86 Morpheme Structure Constraints

Geert Booij

1 Introduction

Morpheme structure constraints are constraints on the segmental make-up of the morphemes of a language. A textbook example of such a constraint is that *bnik* is an impossible morpheme of English, whereas *blik* is a possible English morpheme that happens not to exist. Hence, *bnik* is a systematic gap in the morpheme inventory of English, whereas *blik* is an accidental gap in this inventory. This can be taken to imply that there is a morpheme structure constraint that prevents English morphemes from beginning with a /b/ followed by a nasal consonant.

Halle (1959: 38) proposed to account for such distributional generalizations by means of morpheme structure rules, which define the class of possible morphemes of a language. Morpheme structure rules were conceived of as rules that fill in predictable specifications of the sound segments of a morpheme. For instance, in the case of English morphemes that begin with the consonant cluster *b*C, such as *brick*, it is predictable that the C must be a liquid, i.e. a non-nasal sonorant consonant. That is, the feature specifications [–nasal] and [+sonorant] of the second consonant of *brick* are predictable. They can therefore be omitted in the lexical phonological specification of the relevant morphemes. Morpheme structure rules fill in the blank cells of the lexical phonological matrix, and thus turn this underspecified matrix into a systematic phonological matrix, with all feature values of its segments specified. This is the underlying phonological form of a morpheme to which the phonological rules of a language apply. In sum, morpheme structure rules function as redundancy rules that specify predictable information, and at the same time they define the set of possible morphemes of a language.

Stanley (1967) proposed to replace Halle's notion "morpheme structure rule" by the notion "morpheme structure condition" (MSC). All morpheme structure conditions function as redundancy statements with respect to fully specified lexical phonological matrices, which form the input for the phonological rules (P-rules).

The notion "morpheme structure condition" as discussed above forms part of the theoretical machinery of classical generative phonology, but has been subject to debate. In this chapter this debate will be summarized. Before doing that, I will provide some examples of phonotactic properties of morphemes in §2. The problems raised by the concept of "morpheme structure condition" will then be discussed in §3–§5. These problems are the following:

i. The redundancy problem: is there any need for a specific set of morpheme structure conditions, or can they be made to follow from other types of phonological rules or constraints? (§3 and §4)
ii. The duplication problem: how can we avoid the same distributional generalization (for example the homorganicity of a nasal consonant and a following obstruent in consonant clusters) being expressed by both an MSC and a phonological rule (P-rule), and thus making the grammar unnecessarily complex? (§4)
iii. The status of MSCs: are they absolute constraints, or statistical tendencies only? (§5)

The chapter will conclude with some observations on the expressive functions of morphemes with specific phonotactic properties (§6) and a summary of our findings on the status of MSCs (§7).

2 Morpheme structure conditions

The unequal distribution of phonemes across words and morphemes was an important topic of research in structuralist phonology, because distributional facts were interpreted as signaling the presence or absence of grammatical boundaries, as in the work of Trubetzkoy (Wiese 2001). For instance, in German the phoneme /j/ only occurs at the beginning of lexical morphemes (as in *jagen* 'to hunt' and its derivative *verjagen* 'to chase away'). Hence, the presence of the /j/ is a positive signal of a left-edge morpheme boundary (van Wijk 1939: 125). Such facts show that the phonological and grammatical dimensions of linguistic structure are not completely autonomous, but are related in a systematic fashion (Jakobson 1949). The relation between the distribution of phonemes and grammatical units such as morphemes and words is therefore an aspect of the interface between phonology and morphology. Jakobson (1949) drew attention to the fact that different grammatical units may have different phonotactic properties. For instance, he observed that of the 23 Czech consonants, only eight are found in inflectional suffixes. Jakobson also mentioned that only the following consonants appear in the inflectional suffixes of English: /z d n ŋ/.

Dutch exhibits a number of such asymmetries between lexical morphemes on the one hand and derivational and inflectional suffixes on the other (Booij 1995). For Dutch, the following generalizations hold:

(1) a. Suffixes may consist of consonants only (/s/, /t/ or a combination thereof).
b. Suffixes may begin with the vowel /ə/.
c. Suffixes may have /ə/ as their only vowel.

Lexical morphemes of Dutch, on the other hand, do not have the phonotactic possibilities listed in (1) for suffixes, and require the presence of at least one full vowel (that is, a vowel that is not /ə/; see CHAPTER 26: SCHWA), and cannot be schwa-initial. Dutch prefixes cannot begin with a schwa either, but can have the schwa as their only vowel, as is the case for the Dutch prefix *be-* /bə/. Thus, we can sometimes tell from the phonological make-up of a morpheme whether it is a lexical morpheme or an affix.

A famous type of morpheme structure constraint is the restricted distribution of consonants in Semitic roots (see CHAPTER 108: SEMITIC TEMPLATES). Most Semitic roots are triliteral, that is, they contain three consonants, the consonantal skeleton. These skeletons are intercalated with vowels, and these vowel patterns are the exponents of grammatical information. Greenberg (1950) observed that the first two consonants of a Semitic CCC skeleton cannot be identical, whereas the last two can. Furthermore, homorganic consonants, i.e. consonants with the same place of articulation, are excluded, unless they are identical, even if they are the last two consonants. This is exemplified by the following distributional patterns in Arabic:

(2)	*m-m-d	
	m-d-d	'to stretch'
	f-r-r	'to flee'
	*b-m-C	
	*C-b-m	
	*g-k-C	
	ʃ-k-k	'to split'
	*ʃ-k-g	

Similar facts are reported for Modern Hebrew in Bar-Lev (1978: 321): "well-formed roots contain only consonants from different places of articulation." For instance, the following patterns can be observed in existing Modern Hebrew roots (cf. also Berent *et al.* 2002):

(3)	labial – velar – dental	*bagad*	'to betray'
	velar – labial – dental	*gibor*	'hero'
	dental – velar – labial	*dégem*	'model'
	labial – dental – velar	*mélex*	'king'

Crucially, these constraints apply to morphemes only; hence they are tautomorphemic constraints. As McCarthy (1986: 209) observes, there is no Arabic root *tatak*, with two identical consonants, but there are inflected forms of verbs like *ta-takallam* 'you converse' in which the first *t* belongs to a prefix, and hence does not belong to the same morpheme as the second *t*. Thus the prohibition on identical consonants is not violated.

McCarthy (1986) proposes to analyze this constraint as following from the Obligatory Contour Principle (OCP), which states that identical elements on the melodic tier of a morpheme are not permitted (see CHAPTER 14: AUTOSEGMENTS). If the OCP applies to the lexical representations of Arabic roots, and assuming that all autosegmental spreading in Arabic is rightwards, it is predicted that, of the following three structures, only the third one is well-formed (McCarthy 1986: 209):

```
(4)  a.  C  V  C  V  C      b.  C  V  C  V  C      c.  C  V  C  V  C
         |     |     |          |\_____/     |          |     |\____/
         s     m     m          s            m          s     m
```

In (4a), the OCP is violated on the melodic tier, whereas in (4b) leftward spreading has taken place. Hence, if the V is /a/, the root **sasam* is excluded, and only the root *samam* is well-formed. In order to exclude sequences of homorganic but not identical consonants as well, the OCP must be interpreted here as OCP-Place. That is, assuming that there is a separate tier for Place specifications of sounds, OCP-Place forbids adjacent identical specifications on the Place tier, and this excludes adjacent homorganic consonants (including identical ones, which by definition have the same specification on the Place tier).

