Morphology and Phonotactics

Maria Gouskova

1 Summary

Phonotactics is the study of restrictions on possible sound sequences in a language.[1] In any language, some phonotactic constraints can be stated without reference to morphology, but many of the more nuanced phonotactic generalizations do make use of morphosyntactic and lexical information. At the most basic level, many languages mark edges of words in some phonological way. Different phonotactic constraints hold of sounds that belong to the same morpheme as opposed to sounds that are separated by a morpheme boundary. Different phonotactic constraints may apply to morphemes of different types (such as roots vs. affixes). There are also correlations between phonotactic shapes and following certain morphosyntactic and phonological rules, which may correlate to syntactic category, declension class, or etymological origins.

Approaches to the interaction between phonotactics and morphology address two questions: how to account for rules that are sensitive to morpheme boundaries and structure, and what is the status of phonotactic constraints that are associated with only some morphemes. Theories differ as to how much morphological information phonology is allowed to access. In some theories of phonology, any reference to the specific identities or subclasses of morphemes would exclude a rule from the domain of phonology proper. These rules are either part of the morphology or are not given the status of a rule at all. Other theories allow the phonological grammar to refer to detailed morphological and lexical information. Depending on the theory, phonotactic differences between morphemes may receive direct explanations or be seen as the residue of historical change

and not something that constitutes grammatical knowledge in the speaker's mind.

2 Keywords

phonotactics, morpheme structure constraints, markedness, faithfulness, sublexicon, derived environment effects, lexical phonology, lexical strata

3 Morphology and phonotactics: Empirical overview

Phonotactic generalizations can apply generally in a language or be specific to morphological contexts. An example of a generalization that is not morphologically sensitive is the statement that Samoan has only open (C)V syllables without consonant clusters; this is true of all Samoan speech. Similarly, in American English, regardless of morphological context, the flap [ɾ] occurs only between two syllabic sonorants, the second of which is unstressed. But many phonotactic generalizations require further qualifications, since they apply differentially to morphemes of different types and are often sensitive to word boundaries, which usually coincide with morpheme boundaries. As observed originally by Trubetzkoy (see §4.1), phonotactic generalizations are a clue to such boundaries—in other words, they serve a delimitative or demarcative function in speech.

3.1 Morpheme and word boundary effects

Phonotactic constraints often make special reference to morpheme boundaries (junctures). For example, in languages with word-internal voicing assimilation, sequences of voiceless and voiced obstruents may occur only at word boundaries. This is the case in Hungarian: word-internal sequences agree in voicing (e.g., [ka:d] 'tub'∼[ka:d-bɒn] 'in the tub', [ku:t] 'well'∼[ku:d-bɒn] 'in the well'). Sequences at word edges may disagree, so if [ka:d] is followed by [ku:t], [dk] will not necessarily assimilate.

Relatedly, phonotactic patterns may apply differently in underived words as opposed to morphologically complex sequences: for example, in English, nasals assimilate in place of articulation to the following stops inside morphemes, but at affix boundaries, assimilation may be optional or even impossible. Thus, morpheme-internal nasal sequences tend to be assimilated: "camp" [mp] and "bank" [ŋk], but assimilation does not apply in sequences that straddle the suffixal morpheme boundary in "time-d" [md] or "hang-ed" [ŋd], or across the prefix boundary in "un-canny" ([nk] possible) and "un-met" ([nm] possible).

In an extreme example of a boundary-sensitive phonotactic pattern, certain segments may occur only at morpheme boundaries: in Russian, the palatalized voiced fricative [ʒʒʲ] (Jones and Ward 1969:142) does not occur morpheme-internally but appears where one might expect a consonant cluster followed by a morpheme boundary (compare /doʐdʲ/ →[doʂtʲ] 'rain', /doʐdʲ-i / → [doʒʒʲ-i] 'rains', vs. [doʐd-lʲivɨj] 'rainy') Roots such as *[oʒʒʲip] do not exist. This kind of pattern is an example of a morphologically derived environment effect (see §4.6).

3.2 Domain-sensitive phonotactic generalizations

Related to boundary effects is the observation that many phonotactic generalizations hold only within a domain that is determined morphosyntactically. For example, in languages such as Russian or Thai, a word cannot end in a voiced stop. Here, "word" means phonological or prosodic word, which is coextensive with the morpheme in isolating languages (such as Thai) and with the extended projection of the verb, noun, adjective, or some other syntactic constituent in synthetic or agglutinating languages.

Morphosyntactic domains also normally delimit the application of rules such as stress assignment, vowel harmony, and consonant harmony. In Turkish, for example, stress assignment, vowel harmony, and syllabification operate over the same morphosyntactically defined domains: within stems but not across compounds or phonological words (Kabak and Vogel 2001). The aspiration dissimilation rule known as Grassmann's Law is also usually restricted to apply within morphosyntactic domains: in Ancient Greek and Sanskrit, roots cannot contain more than one aspirated stop (Steriade 1982). In Ofo, the co-occurrence restriction on aspirated stops holds within roots as a static constraint but is also enforced by a rule within the phonological word, e.g., [ˈoskʰa]'the crane' + [aˈfʰa] 'white' > 'the white or American egret' (*oskʰafʰa) (MacEachern 1997:39).

It is very common for phonotactic constraints to apply differently to words with one root versus compounds—words with more than one root. Thus, in Russian (Gouskova and Roon 2013) and Polish (Newlin-Łukowicz 2012), as well as many other languages, there is one stress per phonological word regardless of length unless that word is a compound, in which case a weaker, secondary stress appears on the left-hand stem of the compound.

(1) Russian: stress in single-root words vs. compounds

kristəl-iz-av-ˈa-t-sə	"to crystallize" (single root)
crystal-verbalizer-impf.-conjug.marker-infin.-intrans.	
kristˌal-ə-grˈaf-ij-ə	"crystallography" (compound)
crystal-compound.linker-graph-noun-nom.sg	

(2) Polish: stress in single-root words vs. compounds

banan-ɔv-ˈego	"banana-adj" (single root)
banana-adj-gen.masc.sg	
banan-ˌɔv-ɔ-arbuz-ˈɔv-ɨ	"banana-watermelon" (compound)
banana-adj-compound.linker-watermelon-adj-masc.nom.sg	

3.3 Phonotactic differences between morpheme and word types

It is common for phonotactic patterns to differ between roots and affixes (McCarthy and Prince 1994b, Beckman 1998, Urbanczyk 2006). Roots are often subject to a minimal size requirement, whereas affixes may not be restricted in size or may be restricted to a maximal size requirement (e.g. "no bigger than a syllable", McCarthy and Prince 1986, Crowhurst 1991, and others). Thus, in English, a root cannot be smaller than a heavy syllable (thus, [bɛt] 'bet' but not *[bɛ]), while affixes do not have to be syllabic at all (e.g., non-syllabic consonants in the past tense suffix [-d] and plural [-z] alongside the heavy syllable [-ɪŋ] '-ing'). Requirements on roots are distinct from requirements on words: English free function morphemes are at least a syllable but do not have to be heavy (e.g., the light syllable in [ðə] 'the' and the heavy in [wɪθ] 'with'), although sometimes they have a heavy allomorph when pronounced under focus or in isolation (Selkirk 1995b).

As documented extensively by Beckman (1998, ch. 4), roots often contain a different (usually larger) set of segments than affixes—e.g., in Quechua, roots may have ejective, aspirated, and plain stops (e.g., [p', pʰ, p]), but affixes may ony have plain stops ([p] but not [pʰ, p'] (Parker and Weber 1996, Gallagher 2016). Occasionally, the asymmetry between roots and affixes is in the other direction; Albright (2004) reports that Lakhota roots cannot contain codas but affixes and function words can.

