohala, 325–347. Cambridge: Cambridge University Press.

Pinker, Steven and Ray Jakendoff. 2005. The faculty of language: What's special about it? *Cognition* 95. 201–236.

Raimy Eric. 2000. *The phonology and morphology of reduplication*. Berlin: Mouton de Gruyter.

Sachs, Jacqueline. 1985. Prelinguistic development. *The development of language*, ed. by Jean Berko Gleason, 37–60. Columbus, OH: Charles E. Merrill Publishing Co.

Samuels, Bridget. 2009. The third factor in phonology. *Biolinguistics* 3.2. 355–382.

Samuels, Bridget. 2011. *Phonological architecture: A biolinguistic perspective*. Oxford: Oxford University Press.

Scheer, Tobias. 2004. *A lateral theory of phonology: What is CVCV and why should it be?* Berlin: Mouton de Gruyter.

Scheer, Tobias. 2008. Why the prosodic hierarchy is a diacritic and why the interface must be direct. *Sounds of silence: Empty elements in syntax and phonology*, ed. by Jutta M. Hartmann, Veronika Hegedüs and Henk van Riemsdijk, 145–192. Oxford: Elsevier.

Scheer, Tobias. 2011. *A guide to morphosyntax-phonology interface theories: How extra-phonological information is treated in phonology since Trubetzkoy's Grenzsignale*. Berlin: Mouton de Gruyter.

Schreuder, Maartje, Dicky Gilbers and Hugo Quené. 2009. Recursion in phonology. *Lingua* 119. 1243–1252.

Selkirk, Elisabeth O. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Reuland, Eric. 2009. Language: Symbolization and beyond. *The prehistory of language*, ed. by Rudolf Botha and Chris Knight, 201–224. Oxford: Oxford University Press.