Phonotactic properties of morphemes may also reveal that they belong to a particular stratum of the lexicon of that language, and may differentiate between native and borrowed morphemes (see CHAPTER 95: LOANWORD PHONOLOGY). Itô and Mester (1995: 819) mention examples of constraints from several languages that are specific for native morphemes of those languages. Japanese is interesting in that its morphemes can be divided into four subclasses: Yamato (native), Sino-Japanese, Foreign and Mimetic. Each subclass is characterized by a set of constraints, some of which are valid for more than one subclass. For instance, Sino-Japanese roots consist of one syllable only, and Lyman's Law (morphemes contain at most one voiced obstruent) holds for Yamato morphemes only (Itô and Mester 1995).

Dutch words that begin with *pn-* in the spelling, such as *pneumatisch* 'pneumatic' and *pneumonie* 'pneumonia', betray their non-native origin: *pn-* is a well-formed word-initial consonant cluster in Greek, but not in Dutch. This suggests that morphemes with *pn-* do not belong to the set of possible native morphemes of Dutch, and the constraint *[pn- partially characterizes the set of native Dutch morphemes. In English this constraint applies to all morphemes, and hence the combination *pn-* is realized as /n/.

The range of phonotactic patterns found in morphemes may be smaller than those in words. English morphemes, for instance, never end in a cluster of voiced obstruents (there are no morphemes like **lovd* or **dubd*), whereas such clusters do occur in complex words like past tense forms of verbs (as in *loved* /vd/, and *dubbed* /bd/). Dutch morphemes are subject to the constraint that voiced obstruent clusters only occur in complex words: morpheme-internally we only find clusters like /pt/ and /st/, but in complex words we find clusters like /bd/ and /zd/, as in the past tense forms *eb-de* /ɛbdə/ 'receded' and *raas-de* /raːzdə/ 'raged'. The only exceptions to this Dutch MSC are a few loanwords like *labda* /lɑbdaː/ 'lambda' and *budget* /[bœʤɛt/ 'budget'. Hence, the occurrence of voiced obstruent clusters morpheme-internally makes the relevant Dutch morphemes recognizable as loans (Zonneveld 1983).

As observed by Shibatani (1973), MSCs may have a different status from phonological rules or constraints, in that loanwords are not necessarily adapted to the MSCs of a borrowing language, whereas the application of the phonological constraints cannot be suppressed. Hence, the Dutch loan *labda* keeps its voiced obstruent cluster, and is not pronounced as *la*[pt]*a*. Dressler (1985: 219–245) also provides examples from various languages of distributional patterns that are characteristic of morphemes.

In sum, there are distributional constraints that are characteristic of morphemes, but not of words in general. The question is whether and how they have to be accounted for by a specific type of constraint, the MSC.

3 The redundancy problem

As we saw above, in classical generative phonology constraints on the segmental composition of (lexical) morphemes are interpreted as lexical redundancy rules or morpheme structure constraints (MSCs) (Halle 1959, 1964; Stanley 1967; Chomsky and Halle 1968). For instance, in many languages, nasal consonants in morpheme-internal clusters share their place of articulation with a following consonant. This generalization can be expressed by omitting the place of articulation of the nasals in the lexical representation of the relevant morphemes. A lexical redundancy rule will then fill in the proper value for the feature [place], and thus derive fully specified underlying phonological representations to which the phonological rules of a language apply. Thus, in the phonological component of the grammar the set of rules that express static phonotactic generalizations is ordered to apply before the set of phonological rules that account for alternations. The two sets of rules (MSCs and phonological rules) together are considered to express all the phonotactic regularities of a language (Postal 1968: 214; see also CHAPTER 7: FEATURE SPECIFICATION AND UNDERSPECIFICATION).

The role and importance of MSCs have been questioned for a number of reasons. In the first place, as pointed out by Hooper (1972), the role of the syllable as a domain of phonotactic generalizations cannot be ignored. The notion "syllable" does not play any formal role in the type of generative phonology codified in Chomsky and Halle (1968), but since then a wealth of evidence for the crucial role of the syllable (and larger prosodic units such as the foot and the prosodic word) in phonological analysis has been amassed (Nespor and Vogel 1986). Constraints on syllable structure are by definition constraints on how phonemes can combine into larger units. Hence, a lot of constraints on phoneme sequences are in fact syllable structure constraints (Hooper 1972). For example, the constraint that an English word cannot begin with a consonant cluster of the type nasal + obstruent follows from the universal principle of syllable structure that the sonority of consonants must decrease towards the edges of the syllable (see CHAPTER 49: SONORITY). Thus, sequences like **mpat* and **ntak* are impossible English morphemes. This means that morphemes must have a phonological composition that will lead to well-formed prosodic structures.

A second argument for the syllable as a phonotactic unit is that we cannot determine whether a particular segmental string is ill-formed without taking syllabification into account. For instance, the consonant sequence /bkm/ is always phonotactically ill-formed in English, because there is no possible division across two syllables that leads to a sequence of well-formed syllables. On the other hand, the consonant cluster /kn/, which does not appear word-initially in English words, can appear word-internally, as in *acne*, because this word can be syllabified as *ak.ne*, with two well-formed syllables (dots indicate syllable boundaries). That is, the following generalization holds:

(5) A (grammatical) word is phonotactically well-formed *iff* it can be parsed exhaustively into one or more well-formed prosodic constituents.

The class of well-formed syllables and higher prosodic constituents (foot, prosodic word) can be defined by the prosodification algorithm of each language, which is partially universal and partially language-specific. This algorithm groups the sounds of words into syllables, feet, and prosodic words (Rubach and Booij 1990). Alternatively, the grammar contains a set of ranked prosodic constraints that determine the optimal prosodification of a string of sounds, as in Optimality Theory (Kager 1999).

We might therefore claim that MSCs are superfluous because phonotactic restrictions on morphemes can be seen as the effects of phonological constraints on the output forms of words. For instance, English does not have a morpheme *abkmer*, since this morpheme cannot be prosodified exhaustively: the /k/ cannot be made part of the first or the second syllable. Similarly, the reason that a Dutch lexical morpheme requires the presence of at least one full vowel is that, otherwise, such a morpheme cannot yield a well-formed prosodic word. Dutch suffixes, on the other hand, can be vowelless or contain the vowel /ə/ only, as mentioned in (1), because they always combine with a lexical morpheme. That is, the phonotactic shape of Dutch suffixes has to do with their being dependent on a host morpheme to which they attach. We might therefore consider these suffix-specific phonotactic properties as something that need not be expressed separately, since it follows from the mapping of morphological structure onto prosodic structure.

Hooper (1972) offers a second argument against MSCs, that formulating phonotactic constraints with the morpheme as domain may also lead to spurious generalizations. For instance, Dutch lexical morphemes of Romance origin may end in obstruent clusters that are unpronounceable in isolation:

(6)	*castr-eer*	'castrate'
	celebr-eer	'celebrate'
	emigr-eer	'emigrate'
	penetr-eer	'penetrate'

One might conclude that Dutch morphemes can end in clusters of the type /Cr/, but this generalization does not reveal what is really at stake: those morphemes are only acceptable because they are bound roots, obligatorily followed by a vowel-initial suffix. Hence, these consonant clusters will form proper syllable onsets, as in *pe.ne.treer*. Similar observations can be found in Kenstowicz and Kisseberth (1977: 145) for Tunica, and they therefore concluded that in such cases it is the word rather than the morpheme that is the domain of phonotactic constraints. Nevertheless, the occurrence of such root-final consonant clusters is revealing in the sense that they betray the Romance origin of those roots: Germanic roots of Dutch never have this form because they can be used as words without further suffixation.