Phonotactic patterns may differ for words of different syntactic categories (see Smith 2011 for a detailed overview). The most common type of category difference is between nouns and verbs: usually, nouns have more phonotactic variety than verbs do. Smith (1998) shows that several Japanese dialects allow pitch accent contrasts in nouns but have predictable stress in verbs. Phonotactic differences between syntactic categories often have a subtle and gradient character: thus, in English, stress is semi-predictable overall, but it is more predictable and restricted in verbs (where it tends to be final) than in nouns, where it tends to be non-final (see, among others, Chomsky and Halle 1968, Kelly 1992, Burzio 1994, Hayes 1995, Guion et al. 2003, Albright 2008). In a non-stress example, Shona has more vowel co-occurrence possibilities in nouns than in verbs; verbal affixes display vowel harmony (Fortune 1980, Hayes and Wilson 2008, Gouskova and Gallagher 2017).

Morphological distinctions such as declension class can align with phonotactic differences as well. For example, in Russian, noun stems of the 4th declension always end in a palatalized consonant or the retroflex fricatives [ʐ, ʂ] (e.g., [krof̠ʲ] 'blood', [nov-əst̠ʲ] 'news' *new-nominalizer*); noun stems of other declensions can end in either palatalized or velarized consonants, e.g., the masculine 2nd declension [ʐɨraf] 'giraffe', [gostʲ] 'guest' (see Timberlake 2004, Steriopolo 2008). Gender has been shown to correlate with phonotactics in many languages (Köpcke and Zubin

1984 and others; see LINK TO GENDER/NOUN CLASS ARTICLE). It is quite common also for morphemes of different etymological origins to follow different phonotactic patterns. Thus, in English, morpheme- and word-initial [sf] occurs only in Greek loanwords. In Tagalog, [f] occurs only in loanwords such as [fiesta] 'feast' (Zuraw 2006, Jurgec 2012); when affixes are attached to such words, [f] is replaced with [p] (as in [pam-piesta] 'feast-instr.').

3.4 Lexical variation in phonotactics

Divergent phonotactic patterns such as those discussed above can arise from historical change or through borrowing. When a phonological rule loses its productivity, it can result in a trend toward different distributions of segments and structures among morphemes that used to undergo a rule as opposed to morphemes that never did. In English, fricatives used to undergo productive intervocalic voicing, but some of the vowels were lost (for example, the final vowel in "believ<e>"). The once productive alternation turned into a statistical difference between nouns and verbs, with final voiced fricatives being more frequent in the verbs: "belief/believe", "half/halve", "shelf/shelve", "a hou[s]e/to hou[z]e". The difference is not a categorical one, as seen in the nouns "sieve", "gauze", and verbs such as "goof", "cough", and "lease".

Another diacronic source of phonotactic differences between morphosyntactic categories lies in how the categories are used. For example, if the position of the verb in the language is relatively fixed but the position of nouns is less rigid, this can result in prosodic differences between verbs and nouns. As a result of how sentence prosody interacts with word phonology and factors such as stress clash and lapse avoidance, verbs would have fewer stress possiblities than nouns (Kelly 1992, Smith 2011).

Borrowing from other languages can strengthen existing trends or introduce new morphologically sensitive phonotactic patterns. In the English example of fricative voicing, many of the verbs that do not follow the historical pattern are borrowed or derived from loanwords ("to increa[s]e"—Latinate, "to photogra[f]"—Greek, denominal). In other languages, borrowing can also be a source of phonotactic regularities: for example, in Japanese, the labial fricative [ɸ] is restricted in the native vocabulary so that it occurs only before [u], but it occurs freely before other vowels in loanwords (e.g., [ɸoroː] 'follow-up', Ito and Mester 1999, Moreton and Amano 1999). Since languages generally borrow nouns more freely than verbal stems, this kind of borrowing can over time become a source of noun-verb asymmetries.

When a language borrows words with atypical phonotactics, these words can either be integrated into the normal morphological system of the language or fail to follow native morphological patterns. In Japanese, for example, the class of declinable verbs is essentially a closed one, containing only verbs with native roots. For loanwords in a verbal context, Japanese uses periphrastic light verb constructions: e.g., [goruɸu] 'golf', [goruɸu-o suru] 'to play golf' (*golf-acc-do*) (Miyamoto 2000). In Russian, most loanword nouns are adapted into the morphological system in part based their phonological shape: e.g., [ərəngután] 'orangutan' is masculine because it looks like a masculine stem of class 2, ending in a velarized consonant; [garílə] 'gorilla' is borrowed as a feminine because it ends in the same vowel as nominative feminines of class 1. In the genitive singular, these nouns are [ərəngután-ə] and [garíl-ɨ], following different declension patterns. But there is a class of loanword nouns that are borrowed as indeclinable, so [kengurú] 'kangaroo', [men^{j}ú] 'menu', [parí] 'bet (from French)' remain the same in all case forms (Corbett 1982). Even though all three trigger neuter agreement, they do not combine with the case suffixes of neuters, because these stems are phonotactically atypical for that class.

4 Theoretical approaches to morphology-phonotactics interactions

4.1 Trubetzkoy's early work on the interaction of morphology and phonotactics

Grenzsignale (boundary signals). The earliest observations about the interaction between morphology and phonotactics are due to Trubetzkoy (1939). Trubetzkoy notes that languages may limit certain sound sequences so that they only occur at morpheme or word boundaries (his *positive Grenzsignale,* “positive boundary signals”) or never occur at morpheme/word boundaries (*negative Grenzsignale*, “negative boundary signals”). As an example of a positive word boundary signal, Trubetzkoy cites word-initial stress in Finnish. As an example of a negative word boundary signal, he gives Finnish [ŋ] and [d], which occur only word-internally.

Trubetzkoy distinguishes between individual segments that are positionally restricted and sequences of such segments; he terms the sequences “group signals”. An example of a sequence that occurs only at morpheme boundaries would be [kɯ] in Russian: [kɯzbam], /k-izb-am/ ‘to huts-dative’ must contain a morpheme boundary after [k] because [kɯ] never occurs morpheme-internally in the language. On the other hand, [kosam] is ambiguous between /k-os-am/ ‘to wasps-dative’ and /kos-am/ ‘braids-dative’, since the sequence [ko] is not a boundary signal of any kind. A non-segmental boundary marker that can nonetheless affect the qualities of multiple segments is fixed stress. The distribution of fixed stress is defined morphosyntactically, with respect to the edges of words, so it serves a delimitative function in languages that have it.

Trubetzkoy notes that boundary positions are often sites of expanded contrast possibilities—for example, laryngeal contrast options are greatest in word/morpheme-initial position in languages such as Thai. Conversely, boundaries can be sites of obligatory contrast neutralization; for example, in Thai, only plain stops are possible word/morpheme-finally. In Upper Sorbian, there is an alternation between [x] and [k͡x]; of these, [k͡x] is restricted to morpheme-initial position. Subsequent typological work on positional neutralization in Optimality Theory has largely confirmed Trubetzkoy’s observations (Beckman 1998, Lombardi 1999, Zoll 1998, et seq.). Relevant here is also work on morphonotactics, which identifies some generalizations about correlations between morphological type and phonotactic complexity (Dressler 1985, Dressler and Dziubalska-Kolaczyk 2006; see also §6 and LINK TO NATURAL MORPHOLOGY ARTICLE).

Differences between languages in the use of boundary signals. Another observation in Trubetzkoy’s Gründzuge der Phonologie is that languages vary in how much actual use they make of boundary signals: for example, Finnish marks boundaries of words well, but not boundaries of morphemes, whereas German marks both types of boundaries—although even in German, not every sequence will be unambiguously segmentable into words just on the basis of boundary signals (see section 5 for experimental and computational studies that confirm Trubetzkoy's observations). Somewhat more speculatively, Trubetzkoy suggests that phonotactic differences between languages may limit loanword borrowing. If a language has done a significant amount of borrowing despite phonotactic inconsistencies, the usefulness of its boundary signals diminishes in stylistically mixed contexts, where loanwords are mixed with native vocabulary.

Phonotactic generalizations about the shapes of morphemes. While Trubetzkoy devotes far more attention to the interaction between phonotactics and morphological boundaries, he also identifies another important and at the time unstudied area where morphology and phonotactics intersect: the phonological shapes of morphemes. He points out that in Russian, noun roots always end in a consonant, and that Semitic verbal roots are for the most part triconsonantal. This topic of phonological shapes of morphemes has been the subject of fruitful research, from Prosodic Morphology onward (see section 4.3).

4.2 The Sound Pattern of English

The Sound Pattern of English (SPE, Chomsky and Halle 1968) offers a complete theory of the interface between morphology and phonology. It accounts for boundary effects by reifying morpheme and word boundaries to the status of representational elements that are available to rules. It accounts for phonological differences between morphemes and morpheme classes with its diacritic theory of exceptionality. It furthermore can capture domain sensitivity of certain rules via cyclicity, the notion that rules apply in stages to successively larger morphosyntactic domains.

Boundary symbols. The SPE distinguishes between three kinds of boundaries: +, #, and ##. The # boundary is inserted automatically between morphemes whenever the plus is absent, and the ## boundary marks edges of words in an utterance. The + boundary is special—it is weaker than # and associated with morphemes that are lexically listed with it. Rules can refer to these boundaries as either obligatory or optional context, allowing the theory to capture a wide range of morphophonological interactions. Rules that obligatorily mention #, for example, serve to phonologically demarcate morphosyntactic boundaries—e.g., the Russian rule of word-final devoicing ([koz̲-u/kos̲] 'goat (acc sg.)/(gen. pl.)', cf. [kos̲-u/kos̲] 'braid (acc. sg.)/(gen.pl.)') would obligatorily apply in the context /___#. Rules that optionally mention boundaries would apply across the board to any string involving two segments, enforcing context-free phonotactic generalizations—for example, in Korean, any obstruent-nasal sequence becomes nasal-nasal, /p n/ →[m n] (Zsiga 2011), even across word boundaries, so the rule would mention (#) as an optional part of the context. Finally, rules that apply only at some morpheme boundaries would mention + but not #, and rules that apply except at morpheme boundaries would be stated over segmental contexts only, with no boundaries in the context. The difference between + and # can be demonstrated on English nasal assimilation: it applies morpheme-internally and across "weak" boundaries, so its context is [+nasal](+)___[-son]. The prefix "in-" in "incoherent" would be listed as /ɪn+/, whereas "un-" gets an automatic # and would fail to undergo nasal place assimilation.

The cycle. The cycle is a device for encoding the domain-sensitivity of phonotactic generalizations. The cycle allows morpheme boundaries to condition the application of certain rules, but the boundaries can be ignored by other, postcyclic rules because boundaries are eliminated once the cyclic rules have applied. A classic example of a cyclic rule is stress assignment—in many languages, stress tracks the morphosyntactic structure of words, and is reconfigured only minimally or not at all in sentences (Liberman and Prince 1977, Selkirk 1995a, Wagner 2010, [INSERT LINK HERE TO METRICAL STRUCTURE ARTICLE]). Rules such as vowel harmony can also be formulated cyclically, to capture the observation that harmony often applies inside words but not across words and not in compounds.

Morpheme structure constraints. Morpheme structure constraints were introduced by Halle (1959) and elaborated in the Sound Pattern of English. The SPE theory of phonology offers two kinds of explanation for phonotactic generalizations: first, certain surface sound sequences are absent because rules systematically eliminate them, and second, certain surface sequences are absent because they were never there in the underlying representations of morphemes. Morpheme structure constraints ensure this latter part, asserting that there are phonotactic generalizations over underlying representations (URs). In the SPE, the unit of lexical storage is the morpheme, so constraints on URs are constraints on morphemes.

For example, in the SPE treatment of English [ŋ], it is assumed to always be derived by rule from an underlying nasal-velar sequence. The UR posited for 'long' has /Ng/, with the place of the nasal unspecified, and nasal place assimilation along with non-coronal voiced stop deletion derive [ŋ]. This assumption explains the alternation between [g] and zero in "lo[ŋ]"/"lo[ŋg]er", and it also explains the restricted distribution of [ŋ] in English: unlike [m] and [n], it cannot start a morpheme or form a cluster with [s] (thus, #[sm], #[sn] are attested but *[sŋ] is not). There must be a constraint on English URs whereby nasals cannot be dorsal, [+nasal] →[+anterior]. Since [ŋ] does occur in a few morpheme-internal contexts, e.g., [hæŋɚ] 'hangar' and [dɪŋi] 'dinghy', Chomsky and Halle posit an abstract velar consonant /x/ following the nasal in the URs of these words, which deletes after conditioning place assimilation. In this analysis of English, there is a two-way discrepancy between the segmental inventory of URs and surface representations: /x/ occurs in underlying representations but not on the surface, and [ŋ] occurs on the surface but not in underlying representations. This illustrates the two mechanisms for deriving phonotactic generalizations in the SPE; the rule component systematically eliminates /x/, and morpheme structure constraints together with rules derive the restricted distribution of [ŋ].

Morpheme structure constraints play a role in SPE accounts of certain types of derived environment effects (see §3.1 and §4.6): if a particular segment or structure occurs only at boundaries, it is because it is banned from URs or morphemes by these constraints and is only created by rules in the phonological component. Another, related function of morpheme structure constraints is to characterize the contrastive inventory of the language, regardless of the inventory of surface sounds. Thus, in Russian, a language with obstruent voicing assimilation, the voicing contrast is balanced among stops (e.g., b/p, d/t, g/k), but the affricates [ʦ] and [ʧ] have voiced counterparts only as allophones. The voiced affricates do not occur in any environments where voiceless obstruents are found—they appear only before voiced obstruents, to satisfy the requirement that clusters of obstruents agree in voicing:

(3) Russian voicing contrasts and inventory: [ʣ, ʤ] as allophones of [ʦ, ʧ]

	__V	__C[+son]	__C[-voice]	__C[+voice,-son]
b, d, g	✓	✓	✗	✓
p, t, k	✓	✓	✓	✗
ʣ, ʤ	✗	✗	✗	✓
ʦ, ʧ	✓	✓	✓	✗

The only environment where [ʦ] and [ʧ] come into contact with voiced obstruents is at morpheme boundaries: /oteʦ-bɨ/ →[ateʣbɨ] 'father-irrealis', /m^{j}aʧ-ʐe/ → [m^{j}aʤʐə] 'ball-discourse particle'. There is just one loanword exception that cannot be analyzed as containing a morpheme boundary, [man^{j}ʤʐur] 'Manchu'; the rest are morphologically complex: [biʤ-bol] 'beachball' (cf. [pianer-bol] 'Pioneer ball'), /erʦ-gerʦog/ → [erʣ-gerʦək] 'archduke'. An SPE-style account of these facts would exclude the voiced affricates from URs, positing a morpheme structure constraint that requires all underlying affricates to be voiceless. A general rule of voicing assimilation would then introduce voiced counterparts to all eligible obstruents. This accounts for the disparity between what can occur inside morphemes as opposed to at morpheme boundaries.

Morpheme structure constraints explicitly divorce statements about morpheme phonotactics from alternations, which allows the theory to encode generalizations that hold statically of morphemes but not of morpheme sequences. An example of such a phonotactic generalization comes from Quechua. If a Quechua morpheme contains more than one dorsal, they must all be uvular (e.g., [q^{h}aɴqa] 'ravine') or all velar ([kuʎku] 'type of bird'), but no morpheme can contain a mix: *[qaŋka]. Mixed velar-uvular sequences are attested across morpheme boundaries, however: [k^{h}aska-rqa] 'he gnawed' (Wilson and Gallagher to appear). An SPE-style account of this pattern would posit a morpheme structure redundancy rule that enforces height agreement for all the dorsals in a morpheme, but no further rules would apply to enforce this generalization at the level of word phonology, after morphemes have been concatenated.