A third example of the role of prosody in the phonotactics of morphemes is that in many languages lexical morphemes are subject to prosodic minimality conditions. For instance, Dutch lexical morphemes are subject to the constraint that they consist of at least one heavy syllable (with either a long vowel or a short vowel followed by a consonant). That is, a lexical morpheme cannot consist of a light syllable only; bimoraicity is required. It is only in exclamations like *hè* /hɛ/ that the use of such light syllables with a short vowel is possible.

Prosodic conditions on morphemes create a problem for the classical MSCs: the syllable structure of a morpheme is not part of its lexical representations, but a derived property. Therefore, MSCs cannot refer to derived prosodic properties such as bimoraicity (McCarthy 1998). The only way to circumvent this problem is to phrase the constraint in terms of segment sequences: a lexical morpheme must contain either a long vowel, or a short vowel followed by at least once consonant. However, we then miss the generalization that it is a prosodic syllable weight condition that is involved. Once more, this suggests that the segmental composition of morphemes is governed by phonological output conditions. A similar problem occurs when we want to express the following generalization for Dutch: "In mono-morphemic forms we do not find sequences of schwa-headed syllables" (van Oostendorp 1995: 141). Again, this MSC refers to the derived property of syllable structure (cf. Downing 2006 for a cross-linguistic survey of prosodic minimality conditions).

In sum, we have to find a way in which prosodic constraints can account for at least part of the phonotactic constraints on morphemes.

3.1 Non-syllabic sequential constraints

Not all constraints on segmental sequencing can be reduced to syllable structure or prosodic minimality requirements. There are sequential constraints on consonant clusters that hold independently from the tautosyllabic or heterosyllabic status of these clusters. For instance, Yip (1991) proposes the following generalization for English (see also CHAPTER 12: CORONALS):

(7) *Consonant Cluster Condition*

In consonant clusters, consonants may have at most one other articulator feature than Coronal.

Thus, we find English clusters like /pt/ and /kt/ (*apt*, *act*), but not (tauto- or heterosyllabic) clusters like /kp/, /pk/, /km/, /mk/, /xm/, and /gm/ (loanwords like *drachma* and *stigma* are exceptions to this generalization). Note that the ill-formedness of such clusters does not follow from syllable structure constraints since they could be heterosyllabic. Yet they do not occur. If we come across such clusters in words (as in *zipcode* and *backpack)*, we can conclude that these words must be compounds, consisting of more than one lexical morpheme.

An example of a sequential constraint that holds both for tautosyllabic and heterosyllabic sound sequences, observed for English by Davis (1991), and for Dutch by Booij (1995: 46), is that in the sequence *s*CVC the two Cs should not be identical, unless they are coronal. Here are some Dutch examples with labial and coronal consonants (such sequences of velar consonants do not occur for independent reasons):

(8)

CVC			*s*CVC		
poep	/pup/	'shit'	**spoep*	/spup/	
mam	/mɑm/	'mother'	**smam*	/smɑm/	
toet	/tut/	'face'	*stoet*	/stut/	'procession'

This constraint is also valid for heterosyllabic sequences: they are not acceptable when followed by a vowel, as shown by forms like **spupo* and **smama*.

The point that not all phonotactic constraints can be reduced to syllable structure constraints is particularly clear for word-edge constraints, which are discussed in the next subsection.

3.2 Word edges

The difference between syllable structure constraints and sequential constraints is stressed by Kristoffersen (2000: 46–48), in relation to the distribution of consonants at word edges. In Norwegian, the cluster *tl-* is a proper syllable onset. It occurs word-internally in words like *Be.tlem* 'Bethlehem' and *A.tle* (proper name). Yet Norwegian words never begin with this cluster. Kristoffersen also observed that, although Norwegian words never begin with *pn-*, a cluster that does not violate the Sonority Hierarchy Constraint on syllable structure, Norwegians have no difficulty in pronouncing loan words like *pneumatisk* 'pneumatic'. These observations imply that /tl-/ and /pn-/ are proper syllable onsets in Norwegian, and that the non-occurrence of initial /tl-/ and /pn-/ is not due to a syllable structure constraint, but to a constraint that holds for the left edge of Norwegian root morphemes or prosodic words. A similar example from Dutch is that lexical morphemes do not begin with *pj-*, *tj-*, or *kj-*; however, the diminutive suffix allomorphs *-pje*, *-tje*, *-kje* begin with these clusters, and hence these clusters do appear in word-internal syllable onsets, as in *riem-pje* 'belt-DIM' with the prosodic structure $((\text{rim})_\sigma(\text{pjə})_\sigma)_\omega$ (ω = prosodic word, σ = syllable). Therefore, the non-occurrence of these clusters cannot be attributed to a syllable structure constraint. The word-initial sequences /pj- tj- kj-/ do occur in borrowed proper names for male persons, such as *Pjotr*, *Tjeerd*, *Kjeld*, and they do not cause pronunciation problems for speakers of Dutch.

The edges of words may have special phonotactic properties, since they may either impose more restrictions than what syllable well-formedness requires, or allow for extra consonants compared to what is possible in syllables in general. The Norwegian examples above (no *tl-* or *pn-* at the beginning of a word) are a case in point. Other examples of more restricted phonotactics at word edges can be found in Booij (1983): in Huichol, for example, words cannot end in a consonant but syllables can (source: Bell 1976).

In Polish, extra consonants may be added in word-initial position that violate the universal Sonority Sequencing constraint (Rubach and Booij 1990: 434; see also CHAPTER 109: POLISH SYLLABLE STRUCTURE):

(9) *rwać* 'tear'
rfza 'rust'
lgnąć 'stick'
mdły 'tasteless'
mnich 'monk'

In these words, a sonorant consonant is followed by a consonant of the same or higher degree of sonority, in violation of the Sonority Sequencing requirement that the sonority of consonants must increase towards the direction of the nucleus. The account that Rubach and Booij (1990) propose is that Polish prosodic words have an extra optional word-initial slot for an extrasyllabic consonant preceding the regular syllables, which is exempt from the requirements of the Sonority

Sequencing condition. This analysis implies that allowing for these marked consonant clusters is not to be seen as a property of lexical morphemes, but of the prosodic words that corresponds with such morphemes.

The special phonotactics of word edges is dealt with in Optimality Theory in the form of alignment constraints (McCarthy and Prince 1993). The basic idea of this approach, which makes crucial use of ranked output constraints in computing the phonetic form of words, is that there are alignment constraints that require the alignment of prosodic and grammatical boundaries. According to McCarthy and Prince (1993), the language Axininca Campa has word-initial onsetless syllables, whereas word-internally a vowel hiatus must always be filled by an epenthetic consonant. The relevant alignment constraint blocks the insertion of an epenthetic consonant in word-initial position. If epenthesis took place, there would be no alignment of the left edge of the prosodic word with the left edge of the (vowel-initial) morpheme. That is, the alignment constraint is ranked higher than the constraint that penalizes empty onsets. Note, however, that this analysis does not directly express that the left edges of Axininca Campa morphemes can begin with a vowel, even though syllables in this language normally begin with a consonant. The alignment mechanism allows for a difference in make-up between the edges of morphemes and syllables, but does not express it.