The lack of connection between alternations and phonotactics in the SPE has sometimes been criticized, because there are cases where rules derive the same generalizations as those that hold statically of morphemes (Clayton 1976, Kisseberth 1970, McCarthy 2002:68–91). For example, both Russian and Thai allow only one type of stop in word-final position. In Russian, there are morphophonological alternations between voiced and voiceless stops. In Thai, there are no alternations, just a static restriction that voiced stops are not allowed word-finally. In the SPE, these languages would require two separate mechanisms: Russian has a grammatical rule that applies in the phonological component, whereas Thai has a morpheme structure constraint in its lexicon. Kisseberth's (1970) original example is a similar situation in a single language, Yawelmani, where no morpheme appears to have an underlying three-consonant cluster, and when such clusters are created by morpheme concatenation, they are eliminated by various phonological rules. Kisseberth points out that an SPE account of this pattern involves several rules and a morpheme structure constraint that are functionally united and encode the same generalization, namely, that consonant clusters are not allowed. While the SPE allows to simplify formally similar rules, it has no way to unify functionally related ones, and morpheme structure constraints end up duplicating some of the generalizations that the phonological rules enforce. While this lack of connection between rules and permissible morpheme shapes may be a liability in cases like Yawelmani, it does make for a simpler analysis of patterns such as Quechua dorsal assimilation, where a generalization holds over morphemes but it is not supported by an alternation. In a theory such as OT, where all phonotactic generalizations are derived from surface constraints, an analysis of Quechua would require additional mechanisms (such as a stipulation that the constraint against dorsal co-occurrence is limited to the domain of a morpheme rather than the phonological word).

4.3 Prosodic Morphology

Prosodic Morphology studies morphological rules and patterns that either impose phonological regularities or depend on them. Within this theory's domain are such phenomena as reduplication, templatic truncation, infixation, phonological selectional restrictions/subcategorization, and phonotactic generalizations about morpheme shapes such as minimal and maximal phonological size requirements (Broselow and McCarthy 1983, McCarthy and Prince 1986, Mester 1990, Weeda 1992, McCarthy and Prince 1993b, 1995, Urbanczyk 1996 et seq.). Since these phenomena often implicate units such as syllables and metrical feet, the theory of Prosodic Morphology introduces a unifying hypothesis about all of these phenomena: the phonological shape or positioning of certain morphemes is specified in the grammars in terms of prosodic structure. This is an extension of proposals such as McCarthy (1979) and Marantz (1982), which hypothesize that the URs of Semitic templates and reduplicative morphemes are specified partially as consisting of consonants and vowels but lacking in segmental/featural content.

In many languages, reduplicative affixes copy a consonant-vowel sequence from the stem; e.g., Samoan [no-nofo] 'sit (pl.)', cf. [nofo] 'sit (sg.)'. In a Prosodic Morphology analysis, the underlying representation for the reduplicant would be an empty syllable, possibly specified as light, whose content is supplied by a special copying operation. In other languages, the reduplicant may be a heavy syllable, or two syllables, in which case the reduplicant is hypothesized to be a foot or a minimal phonological word. This same unit would be involved in truncation rules, such as English hypocoristic (nickname) formation: both "Pat" from "Patrick" and "Sue" from "Susan" are heavy syllables, which also happen to be metrical feet in English.

Under the strong version of Prosodic Morphology theory, the connection between phonological units available to the morphology and the structures used in the phonology proper is not accidental. For example, if a language uses iambic feet in its stress system, then it should also use iambic feet in its truncation, reduplication (if larger than a syllable), and so on. An even stronger claim is that certain properties of rules such as infixation, as well as templates, arise from universal phonological principles (McCarthy and Prince 1994b). For example, it is common for vowel-initial infixes to be positioned after the first consonant of the stem, as in Tagalog [sulat] 'write' ∼ [sumulat] 'write (agent focus)' (French 1988). Prosodic Morphology connects this positioning to the prohibition on syllables that end in codas. In an Optimality-Theoretic analysis, the NoCoda constraint is violable, so even if the language does not show evidence of coda avoidance (Tagalog does not), the constraint can still drive infixation or be part of the explanation for why

CV is such a common reduplication shape (Steriade 1988, McCarthy and Prince 1993a, 1994a). This version of Prosodic Morphology does not separate the theory of general language phonotactics from phonotactic generalizations about morphemes: any differences between language-general and context-specific generalizations arise from the details of a language's ranking of universal constraints such as faithfulness.

While the intuition behind the main proposal is appealing, its empirical foundation has been challenged. In an extensive survey, Downing (2006) demonstrates that there are many languages in which the assignment of stress would require an iambic analysis but reduplication or truncation use a trochaic template, and vice versa. These facts point to the conclusion that it is profitable to recruit phonological units such as syllables and feet in analyzing prosodic invariance in morpheme shape and morphotactics (see, e.g., McCarthy et al. 2012), but they do not need to be and indeed sometimes cannot be connected to the language's other phonological patterns.

4.4 Lexical Phonology and Morphology and Its Descendants

Lexical Phonology and Morphology is a theory of the morphology-phonology interface whereby morphological and phonological rules are interspersed. Some phonological rules apply before certain kinds of morpheme concatenation, whereas others apply repeatedly before and after concatenation. This allows the theory to explain the differences between morphemes with respect to stress shift and boundary effects, as well as the resistance of certain morphemes to otherwise general rules. Phonotactic differences between morphemes are therefore hypothesized to be due to relatively earlier or relatively later morphological attachment.

Table 1 sketches an example of a Lexical Phonology analysis of phonotactic differences between morphemes. English nasal place assimilation applies to certain prefixes such as [ɪm-/ɪn-/ɪŋ] (*imperfect, intolerant, incoherent*) and [əm-/ən-/əŋ-] (*embed, entomb, encode*), but not to *un-*, which generally either fails to assimilate or does so only in rapid, informal speech (thus, normally *u[nb]ox,* not *u[mb]ox*). Nasal place assimilation holds as a static phonotactic morpheme-internally, with a few exceptions, but notably also does not apply before past tense suffixes, as in [hɛm-d] 'hemmed' and [hæŋ-d] 'hanged'. A Lexical Phonology account would divide these affixes and the rules that attach them into two levels, Level I (early attachment, including *in-* and *en-*) and Level II (late attachment, including *un-* and *-d*). The rule of nasal place assimilation applies after the Level I prefixes *in-* and *en-* are attached, but it is absent at Level II, where *un-* and *-d* are attached. This captures the phonotactic tendency for morpheme-internal sequences to obey nasal place assimilation but allows for assimilation to not apply at certain morpheme boundaries.

Input	/pæNdə/	/pɚfəkt/	/bɑks/	/hɛm/
Level I affixation	—	ɪn-pɚfəkt	—	—
Nasal Place Assimilation	pændə	ɪm-pɚfəkt	—	—
Level II affixation		—	ʌn-bɑks	hɛm-d
Output	[pændə]	[ɪm-pɚfəkt]	[ʌn-bɑks]	[hɛm-d]
	"panda"	"imperfect"	"unbox"	"hemmed"

Table 1: A sketch of a Lexical Phonology account of English Nasal Place assimilation (following Mohanan 1982)

Lexical Phonology distinguishes between *lexical* and *postlexical* rules. **Lexical rules** are phonological rules that apply in the lexicon, interspersed with morphological derivations. Lexical rules encode phonotactic generalizations that are not fully general in the language—they either hold of subclasses of morphemes and words (e.g., compounds but not one-root content words) or have unprincipled exceptions. **Postlexical rules** apply indiscriminately after all morphological derivations are completed. From the point of view of general language phonotactics, postlexical rules enforce generalizations that have no exceptions in the language—for example, certain facts about English flapping (e.g., no flaps in the onsets of stressed syllables) follow from the postlexical nature of the rule that derives it. In at least some proposals, postlexical rules are associated with several distinctive characteristics. They are not structure-preserving, which means that they may derive segments and structures that may be absent from underlying representations. For example, in a classic SPE-style account of English flapping, the segment is absent from lexical representations and is always derived by rule. The same may be true of secondary stress in English, which has a predictable distribution except where it is derived from hypothesized primary stresses (Benua 1997, Pater 2000, and others). Postlexical rules may also be sensitive to speech rate; this would be the explanation for why "unbox" may undergo optional place assimilation in fast speech.