3.3 Phonotactic differences between simplex and complex words

As briefly mentioned at the end of §2, the range of phonotactic patterns in morphemes may be smaller than in complex words. Harris (1994) presents a number of observations on the phonotactic differences between simplex words and complex words in English. For instance, one will not find a heterosyllabic sequence /pt/, as in *laptop*, within a morpheme (except for loans like *helicopter*), even though a heterosyllabic cluster /pt/ would not violate the syllable structure constraints of English: a syllable can end in a /p/, and begin with a /t/. The same applies to the cluster /pw/: a proper name like *Sopwith*, which is historically a compound, is exceptional in this respect, and thus betrays its historical origin as a compound. Such opaque compounds tend to be adapted to the phonotactic patterns. The proper name *Greenwich* with the sequence /nw/ is now pronounced without the /w/, thus adapting to the phonotactic constraints for monomorphemic words (Harris 1994: 51).

The observation that certain consonant clusters only occur at morpheme boundaries is often used in linguistic analyses for assigning multi-morphemic status to words (see CHAPTER 46: POSITIONAL EFFECTS IN CONSONANT CLUSTERS). For instance, many words in the Amerindian language Athapaskan are considered to be compounds, even though the constituents do not occur as words by themselves, because they contain consonant clusters that are characteristic of morpheme boundaries (Rice 2009: 546).

Phonotactic differences between root morphemes and complex words have also been observed for vowel harmony (cf. §4 for a more detailed discussion of such facts for Turkish). The necessity of a separate morpheme structure condition on vowel combinations in roots is explicitly defended in the analysis of Hungarian vowel harmony in Vago (1976); see also CHAPTER 123: HUNGARIAN VOWEL HARMONY. Harvey and Baker (2005: 1459) observed that in the Australian language

Warlpiri, a language with vowel harmony with respect to the feature [round], the sequence [–round][+round] is not permitted for two consecutive vowels (with intermediate consonants) within roots, whereas the disharmonic sequence [+round] [–round] is. They account for this difference not by assuming an agreement constraint, but by proposing separate constraints for each type of disharmonic sequence. In addition, there is a constraint of root identity that requires the feature specifications for [round] to be preserved in the output. Thus they do not need to assume two rules of vowel harmony, a morpheme structure constraint and a phonological constraint that applies to complex words, and the duplication problem is avoided. Note, however, that this analysis requires reference to the root, a type of morpheme, as the domain of an identity constraint. That is, reference to morphemes in phonological constraints is still required.

Different phonotactics may also play a role in recognizing the lexical category of a word. In Dutch, there is a marked difference in phonological make-up between simplex nouns and simplex verbs. Verbs tend to consist of at most two syllables; if there is a second syllable, it will end in a schwa followed by a liquid. Nouns, on the other hand, allow for a larger variety of phonological structures, such as those consisting of three or more syllables, or ending in a full vowel. It appears that speakers of Dutch are able to categorize words as nouns or verbs on the basis of such phonotactic knowledge (Don and Erkelens 2006).

In sum, the distributional properties of segments within morphemes relate to the phonological rules or constraints of the relevant language, but not all morpheme-internal phonotactics can be reduced to these more general phonological regularities. In the words of Stanley (1967: 397): "The constraints holding within single morphemes are more restrictive than the constraints which characterize larger units."

4 The duplication problem

The problem that the assumption of both MSCs and P-rules seems to lead to unnecessary complications of the grammar was noted by Stanley (1967). For instance, Turkish has two general P-rules of vowel harmony that also predict the distribution of vowels within morphemes: all vowels agree in backness, and high vowels agree in roundness (see CHAPTER 118: TURKISH VOWEL HARMONY). As Zimmer (1969: 310) points out:

> The restrictions on vowel co-occurrence within almost all bases of Turkic origin are nearly the same as those just described for suffix vowels; thus for the "harmonic" part of the lexicon, there are two MSC's which replicate, to a great extent, the vowel-harmony rules that determine the selection of vowels in suffixes. There is, however, a large number of loanwords to which these vowel harmony MSC's do not apply, – e.g. /günah/ 'sin', /kalem/ 'pen', /sosis/ 'sausage', /viraž/ 'curve'.

In addition, there is an MSC that does not double as a P-rule, the Labial Consonant MSC (Zimmer 1969: 312):

(10) After /a/, a [+high] vowel agrees in labiality with a preceding [+labial] consonant.

An example of a morpheme that obeys MSC (10) is *karpuz* 'watermelon', in which the second vowel is round, even though the first vowel is non-round. This is an interesting MSC for the debate on the redundancy of MSCs, since it has no P-rule counterpart.

This duplication problem, already noted by Stanley (1967), is discussed by Anderson (1974: ch. 16). Anderson observes that many Turkish morphemes such as *kitap* 'book' are disharmonic, but do not block the application of vowel harmony rules once they have been suffixed. Anderson therefore concludes that we need both MSCs and P-rules for vowel harmony in Turkish, since they may be subject to different idiosyncrasies. In Anderson's view, the relation between MSCs and P-rules dealing with vowel harmony is a functional one, which need not and cannot be expressed formally by unifying them into one rule.

Shibatani (1973) proposes that such constraints should be considered to be both MSCs and Surface Phonetic Constraints (SPCs). A constraint can be marked as an MSC, a SPC, or both an MSC and a SPC. Clayton (1976) argues that constraints that hold for underlying forms of morphemes only are unmotivated, and do not reflect the speaker's knowledge of his/her language. Therefore, Clayton claims that Surface Phonetic constraints suffice.

The duplication problem is also considered by Kiparsky (1982: 167–170), in the framework of Lexical Phonology (see also Kaisse and Shaw 1985: 25; CHAPTER 94: LEXICAL PHONOLOGY AND THE LEXICAL SYNDROME). In this framework, phonological rules apply cyclically. Rules apply either in a structure-adding fashion or in a structure-changing one. Rules only apply in a structure-changing fashion in derived environments, i.e. in environments created by the previous application of a morphological or a phonological rule (see Booij 2000 for a survey of this theory). Kiparsky (1982) proposed that there are no Morpheme Structure Rules. The lexical representations of morphemes are underspecified; that is, predictable properties are omitted. On the first cycle, phonological rules specify these features; i.e. they fill in the blanks. If a word is complex, the same rule can apply in a structure-changing fashion on the next cycle, since the complex word is a derived environment. For instance, the Dutch rule that requires obstruents to have the same specification for the feature [voice] as an adjacent obstruent can be applied as a blank-filling rule to the underspecified feature matrix for /x/ in a word like *achter* /ɑxtər/ 'back', and it can be applied in a structure-changing fashion in a complex word like *as-bak* 'ash-tray' (underlying /ɑs-bɑk/; phonetic form [ɑzbɑk]). In English, the place of articulation of the nasal consonant in *damp* can be left unspecified, and filled in by the rule of Nasal Place assimilation, whereas the same rule will change the underlying coronal nasal /n/ into [m] in the derived word *compress* (underlyingly *con-press*).