One of the challenges for Lexical Phonology and Morphology is aligning morphological facts with the requirements of phonological analyses. Mohanan (1982) observes that the morphological derivation of forms such as "ungrammaticality" requires that "-ity", a Level I suffix, attach after "un-", a Level II suffix. The phonological analysis of stress in this word requires that "un-" attach last. This and other bracketing paradoxes cause Mohanan to posit a device called The Loop, which considerably weakens the argument for morphological level ordering (see Pesetsky 1979, Benua 1997). One of the original reasons for Lexical Phonology's rise was dissatisfaction with the SPE's abstractness: extrinsic rule ordering, arbitrary distinctions between morpheme boundary types, diacritic exceptionality, and the cycle (see Kiparsky 1972, 1973c,a,b). Lexical Phonology does not entirely succeed in doing away with all of these devices. Counterarguments against the global architectural assumptions of Lexical Phonology point out that "lexical words" are not the right level of representation for either syntactic derivation or accounts of phonological exceptionality (Hayes 1990, Marantz 1997, Linzen et al. 2013). Some of these issues are inherited by Lexical Phonology's descendant, Stratal Optimality Theory (Rubach 2000, Inkelas and Zoll 2007, Bermúdez-Otero 2010; see Wolf 2008 for some discussion).

4.5 Lexically indexed constraints and sublexical phonotactics

Lexically indexed constraint theory is a version of Optimality Theory. According to lexically indexed constraint theory, a language's constraint hierarchy may include constraints that come in two or more versions: thus, a faithfulness constraint F can appear as generic $F_{Generic}$, violated by all unfaithful mappings, and a morpheme-specific $F_{Specific}$, violated only by those unfaithful mappings that involve phonological material from specific morphemes. Depending on the proposal, indexation can apply to faithfulness only (Ito and Mester, 1995, Fukazawa et al., 1998) or to markedness constraints as well (Pater 2000, Flack 2007, Gouskova 2007, Jurgec 2010). There are also versions of the theory that have at least some generic constraints (Pater 2008, Gouskova

2012), vs. versions where the constraint in question is split into two or more cloned constraints, all of which are morpheme-specific (Becker 2009). The limitation to faithfulness constraints is also espoused in proposals that extend indexation or similar mechanisms to account for stylistic variation (Coetzee and Pater 2011, Coetzee 2016, though cf. Linzen et al. 2013).

Here is an example of an indexed constraint analysis, applied to English nasal place assimilation. The basic distinction is between morphemes that undergo assimilation and morphemes that do not: most morphemes have assimilation internally (e.g., bu[mp], -me[nt]), a few affixes show alternations (for example, "em-bed" [mb] vs. "en-tomb" [nt], "in-tolerant" [nt] vs. "im-perfect" [mp]), but unassimilated sequences can occur both morpheme-internally (as in "Ca[nb]erra") and across certain morpheme boundaries, "un-conscious" [nk], "hang-ed" [ŋd]. An analysis in terms of specific and general constraints must commit to treating some of these cases as normal and others as exceptional; since most of the sequences appear to be assimilated, we can take it that the markedness constraint NasalPlaceAssimilation "assign a violation mark for every nasal that does not share the place of articulation of the following stop" outranks faithfulness to place of articulation, Ident-place. Words like "Canberra" can be treated as exceptions, as can "un-". Exceptions would be placed in a class bearing some diacritic L, and are then indexed to a special faithfulness constraint Ident-place-L, violated whenever a segment belonging to an L morpheme is unfaithful to its place of articulation: Ident-Place-L≫NasalPlaceAssimilation≫Ident-Place. Since NasalPlaceAssimilation is violated routinely across the past tense suffix boundary and in compounds ("hang-ed", "hemm-ed", "horn-blower") the analysis could introduce a constraint that forbids the sharing of place features across prosodic word boundaries (CrispEdge—Ito and Mester 1994 and others; assuming the past tense suffix is an appendix to the prosodic word).

example	agreement?	status	Representation	analysis
bump	yes	normal (root)	[bʌmp]	NPA≫Ident-Place
Canberra	no	exception (root)	[kænbəɹə$_L$]	Ident-Place-L≫NPA
imperfect	yes	normal (prefix)	[ɪm-pɚfəkt]	NPA≫Ident-Place
unconscious	no (optional)	exception (prefix)	[ʌn$_L$-kɑnʃ-əs]	Ident-Place-L≫NPA
hemmed	no (obligatory)	special structure, normal	[[hɛm]$_{Pwd}$d]$_{Pwd}$	CrispEdge(Pwd)≫NPA

Table 2: A summary of an indexed constraint analysis of English nasal place assimilation

Embedded in a sufficiently rich theory of variation, this analysis allows for the individual rankings to interact with other factors, such as speech rate and register—for example, the fact that "un-" optionally assimilates but "hemmed" never does could be captured by translating this analysis into a weighted constraint framework and giving CrispEdge a much higher weight than NPA, but Ident-Place-L would have a weight closer to that of NPA (see Coetzee and Pater 2011, Linzen et al. 2013).

The most similar alternative to this theory is Cophonology Theory, which is a version of Stratal OT that posits different rankings associated with different morphemes and grammatical constructions (Inkelas et al. 1997, Inkelas 1998, Inkelas and Zoll 2005, 2007). As Pater (2006) points out, the theories are notational variants of each other in many ways, but they differ in their approach to locality and in learnability. Pater proposes that an indexed constraint is violated when the offending structure is part of the morpheme in question: thus, in "un-kind", there are two nasal-stop sequences. Both are evaluated by NPA, but only the nasal in "un-" is evaluated by the indexed faithfulness constraint Ident-Place-L, since only that nasal is affiliated with the prefix. The ranking Ident-Place-L≫NPA≫Ident-Place would allow for nasal place assimilation to apply differently to the morpheme boundary and to the morpheme-internal sequence, without referring to boundaries directly. A cophonology account would have to assume a serial derivation that flips the rankings of NPA and Ident-Place: the roots have NPA≫Ident-Place and prefixes have Ident-Place≫NPA, as presumably do words (witness "he[m-d]"). But it is not clear how this account would explain why assimilation is an option for "un-" but not for "hem". One argument for Cophonology Theory is that it avoids a proliferation of constraint types, but this theoretical economy comes at the price of explanatory power.

4.6 Approaches to Derived Environment Effects

Various theories of the morphology-phonology interface offer accounts of morphologically derived environment effects (see Lexical Phonology and Morphology and Its Descendants, Lexically indexed constraints and sublexical phonotactics, The Sound Pattern of English). The SPE theory makes direct use of morpheme boundary symbols to capture such effects. Lexical Phonology and Morphology explains derived environment effects through stipulations on rule application such as the Strict Cycle Condition or its descendants; see Kiparsky (1985), Kenstowicz (1994, ch. 5). Lexical Indexation theory capitalizes on the observation that specific morphemes are usually involved; these morphemes either (a) resist phonotactic generalizations enforced elsewhere because of special faithfulness constraints or (b) obey special phonotactic restrictions not seen elsewhere in the language because of special markedness constraints (see especially Pater 2008, Mahanta 2012).

An alternative take on derived environment effects is to view them as a kind of feature affixation. Wolf (2007) proposes a rich theory of featural changes that accompany affixation that explains both why these changes do not appear to be phonotactically motivated and why they are usually affix-specific. Slavic First Velar Palatalization is a well-known example. Like other modern Slavic languages, Polish has a rule whereby dorsals become coronal fricatives: [stra̠x] 'fear' ∼[straʂ-t͡ɕ] 'frighten', [krɔ̠k] 'step' ∼ [krɔ̠t͡ʂ-ɛk] 'step (dim.)', [rɔɡ̠-u] 'horn (gen sg)' ∼ [rɔʐ-ɛk] 'horn (dim)' (Łubowicz 2002, Gussmann 2007, and others). Łubowicz (2002) points out that this is not phonotactically motivated in the modern language since underlying /d͡ʐ/ does not undergo palatalization ([brid͡ʐ-ɛk] 'bridge (dim.)', not *[briʐ-ɛk]), even though the mapping from /d͡ʐ/ to [ʐ] involves fewer feature changes than the mapping from /ɡ/ to [ʐ]. Gussmann (2007) collects extensive evidence that the rules are affix-specific: cf. [pɔ-mɔɡ̠-ɛ] 'I will help' and [pɔ-mɔʐ-ɛ] '(s)he will help' (cf. [po-mot͡ɕ] 'to help' and [mɔɡ-w-ɛm] 'I could', both containing the same root /mog-/). In a mutation analysis à la Wolf (2007), the difference between the two [-ɛ] suffixes in [pɔ-mɔɡ̠-ɛ]∼[pɔ-mɔʐ-ɛ] would be that the latter comes with floating features as part of its representation: /[+strident], [-anterior], -ɛ/. These features are required to dock onto a segment that isn't part of the affix itself, and vacuous docking onto /d͡ʐ/ would be ruled out by another, independently motivated constraint. In the case of [pɔ-mɔt͡ɕ], the affix would consist entirely of floating features, /pɔ-mɔɡ, [+strident, -continuant]/. Additional markedness constraints would rule out the mapping of [d͡ʐ] in the same system where /k/ maps to [t͡ɕ] This analysis largely divorces phonotactic generalizations about Polish phonology as a whole from the analysis of affix-induced featural mutation, which seems right in light of Gussmann's observations.

It can be difficult to work out the precise details of such mappings in a satisfying way, so some phonologists conclude that the mappings are not phonological at all but are rather instances of phonologically conditioned suppletion or are handled in a special morphophonological component that is separated from the phonology proper (Rubach and Booij 2001, Green 2007, Gussmann 2007). The suppletive analysis would list multiple allomorphs in the lexicon, /mɔg/ ∼/mɔʐ/ ∼/mɔt͡ɕ/ and associate them in the lexicon using some sort of a relatedness statement. The appropriate allomorphs can be selected morphologically by affixal context or in the phonology, based on markedness. This approach is sometimes seen as a radical departure and criticized because it misses phonological generalizations or predicts suppletion of roots to be conditioned by their outer phonological environment (Bobaljik 2000, Wolf 2013, Gouskova et al. 2015, Embick 2010, and others). In theories such as Distributed Morphology, morpheme-specific mutations are handled via Readjustment Rules, which are outside the phonological component of the grammar (Halle and Marantz 1993, Siddiqi 2009, Harley and Tubino Blanco 2012). This is an ongoing area of debate so it is too early to formulate any conclusions.

5 Experimental studies and computational models of phonotactics and morphology

A perennial problem in linguistics is distinguishing between generalizations that form a part of speakers' productive grammatical knowledge and generalizations that go unnoticed by speakers or are not productive.

Trubetzkoy's observation that phonotactic generalzations offer a cue to segmenting the speech stream have been tested experimentally in a variety of ways. The "boundary signals" are tested directly in the experiments of McQueen (1998), who finds that Dutch listeners are better at detecting real words embedded in a sequence of nonsense syllables when they are flanked by segments that cannot be part of the same syllable. Thus, in Dutch, a word can end in [lm] but not in [lv]; words can begin in [vr] but not in [mr]. Correspondingly, listeners detect [pil] 'pill' better in [pilvrem] than in [pilmrem]. This knowledge of boundary signals is modeled computationally in Adriaans and Kager (2010)'s StaGe. StaGe is a computational model of human learning of word segmentation that relies on the observation that sequences that are allowed morpheme- and word-internally will be relatively more frequent than sequences that arise only at word boundaries. One of the open questions is to what extent this principle can be used for segmentation in different languages (see, e.g., Daland and Pierrehumbert, 2011 on Russian, Daland and Zuraw 2013 on Korean, Kastner and Adriaans 2017 on Arabic).

There is experimental evidence that listeners use not only boundary signal clusters but also information supplied by rules sensitive to morphosyntactic domains. Finnish speakers use vowel harmony to locate word boundaries in nonce syllable sequences: since [puhymy] does not follow the rules of Finnish backness harmony, listeners are inclined to place a boundary between [pu] and [hy] (Suomi et al. 1997). There is also extensive evidence that segmentation also relies on stress in a variety of languages (Cutler et al. 1986, 1997).

Experimental evidence also demonstrates that people have detailed knowledge of the phonotactics of portions of the lexicon. Japanese is well-known for its stratified lexicon, which is characterized by morphological and phonological differences between morphemes. Compounding tends to occur only within strata; rules such as post-nasal voicing and voicing dissimilation are also stratum-specific. There are also detailed generalizations about structures and sound sequences that are allowed within each stratum (Ito and Mester 1995, Kawahara et al. 2002):

	(V)pV allowed?	Voiced geminates?	r^j, h^j found?	[aː] allowed?	NÇ allowed?
Yamato	No	No	No	No	No
Sino-Japanese	No	No	Yes	No	Yes
Foreign	Yes	Yes	No	Yes	Yes
Mimetic	Yes	No	No	No	Yes

Table 3: Phonotactic generalizations about Japanese lexical strata

Moreton and Amano (1999) test speakers' knowledge of these generalizations in an indirect way. They demonstrate that Japanese listeners know which phonotactic features of Yamato, Sino-Japanese, and loanword vocabulary can co-occur. As shown in the table, [aː] cannot occur in any strata other than Foreign, and [r^j, h^j] cannot occur in any strata other than Sino-Japanese. Moreton and Amano synthesize a continuum of [a] ∼[aː] and put these vowels in different stratal phonotactic contexts, e.g., [r^jotaː] (Sino-Japanese C1, foreign V2) [potaː] (neutral), [poɸaː] (foreign C2, foreign V2). They show that Japanese listeners are more likely to perceive the same indeterminate length vowel as shorter when it co-occurs with a consonant that is not allowed in the foreign stratum: thus, [r^jota(ː)] with a half-long final vowel is more likely to be perceived as short because of the [r^j], which cannot co-occur with a long [aː].

A different kind of evidence on the role of phonotactics in morphology is supplied by experimental studies of phonologically conditioned morphological alternations rules (Bethin 1992, Marcus et al. 1995, Albright and Hayes 2003, Ernestus and Baayen 2003, Becker and Gouskova 2016, Gouskova et al. 2015). Much of this work establishes that people know the productive, generalizable aspects of these rules and can apply this knowledge in an experimental setting to novel words. But the studies also demonstrate that speakers have detailed phonotactic knowledge about the types of words and morphemes that follow different patterns. The Sublexical Phonotactics model (Becker and Allen submitted, Gouskova and Becker 2013, Gouskova et al. 2015, Becker and Gouskova 2016) directly encodes this observation by positing that people have phonotactic grammars about subsets of the lexicon that follow a particular pattern (e.g., English nouns that pluralize with [-aɪ], such as "syllabus"). When asked to generalize the rule, they use the phonotactic grammars as "gatekeepers": in order to follow the [-aɪ] rule, the word has to be phonotactically well-formed within the sublexicon of [-aɪ] pluralizing nouns.