Such an analysis can also deal with exceptions to MSCs. For instance, the Dutch word *imker* /ɪmkər/ 'bee-keeper', which is synchronically a simplex word, violates the constraint on homorganicity of nasal–obstruent clusters. In Kiparsky's proposal, this is no problem: the nasal consonant will be fully specified as being labial, and the rule that predicts the feature [velar] for a nasal followed by /k/ is blocked from applying because feature-changing applications of this rule are allowed in derived environments only.

In the case of Turkish vowel harmony, the same solution would apply. Disharmonic roots are fully specified, and therefore the P-rules of vowel harmony are blocked from applying to these morphemes, whereas they will apply in derived environments, to the vowels of the suffixes.

In short, in Kiparsky's proposal, the duplication problem is solved by abolishing the class of morpheme structure rules, and having P-rules apply in two different fashions. However, not all types of generalizations over the phonological shape of morphemes mentioned above can be expressed this way. This applies in particular to prosodic conditions on the shape of morphemes.

4.2 MSCs in Optimality Theory: Lexicon optimization and output–output faithfulness

Optimality Theory (OT) does not allow for constraints on the inputs of phonological evaluation. Output constraints are the only mechanisms for expressing phonotactic patterns. This idea of OT is referred to as the Richness of the Base hypothesis. For instance, there is no input constraint that forbids the morpheme **bnik* as a morpheme of English. The output constraints will penalize such a form, and evaluate this form in such a way that the optimal output form is not faithful to this form, but different, e.g. *blik*. Since forms such as *bnik* will never surface in English, it does not make sense to store an underlying form *bnik* for *blik*. This is the effect of lexicon optimization. Thus, the phonological output constraints of a language will be reflected by the input forms. This point of view is foreshadowed in Sommerstein (1974: 73), who argued that judgments about whether a sound sequence is a possible morpheme must be made on the basis of surface representations.

This idea is discussed in more detail in McCarthy (1998, 2002, 2005), and can be illustrated as follows. Suppose there is a language with the constraint that obstruents are voiceless at the end of a syllable, and with the suffix /-ən/ as plural ending for nouns, as in [hut] – [hutən] 'hat(s)'. Furthermore, this language has no alternations of the type [hut] – [hudən]. That is, morphemes that end in an obstruent will always end in a voiceless obstruent. Given the word [hut], the Richness of the Base hypothesis implies that we might assume the underlying form /hud/ for the singular form. The correct phonetic form [hut] will be computed anyway. However, in an optimal lexicon the underlying form to be chosen will be /hut/ because of lexicon optimization. This means that of the possible tableaux that select the right form, the most harmonic one will be selected, i.e. the one with the minimal number of violations of constraints. The assumption of the underlying form /hud/ will imply violation of the input–output (IO) faithfulness condition, unlike the underlying form /hut/. IO faithfulness requires the underlying form to be selected as the surface form, unless it is overruled by higher-ranked constraints. The optimal underlying form can be selected by comparison of tableaux, and the selection of the most harmonic one. Thus, lexicon optimization makes restrictions on input forms superfluous.

If there are no MSCs, the question arises of how to account for constraints that hold for morphemes only. One example of a distributional difference between morphemes and words concerns the distribution of nasals in Dutch. Within morphemes, nasal consonants are always homorganic with a following obstruent (with the exception of *imker*; cf. §4.1). Hence, we find *damp* /dɑmp/ 'damp', *tand* /tɑnt/ 'tooth' and *dank* /dɑŋk/ 'thanks', but no morphemes ending in */-mt -mk -ŋp -ŋt/. On the other hand, complex words such as the 3rd singular present forms of verbs always end in /-t/, preceded by all three types of nasals: *klim-t* 'climbs', *zon-t* 'sunbathes', *zing-t* /zɪŋt/ 'sings'. If we assume a markedness constraint NC (nasals are homorganic with a following consonant), this constraint must be blocked from changing a verb form like *klimt* into *klint*. McCarthy (1998) argued that this can be achieved by making use of output–output (OO) correspondence constraints. If we rank OO faithfulness conditions for the relation between a base word and its derivatives higher than the markedness constraint NC, a verb form like *klimt* cannot be changed to *klint*, because this would violate the requirement of correspondence of the stem of this inflected form with the verbal stem *klim*.

Nasal assimilation should not be blocked in all derived environments, however. In a prefixed word like *compress*, the prefix-final /n/ of *con-* does assimilate to the next /p/. This can be accounted for if we assume that the NC constraint is ranked higher than faithfulness to the underlying form of the prefix, /kɔn/. The more general observation is that affixes tend to adapt to roots rather than the other way around. Hence, in OT analyses it is often assumed that faithfulness constraints for affixes rank lower than those for roots (Alderete 2003). This implies that constraints have to be indexed for particular morphological categories such as root and affix. Therefore, we have to allow for reference to morphological domains in a system of phonological output constraints.

4.3 *Domains and strata*

Our conclusion so far is that, even when we do not allow for constraints on underlying forms of morphemes, it should be possible to index a phonological output constraint for a particular morphological domain such as the lexical morpheme. This will make it possible to specify distributional constraints that hold for lexical morphemes only. For instance, the Dutch constraint that lexical morphemes and prefixes cannot begin with /Cj/, whereas suffixes can, is expressible by indexing this constraint for the relevant morphological domains.

As observed in §3, particular phonotactic properties may only hold for certain strata of the lexicon. This is discussed in detail for Japanese by Itô and Mester (1995, 1999, 2001), who argue that "phonological generalizations can be covert by being lexically partial: they hold within a subdomain of the lexical space, but are violated in peripheral areas occupied e.g. by loanwords or onomatopeia" (Itô and Mester 2001: 274). For instance, in Japanese the palatalization constraint that changes /t/ into /ts/ before /i/ (this constraint also excludes tautomorphemic /ti/ sequences) does not affect loanwords like English *tea* and *party*. Therefore, Itô and Mester defend the idea of stratum-specific (ranking of) faithfulness constraints.

A Dutch example of a stratum-specific constraint was mentioned in §2: in native Dutch morphemes morpheme-internal obstruent clusters are always voiceless, but this constraint does not hold for non-native morphemes such as *labda* /lɑbdaː/ 'lambda', or the brand name *Mazda* /mɑzdaː/. Such non-native morphemes preserve their foreign pronunciation. The word *labda*, for instance, will not be changed to [lɑptaː], and Dutch speakers will recognize it as a loan due to this phonological property. Another example is the word-initial cluster *sk-*, which does not occur in Dutch native words, but only in loans from English, e.g. *scan* and *Skype*.

A similar distinction between native and non-native morphemes can be observed for languages with vowel harmony: non-native lexical morphemes may be disharmonic, and this is not changed through application of the vowel harmony constraints. For instance, the Hungarian noun *soför* 'chauffeur' is disharmonic (the first vowel is back, the second one is front), and remains so, even though it selects its suffix vowels in accordance with the frontness/backness vowel harmony constraint.

In sum, phonological constraints may have to be indexed for particular morphological categories or for lexical strata.

5 Absolute constraints or tendencies?

A final important point of debate concerning MSCs is whether they are absolute constraints or just statistical tendencies. Zimmer (1969) investigated the psychological reality of the Labial Consonant MSC of Turkish (10) mentioned above. Recall that this constraint holds for morphemes only, and is not supported by the two P-rules of vowel harmony. Zimmer made up lists of pairs of nonsense words, and asked subjects, Turkish students in California, to determine which word of such a pair sounds more like a word that might actually occur in Turkish. In the case of word pairs where the P-rules of vowel harmony played a role, the results were as expected, with a strong preference for the word in accordance with the vowel harmony constraints. For the pairs that involved the Labial Consonant MSC, on the other hand, there was hardly any difference in number between expected responses (the words in accordance with the MSC) and unexpected responses. Zimmer (1969: 320) therefore concluded that an MSC that is not supported by a P-rule might not be internalized by native speakers of Turkish.

5.1 OCP-Place

The psychological reality of OCP-Place, discussed in §2, which excludes identical adjacent place specifications, has been investigated for speakers of Jordanian Arabic (Frisch and Zawaydeh 2001; Frisch *et al.* 2004). It appears that "Jordanian Arabic speakers do recognize systematic gaps that are violations of OCP-Place as different from accidental gaps involving unrelated consonant pairs" (Frisch and Zawaydeh 2001: 99), even though there are violations of OCP-Place. Frisch *et al.* (2004) argue that OCP-Place is not an absolute, universal constraint. They consider the constraint as reflecting the generalizations that Arabic speakers make on the basis of their lexicon. OCP-Place is claimed to be a gradient constraint, since there are quite a number of words that violate it, but in different degrees. "Forms that violate the constraint to a lesser degree are more frequent than forms that violate the constraint to a greater degree" (Frisch *et al.* 2004: 182). Frisch *et al.* also point out that the co-occurrence of homorganic consonants that are non-adjacent (occurring in the first and third positions) is less restricted than the co-occurrence of adjacent homorganic consonants. In other words, the OCP-Place constraint is a gradient, but psychologically real constraint: "the native speaker knows an abstract but gradient OCP Place constraint ('Roots with repeated homorganic consonants are unusual') based on generalizations over the statistical patterns found in the lexicon" (Frisch *et al.* 2004: 216).

Frisch *et al.* also looked at the effect of OCP-Place on the borrowing of Italian verbs in Maltese, a variety of Arabic with many loans from Italian. The number of Italian verbs whose consonant pattern conforms to OCP-Place is significantly higher than that of the Italian verbs that violate OCP-Place (but these latter verbs may also be borrowed, and adapted to Maltese). This again supports the psychological reality of such a constraint, without it being categorical.

These findings suggest that OCP-Place is a gradient constraint that aims at avoidance of similarity: the more similar adjacent consonants are, the more they are avoided. Speakers are able to make phonotactic generalizations about lexical morphemes, but the corresponding constraints need not be categorical.

Statistical tendencies for the composition of various morphological categories such as the root and the stem have been observed by Wiese (2001): 94% of all German roots begin with a consonant, and 96% of all German roots end in a consonant. In OT, this can be expressed by alignment constraints that require the left and right edges of a root to coincide with the feature [+cons]. For those roots that violate the constraint, IO faithfulness will preserve the vowel at the edges. Note in particular that the tendency to have consonants at the end of roots does not follow from a syllable constraint, since the universally most unmarked syllable is the open syllable. Thus, this type of distribution may function as a boundary signal.

5.2 *Constraints on underlying forms*

Dutch exhibits intriguing constraints on sequences of vowels followed by fricatives. The basic generalization is that a vowel is short before /f s/, whereas it is long before /v z/. Let us call this the VZ constraint. Due to the effect of Final Devoicing, the constraint that obstruents are voiceless at the end of a syllable, this constraint can only be observed directly if the fricative is not morpheme-final. The following morphemes illustrate this constraint:

(11)	*short vowel*			*long vowel*			*excluded*
	effen	/ɛfən/	'even'	*even*	/eːvən/	'even'	/ɛvən, eːfən/
	dissel	/dɪsəl/	'pole'	*vezel*	/veːzəl/	'fiber'	/ɛzəl, eːsəl/

This constraint is violated by a few loanwords like *mazzel* /mɑzəl/ 'good luck' and *puzzel* /pœzəl/ 'puzzle', and by the native morpheme *oefen-* /uːfən/ 'to exercise'. This shows that this constraint is not an absolute condition on pronounceability, but a statistical generalization about morphemes.

This VZ constraint seems to apply to intervocalic sequences only, since we do find long vowels followed by /f s/ at the end of morphemes, as in the singular forms of the following nouns:

(12)	*graaf*	[ɣraːf]	'earl (SG)'	*grav-en*	[ɣraːvən]	(PL)
	kaas	[kaːs]	'cheese (SG)'	*kaz-en*	[kaːzən]	(PL)

However, we can interpret this constraint as also applying to morpheme-final sequences if we assume the constraint to hold for the underlying forms of morphemes. Morphemes like *graaf* and *kaas* end in a voiced fricative underlyingly, as shown by their plural forms, and hence the underlying forms of these morphemes are /ɣraːv/ and /kaːz/, respectively. There are a few exceptions, such as the non-native word *graaf* 'graph' with the plural form *graf-en* [ɣraːfən]. In the case of /s/ *vs.* /z/, the number of exceptions is much higher, since there are a number of verbs like *eis-en* [ɛisən] 'require (INF)' and *ruis-en* [rœysən] 'rustle (INF)', in which the diphthong, which counts as a long vowel, is followed by [s]. Speakers of Dutch may thus recognize a plural form as *grafen* as being non-native, whereas the phonetic form of its singular form *graaf* [ɣraːf] does not betray this stratal property.

If we allow for phonotactic constraints on underlying forms that cannot be observed from their corresponding surface forms, this enables us to make generalizations about the kind of alternations we may find in a language (Booij 1999). This topic is also broached in Ernestus and Baayen (2003). They raise the question to what extent the occurrence of alternations between morpheme-final voiceless and morpheme-final voiced obstruents in Dutch morphemes is predictable. There appear to be clear regularities. For instance, if a Dutch morpheme ends in a long vowel plus a labial stop, that stop is always an underlying /p/; that is, we do not find the alternation [p – b] for such morphemes. In the case of the fricatives discussed above, we saw that the length of the preceding vowel is a strong predictor of whether the final obstruent is underlyingly voiced or voiceless, although stronger for /f v/ than for /s z/. The question is whether language users possess this kind of knowledge. Ernestus and Baayen (2003) tested this by asking subjects to make past tense forms for nonsense words. If the underlying form of the verbal root morpheme ends in a voiceless obstruent, the past tense suffix *-te* /tə/ will be chosen, and *-de* /də/ otherwise. It appeared that language users do make use of the phonotactic tendencies involved: there is a strong correlation between the proportion of *-te*/*-de* choices for nonsense morphemes and the proportion for existing morphemes with a similar phonotactic make-up. Ernestus and Baayen (2003) therefore concluded that the speaker chooses an underlying representation for a nonsense morpheme that makes it resemble similar morphemes in the lexicon. As was the case for the Arabic roots discussed in §5.1, such phonotactic generalizations concerning morphemes may be statistical rather than absolute in nature. Moreover, they may pertain to underlying forms, where properties are present that may not be accessible in surface form.