Computational models of morphophonological generalization such as the Minimal Generalization Learner (Albright and Hayes 2003) use rules based on segmental context: e.g., if an English verb ends in [ŋ], it is more likely to be irregular. But they are also enriched by phonotactic constraints on the products of affixation. In modeling English past tense formation, the Minimal Generalization Learner formulates (among others) a fully general rule that inserts [d] after any stem. This is wrong for stems that end in [d] and [t], such as "need", so a hybrid rule-constraint model generates several options and filters them out using constraints. This allows the model to capture the insight that the schwa in "needed" is inserted because of general phonotactic requirements of English.

6 Further Reading

- Beckman (1998) has a catalog of examples of phonotactic asymmetries between roots and affixes, or initial syllables of words vs. other syllables.
- Chomsky and Halle (1968) is the first complete theory of the morphology-phonology interface, with a detailed analysis of various aspects of English as well as discussion of other languages.
- Smith (2011) contains an overview of various phonological asymmetries between nouns, verbs, and adjectives.
- McCarthy and Prince (1993b) and Downing (2006) are book-length discussions of Prosodic Morphology.
- Mohanan (1982) introduces one version of Lexical Phonology and Morphology, with applications to English and Malayalam.
- Trubetzkoy (1939) is the first to pull together various observations about morphology and phonotactics interacting with each other.
- Wolf (2007, 2008, 2011) takes on analyses of various cases of the type discussed in this article, and presents an encyclopedic review of the literature on the morphology-phonology interface.
- Language Sciences 2014, Vol. 46: Special Issue on Theoretical and Empirical Approaches to Phonotactics and Morphonotactics highlights European work on morphonotactics (Korecky-

Kröll et al. 2014, Dziubalska-Kołaczyk 2014 and others).

Notes

[1] I would like to thank Suzy Ahn and two anonymous reviewers for comments on this article, which have resulted in numerous improvements. The work was supported in part by NSF BCS-1724753.

References

Adriaans, Frans, and René Kager. 2010. Adding generalization to statistical learning: The induction of phonotactics from continuous speech. *Journal of Memory and Language* 62:311–331.

Albright, Adam. 2004. The emergence of the marked: root-domain markedness in Lakhota. Boston. Handout of a talk presented at the 78th Meeting of the Linguistic Society of America.

Albright, Adam. 2008. How many grammars am I holding up? Discovering phonological differences between word classes. In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, ed. Charles B. Chang and Hannah J. Haynie, 1–20. Somerville, MA: Cascadilla Proceedings Project.

Albright, Adam, and Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90:119–161.

Becker, Michael. 2009. Phonological trends in the lexicon: The role of constraints. Doctoral Dissertation, University of Massachusetts, Amherst, MA.

Becker, Michael, and Blake Allen. submitted. Learning alternations from surface forms with sublexical phonology. *Phonology*, URL `http://ling.auf.net/lingbuzz/002503`.

Becker, Michael, and Maria Gouskova. 2016. Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry* 47:391–425.

Beckman, Jill. 1998. *Positional faithfulness*. New York: Routledge.

Benua, Laura. 1997. Transderivational Identity: Phonological Relations between Words. Doctoral Dissertation, University of Massachusetts, Amherst, Amherst, MA.

Bermúdez-Otero, Ricardo. 2010. Stratal Optimality Theory: an overview. http://www.bermudez-otero.com/Stratal_Optimality_Theory.htm. URL `http://www.bermudez-otero.com/Stratal_Optimality_Theory.htm`.

Bethin, Christina. 1992. *Polish Syllables: The role of prosody in phonology and morphology*. Columbus: Slavica Publishers.

Bobaljik, Jonathan. 2000. The ins and outs of contextual allomorphy. In *University of Maryland working papers in linguistics*, ed. K. K. Grohmann and Caro Struijke, volume 10, 35–71. College Park, MD: University of Maryland.

Broselow, Ellen, and John J. McCarthy. 1983. A theory of internal reduplication. *The Linguistic Review* 3:25–98.

Burzio, Luigi. 1994. *Principles of English Stress*. Cambridge: Cambridge University Press.

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

Clayton, Mary L. 1976. The redundance of underlying morpheme-structure conditions. *Language* 52:295–313.

Coetzee, Andries, and Joe Pater. 2011. The place of variation in phonological theory. In *The Handbook of Phonological Theory*, ed. John Goldsmith, Jason Riggle, and Alan Yu, 401–434. Oxford: Blackwell, 2nd edition edition.

Coetzee, Andries W. 2016. A comprehensive model of phonological variation: Grammatical and non-grammatical factors in variable nasal place assimilation. *Phonology* 33:211–246.

Corbett, Greville G. 1982. Gender in Russian: An account of gender specification and its relationship to declension. *Russian linguistics* 197–232.

Crowhurst, Megan. 1991. Minimality and Foot Structure in Metrical Phonology and Prosodic Morphology. Doctoral Dissertation, University of Arizona, Tucson.

Cutler, Anne, Delphine Dahan, and Wilma van Donselaar. 1997. Prosody in the comprehension of spoken language: a literature review. *Language and Speech* 2:141–201.

Cutler, Anne, Jacques Mehler, Dennis Norris, and Juan Segui. 1986. The syllable's differing role in the segmentation of French and English. *Journal of memory and language* 25:385–400.

Daland, Robert, and Janet B. Pierrehumbert. 2011. Learning Diphone-Based Segmentation. *Cognitive Science* 35:119–155.

Daland, Robert, and Kie Zuraw. 2013. Does Korean defeat phonotactic word segmentation? In *ACL (2)*, 873–877.

Downing, Laura. 2006. *Canonical forms in Prosodic Morphology*. Oxford: Oxford University Press.

Dressler, Wolfgang. 1985. On the predictiveness of natural morphology. *Journal of Linguistics* 21:321–337.

Dressler, Wolfgang U, and Katarzyna Dziubalska-Kolaczyk. 2006. Proposing Morphonotactics. *Italian Journal of Linguistics* 18:249.

Dziubalska-Kołaczyk, Katarzyna. 2014. Explaining phonotactics using NAD. *Language Sciences* 46:6–17. URL `http://www.sciencedirect.com/science/article/pii/S0388000114000448`, theoretical and empirical approaches to phonotactics and morphonotactics.

Embick, David. 2010. *Localism versus globalism in morphology and phonology*. Cambridge, MA: MIT Press.

Ernestus, Mirjam, and Harald Baayen. 2003. Predicting the Unpredictable: Interpreting Neutralized Segments in Dutch. *Language* 79:5–38.

Flack, Kathryn. 2007. Templatic morphology and indexed markedness constraints. *Linguistic Inquiry* 38:749–758.

Fortune, George. 1980. *Shona Grammatical Constructions Vol. I*. Harare: Mercury Press, 2nd edition edition.

French, Koleen. 1988. Insights Into Tagalog: Reduplication, Infixation, and Stress from Nonlinear Phonology. Doctoral Dissertation, University of Texas, Arlington. Published by SIL, Dallas and University of Texas at Arlington.

Fukazawa, Haruka, Mafuyu Kitahara, and Mitsuhiko Ota. 1998. Lexical stratification and ranking invariance in constraint-based grammars. In *Proceedings of the 32nd meeting of the Chicago Linguistic Society CLS, Part 2: The Panels*, ed. M. Catherine Gruber, Derrick Higgins, Kenneth Olson, and Tamra Wysocki, 47–62. Chicago, IL: Chicago Linguistic Society.

Gallagher, Gillian. 2016. Asymmetries in the representation of categorical phonotactics. *Language* 92:557–590.