This kind of knowledge about the type of alternations that occur in a language may also be formalized without restricting such MSCs to underlying forms. Consider the singular/plural pairs of nouns in Dutch with a stem-final obstruent such as *hoed* [hut] – *hoed-en* [hudən] 'hat (SG, PL)'. The voice specification of the stem-final obstruent of the morpheme *hoed* can only be determined on the basis of the plural form. The plural form is the most informative form of the paradigm (Albright 2005, 2008), and we may assume that it is stored in lexical memory. The relation between the two forms can be specified by a schema of the following type:

(13) $[x]_{\text{SG}} \leftrightarrow [x - \text{ən}]_{\text{PL}}$

(The symbol ↔ indicates the correlation between the two forms; x is a variable for a string of segments.) The plural form is the only reliable form for the computation of the underlying form, that is, the form on the basis of which new derived words and inflected forms can be computed. For instance, if we were to coin the adjective *hoed-ig* [hudəx] 'hat-like', the stem has to end in a /d/, since the phonetic form [hutəx] is wrong. That is, an underlying form is not necessarily a lexically stored representation, but may be computed when necessary for a morphological operation. In the case of Dutch verbs, we need to compute the underlying form of the verbal stem for the choice of the proper form of the past tense suffix (*-te* or *-de*).

Recall now the generalization for Dutch verbal stems that after a long vowel (VV) there is never a *p/b* alternation: if a singular form ends in [VVp], its plural form will never end in [VVbən]. This generalization also holds for nouns, and this can be expressed by the following subschema of (13) (*y* is a variable for segmental strings):

(14) [*y* VVp]$_{\text{SG}}$ ↔ [*y* VVp-ən]$_{\text{PL}}$

In the case of the Dutch *s/z* alternations discussed above, we might assume subschemas like the following for nouns:

(15) a. [*y* VVs]$_{\text{SG}}$ ↔ [*y* VVz-ən]$_{\text{PL}}$ (as in *kaas – kazen* 'cheese (SG, PL)')
b. [*y* Vs]$_{\text{SG}}$ ↔ [*y* Vs-ən]$_{\text{PL}}$ (as in *kas – kassen* 'greenhouse (SG, PL)')

If such schemas do not apply to all words, that is, if they are statistical generalizations only, they can be given a weight that indicates their probability.

The generalizations expressed in (15) apply almost exceptionlessly to nouns, and are confirmed by irregular pairs of singular and plural nouns with vowel length alternation. Vowel Lengthening is no longer a regular rule of Dutch, but an idiosyncratic alternation, a relic of Prokosch's Law that applied to Early Germanic, illustrated here for the noun *glas*:

(16) glas [ɣlɑs] 'glass (SG)' glaz-en [ɣlaːzən] 'glass (PL)'

The correspondence between the length of the vowel and the [voice] specification given in (15) is maintained in these irregular pairs by the combination of vowel length alternation and choice of obstruent: the forms *[ɣlɑzən] and *[ɣlaːsən] are both ill-formed.

In sum: whether there is an alternation between a voiced and a voiceless stem-final obstruent for Dutch lexical morphemes can only be determined with 100 percent certainty on the basis of inflected forms such as plurals. Yet, segmental composition of the lexical morpheme may give a clue, in some cases with almost 100 percent reliability. This type of knowledge may be modeled by constraints on the underlying forms of lexical morphemes, or by alternation schemas of the type proposed in (14) and (15).

6 The expressive value of phonotactics

A final phenomenon to be discussed is that particular sound sequences may have a specific semantic or pragmatic value. Sound symbolism is the usual term for such phenomena (see Hinton *et al.* 1994 for a number of detailed studies). In particular, there are phonaesthemes, recurring sounds or sound sequences, with a particular value. For instance, Marchand (1969: 397) argued that "/ɩ/ is suggestive of the subjective, emotionally small and is therefore frequent with diminutive and pet suffixes," and Bauer (1996) also found a cross-linguistic tendency for the use of this vowel in diminutive suffixes (see also Nichols 1971 for consonant sound symbolism in diminutives). According to Marchand (1969), the word-initial sequence *fl-* in words like *flick*, *flip*, *flap*, *flash* is expressive of brisk, quick movement, and Marchand provides many examples of such phonaesthemes. Such sound combinations are not to be considered as morphemes by themselves; yet they have a particular value. Hence, one may claim that phonotactic properties of morphemes can have an expressive role.

Japanese has a class of mimetic morphemes, which are sound-imitating or manner-symbolic roots. These morphemes have to be minimally bimoraic, and usually they appear in reduplicated or some other bipodic form (Mester and Itô 1989: 268):

(17)	*poko-poko*	'up and down movement'
	noro-noro	'slow movement'
	paku-paku	'munching'
	pata-pata	'palpitating'

In her study of the expressive value of lexical patterns, Klamer (2002) observed that the violation of general phonotactic constraints in specific classes of lexical items may have an expressive value. That is, a marked semantic value correlates with a marked phonotactic structure. An example from Dutch is the class of monosyllabic words of the type /lVl/, that is, words with the same consonant /l/ in onset and coda. Such words violate the phonotactic constraint or tendency of Dutch that the liquid consonants /l r/ in a syllable should be different. Words with this kind of phonotactics may have marked interpretations, as the following examples from Klamer (2002: 273) illustrate:

(18)	*lal* (v)	'jabber, babble, slur one's words'
	lel (n)	'earlobe, clout, whopper'
	lil (v)	'quiver'
	lol (n)	'fun, lark, trick'
	lul (n)	'prick, jerk'
	lul (v)	'talk nonsense'

Morphemes in which the vowel of /lVl/ is long do not occur at all.

In sum, the expressive value of phonotactic patterns within morphemes may be considered from a different angle: the violation of a constraint may have expressive value.

7 Conclusions

There is no doubt that there are distributional generalizations concerning the phonological make-up of morphemes that need to be expressed somehow in a proper phonological theory. The main theoretical issues are to what extent they can be made to follow from phonological generalizations that also hold for larger units than morphemes, and whether they are absolute constraints, or gradient constraints that express statistical tendencies. MSCs may also reveal different layers of the lexicon. Thus this chapter provides a range of data, observations, and considerations that can be used as a testing ground for the adequacy of theoretical phonological models.

ACKNOWLEDGMENTS

I would like to thank Paulo Chagas de Souza, Moira Yip, two anonymous reviewers, and the editors for their constructive comments and advice on an earlier draft of this chapter. In writing the final version, I also profited from taking part in a phonology seminar on phonotactic learning taught by Adam Albright at MIT and Harvard in the Spring semester of 2010.

REFERENCES

Albright, Adam. 2005. The morphological basis of paradigm leveling. In Downing *et al.* (2005), 17–43. Oxford: Oxford University Press.
Albright, Adam. 2008. Inflectional paradigms have bases too. Evidence from Yiddish. In Asaf Bachrach & Andrew Nevins (eds.) *Inflectional identity*, 271–312. Oxford: Oxford University Press.
Alderete, John. 2003. Structural disparities in Navajo word domains: A case for LexCat-Faithfulness. *The Linguistic Review* 20. 111–157.
Anderson, Stephen R. 1974. *The organization of phonology*. New York: Academic Press.
Bar-Lev, Zev. 1978. The Hebrew morphemes. *Lingua* 45. 319–331.
Bauer, Laurie. 1996. No phonetic iconicity in evaluative morphology. *Studia Linguistica* 50. 189–206.
Bell, Alan. 1976. The distributional syllable. In Alphonse Juilland (ed.) *Linguistic studies offered to Joseph Greenberg*, vol. 2, 249–262. Saratoga, CA: Anma Libri.
Berent, Iris, Gary F. Marcus, Joseph Shimron & Adamantios I. Gafos. 2002. The scope of linguistic generalizations: Evidence from Hebrew word formation. *Cognition* 83. 113–139.
Booij, Geert. 1983. Principles and parameters in prosodic phonology. *Linguistics* 21. 249–280.
Booij, Geert. 1995. *The phonology of Dutch*. Oxford: Clarendon Press.
Booij, Geert. 1999. Morpheme structure constraints and the phonotactics of Dutch. In Harry van der Hulst & Nancy Ritter (eds.) *The syllable: Views and facts*, 53–68. Berlin & New York: Mouton de Gruyter.
Booij, Geert. 2000. The phonology–morphology interface. In Lisa Cheng & Rint Sybesma (eds.) *The first Glot International state-of-the-art book: The latest in linguistics*, 287–306. Berlin & New York: Mouton de Gruyter.
Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.
Clayton, Mary. 1976. The redundancy of underlying morpheme structure conditions. *Language* 52. 295–313.
Davis, Stuart. 1991. Coronals and the phonotactics of nonadjacent consonants in English. In Paradis & Prunet (1991), 49–60.
Don, Jan & Marian Erkelens. 2006. Vorm en categorie. *Taal en Tongval* 19. 40–53.
Downing, Laura J. 2006. *Canonical forms in prosodic morphology*. Oxford: Oxford University Press.
Downing, Laura J., T. A. Hall & Renate Raffelsiefen (eds.) 2005. *Paradigms in phonological theory.* Oxford: Oxford University Press.

Dressler, Wolfgang U. 1985. *Morphonology: The dynamics of derivation*. Ann Arbor, MI: Karoma.
Ernestus, Mirjam & R. Harald Baayen. 2003. Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* 79. 5–38.
Frisch, Stefan A., Janet B. Pierrehumbert & Michael B. Broe. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22. 179–228.
Frisch, Stefan A. & Bushra A. Zawaydeh. 2001. The psychological reality of OCP Place in Arabic. *Language* 77. 91–106.
Greenberg, Joseph H. 1950. The patterning of root morphemes in Semitic. *Word* 6. 162–181.
Halle, Morris. 1959. *The sound pattern of Russian: A linguistic and acoustical investigation.* The Hague: Mouton.
Halle, Morris. 1964. On the basis of phonology. In Jerry A. Fodor & Jerrold J. Katz (eds.) *The structure of language: Readings in the philosophy of language*, 324–333. Englewood Cliffs, NJ: Prentice-Hall.
Harris, John. 1994. *English sound structure*. Oxford: Blackwell.
Harvey, Mark & Brett Baker. 2005. Vowel harmony, directionality and morpheme structure constraints in Warlpiri. *Lingua* 115. 1457–1474.
Hinton, Leanne, Johanna Nicols & John J. Ohala (eds.) 1994. *Sound symbolism.* Cambridge: Cambridge University Press.
Hooper, Joan B. 1972. The syllable in phonological theory. *Language* 48. 525–540.
Itô, Junko & Armin Mester. 1995. Japanese phonology. In John A. Goldsmith (ed.) *The handbook of phonological theory*, 817–838. Cambridge, MA & Oxford: Blackwell.
Itô, Junko & Armin Mester. 1999. The phonological lexicon. In Natsuko Tsujimura (ed.) *The handbook of Japanese linguistics*, 62–100. Malden, MA & Oxford: Blackwell.
Itô, Junko & Armin Mester. 2001. Covert generalizations in Optimality Theory: The role of stratal faithfulness constraints. *Studies in Phonetics, Phonology, and Morphology* 7. 273–299.
Jakobson, Roman. 1949. L'aspect phonologique et l'aspect grammatical du langage dans leurs interrelations. Reprinted 1963 in Roman Jakobson. *Essais de linguistique générale*, 161–175. Paris: Minuit.
Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
Kaisse, Ellen M. & Patricia A. Shaw. 1985. On the theory of Lexical Phonology. *Phonology Yearbook* 2. 1–30.
Kenstowicz, Michael & Charles W. Kisseberth. 1977. *Topics in phonological theory*. New York: Academic Press.
Kiparsky, Paul. 1982. From Cyclic Phonology to Lexical Phonology. In Harry van der Hulst & Norval Smith (eds.) *The structure of phonological representations*, part I, 131–175. Dordrecht: Foris.
Klamer, Marian. 2002. Semantically motivated lexical patterns: A study of Dutch and Kambera expressives. *Language* 78. 258–286.
Kristoffersen, Gjert. 2000. *The phonology of Norwegian*. Oxford: Oxford University Press.
Marchand, Hans. 1969. *The categories and types of present-day English word formation.* Munich: Beck.
McCarthy, John J. 1986. OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17. 207–263.
McCarthy, John J. 1998. Morpheme structure constraints and paradigm occultation. *Papers from the Annual Regional Meeting, Chicago Linguistic Society* 32(2). 123–150.
McCarthy, John J. 2002. *A thematic guide to Optimality Theory*. Oxford: Oxford University Press.
McCarthy, John J. 2005. Optimal paradigms. In Downing *et al.* (2005), 170–210.
McCarthy, John J. & Alan Prince. 1993. Generalized alignment. *Yearbook of Morphology 1993.* 79–153.
Mester, Armin & Junko Itô. 1989. Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* 65. 258–293.

Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology.* Dordrecht: Foris.

Nichols, Johanna. 1971. Diminutive consonant symbolism in western North America. *Language* 47. 826–848.

Oostendorp, Marc van. 1995. Vowel quality and phonological projection. Ph.D. dissertation, University of Tilburg.

Paradis, Carole & Jean-François Prunet (eds.) 1991. *The special status of coronals: Internal and external evidence*. San Diego: Academic Press.

Postal, Paul. 1968. *Aspects of phonological theory.* New York: Harper & Row.

Rice, Keren. 2009. Athapaskan: Slave. In Rochelle Lieber & Pavol Stekauer (eds.) *The Oxford handbook of compounding*, 542–573. Oxford: Oxford University Press.

Rubach, Jerzy & Geert Booij. 1990. Edge of constituent effects in Polish. *Natural Language and Linguistic Theory* 8. 427–463.

Shibatani, Masayoshi. 1973. The role of surface phonetic constraints in generative phonology. *Language* 49. 87–106.

Sommerstein, Alan H. 1974. On phonotactically motivated rules. *Journal of Linguistics* 10. 71–94.

Stanley, Richard S. 1967. Redundancy rules in phonology. *Language* 43. 393–436.

Vago, Robert M. 1976. Theoretical implications of Hungarian vowel harmony. *Linguistic Inquiry* 7. 243–263.

Wiese, Richard. 2001. The structure of the Germanic vocabulary: Edge marking of categories and functional considerations. *Linguistics* 39. 95–115.

Wijk, Nicolaas van. 1939. *Fonologie: Een hoofdstuk uit de structurele taalwetenschap.* The Hague: Martinus Nijhoff.

Yip, Moira. 1991. Coronals, consonant clusters, and the coda condition. In Paradis & Prunet (1991), 61–78.

Zimmer, Karl. 1969. Psychological correlates of some Turkish morpheme structure conditions. *Language* 45. 309–321.

Zonneveld, Wim. 1983. Lexical and phonological properties of Dutch voicing assimilation. In Marcel van den Broecke, Vincent van Heuven & Wim Zonneveld (eds.) *Sound structures: Studies for Antonie Cohen*, 297–312. Dordrecht: Foris.