Gouskova, Maria. 2007. The Reduplicative template in Tonkawa. *Phonology* 24:367–396.

Gouskova, Maria. 2012. Unexceptional segments. *Natural Language and Linguistic Theory* 30:79–133.

Gouskova, Maria, and Michael Becker. 2013. Nonce words show that Russian yer alternations are governed by the grammar. *Natural Language and Linguistic Theory* 31:735–765.

Gouskova, Maria, and Gillian Gallagher. 2017. Inducing nonlocal constraints from baseline phonotactics. Ms., New York University. Currently under review.

Gouskova, Maria, Sofya Kasyanenko, and Luiza Newlin-Łukowicz. 2015. Selectional restrictions as phonotactics over sublexicons. *Lingua* 167:41–81. URL `http://ling.auf.net/lingbuzz/002673`.

Gouskova, Maria, and Kevin Roon. 2013. Gradient clash, faithfulness, and sonority sequencing effects in Russian compound stress. *Laboratory Phonology* 4:383–434.

Green, Anthony Dubach. 2007. *Phonology limited*, volume 27 of *Linguistics in Potsdam*. Potsdam: Universitätsverlag Potsdam.

Guion, Susan G., J.J. Clark, Tetsuo Harada, and Ratree P. Wayland. 2003. Factors affecting stress placement for English nonwords include syllabic structure, lexical class, and stress patterns of phonologically similar words. *Language and Speech* 46:403–426.

Gussmann, Edmund. 2007. *The phonology of Polish*. Oxford University Press.

Halle, Morris. 1959. *The Sound Pattern of Russian*. The Hague: Mouton.

Halle, Morris, and Alec Marantz. 1993. Distributed Morphology. In *The View from Building 20. Essays in Honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser. Cambridge, MA: MIT Press.

Harley, Heidi, and Mercedes Tubino Blanco. 2012. Cycles,vocabulary Items and stem forms in Hiaki. Ms. University of Arizona.

Hayes, Bruce. 1990. Precompiled phrasal phonology. In *The Phonology-Syntax Connection*, ed. Sharon Inkelas and Draga Zec, 85–108. Chicago: University of Chicago Press.

Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: The University of Chicago Press.

Hayes, Bruce, and Colin Wilson. 2008. A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry* 39:379–440.

Inkelas, Sharon. 1998. The theoretical status of morphologically conditioned phonology: a case study of dominance effects. In *Yearbook of Morphology 1997*, 121–155. Dordrecht: Kluwer Academic Publishers.

Inkelas, Sharon, C. Orhan Orgun, and Cheryl Zoll. 1997. The implications of lexical exceptions for the nature of grammar. In *Derivations and Constraints in Phonology*, ed. Iggy Roca, 393–418. Oxford and New York: Oxford University Press.

Inkelas, Sharon, and Cheryl Zoll. 2005. *Reduplication: doubling in morphology*. Cambridge: Cambridge University Press.

Inkelas, Sharon, and Cheryl Zoll. 2007. Is Grammar Dependence Real? A Comparison between Cophonological and Indexed Constraint Approaches to Morphologically Conditioned Phonology. *Linguistics* 45:133–171.

Ito, Junko, and Armin Mester. 1994. Reflections on CodaCond and Alignment. In *Phonology at Santa Cruz*, ed. Jason Merchant, Jaye Padgett, and Rachel Walker, volume 3, 27–46. Santa Cruz: Linguistics Research Center, UC Santa Cruz.

Ito, Junko, and Armin Mester. 1995. The core-periphery structure of the lexicon and constraints on reranking. In *Papers in Optimality Theory II (University of Massachusetts occasional papers in linguistics)*, ed. Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 181–210. Amherst, MA: GLSA Publications.

Ito, Junko, and Armin Mester. 1999. The phonological lexicon. In *The Handbook of Japanese Linguistics*, ed. Natsuko Tsujimura, 62–100. Oxford: Blackwell.

Jones, Daniel, and Dennis Ward. 1969. *The phonetics of Russian.* Cambridge, MA: Cambridge University Press.

Jurgec, Peter. 2010. Disjunctive lexical stratification. *Linguistic Inquiry* 41:149–161.

Jurgec, Peter. 2012. Morphology affects loanword phonology. In *Proceedings of the 43rd meeting of the NorthEast Linguistic Society (NELS)*, ed. Ethan Poole Amanda Rysling Hsin-Lun Huang, volume I, 191–202. Amherst, MA: GLSA Publications.

Kabak, Baris, and Irene Vogel. 2001. The phonological word and stress assignment in Turkish. *Phonology* 18:315–360.

Kastner, Itamar, and Frans Adriaans. 2017. Linguistic Constraints on Statistical Word Segmentation: The Role of Consonants in Arabic and English. *Cognitive Science* .

Kawahara, Shigeto, Kohei Nishimura, and Hajime Ono. 2002. Unveiling the unmarkedness of Sino-Japanese. *Japanese/Korean Linguistics* 12:140–151.

Kelly, Michael H. 1992. Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological review* 99:349–364.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Oxford: Blackwell.

Kiparsky, Paul. 1972. Explanation in phonology. In *Goals of Linguistic Theory*, ed. Stanley Peters and Stanley Peters, 189–227. Englewood Cliffs, NJ: Prentice-Hall.

Kiparsky, Paul. 1973a. Abstractness, opacity and global rules. In *Three Dimensions of Linguistic Theory*, ed. O. Fujimura and O. Fujimura, 57–86. Tokyo: TEC.

Kiparsky, Paul. 1973b. 'Elsewhere' in phonology. In *A Festschrift for Morris Halle*, ed. Stephen R. Anderson and Paul Kiparsky, 93–106. New York: Holt, Rinehart and Winston.

Kiparsky, Paul. 1973c. How abstract is phonology? In *Three Dimensions of Linguistic Theory*, ed. O. Fujimura and O. Fujimura, 5–56. Tokyo: TEC. Reprinted in P. Kiparsky 1982, Explanation in Phonology, Dordrecht: Foris.

Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology* 2:85–138.

Kisseberth, Charles. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1:291–306.

Köpcke, Klaus-Michael, and David A Zubin. 1984. Sechs Prinzipien für die Genuszuweisung im Deutschen: Ein Beitrag zur natürlichen Klassifikation Six principes pour l'attribution du genre en allemand: contribution à la classification naturelle. *Linguistische Berichte* 26–50.

Korecky-Kröll, Katharina, Wolfgang U. Dressler, Eva Maria Freiberger, Eva Reinisch, Karlheinz Mörth, and Gary Libben. 2014. Morphonotactic and phonotactic processing in German-speaking adults. *Language Sciences* 46:48–58. URL `http://www.sciencedirect.com/science/article/pii/S0388000114000473`, theoretical and empirical approaches to phonotactics and morphonotactics.

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

Linzen, Tal, Sofya Kasyanenko, and Maria Gouskova. 2013. Lexical and phonological variation in Russian prepositions. *Phonology* 3:453–515.

Lombardi, Linda. 1999. Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language and Linguistic Theory* 17:267–302.

Łubowicz, Anna. 2002. Derived Environment Effects in Optimality Theory. *Lingua* 112:243–280.

MacEachern, Margaret. 1997. Laryngeal Cooccurrence Restrictions. Doctoral Dissertation, UCLA, Los Angeles.

Mahanta, Shakuntala. 2012. Locality in exceptions and derived environments in vowel harmony. *Natural Language & Linguistic Theory* 30:1109–1146.

Marantz, Alec. 1982. Re Reduplication. *Linguistic Inquiry* 13:483–545.

Marantz, Alec. 1997. No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In *University of Pennsylvania Working Papers in Linguistics: Proceedings of* *the 21st Annual Penn Colloquium*, ed. Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark, and Alexander Williams, volume 4.2, 201–225.

Marcus, Gary, Ursula Brinkmann, Harald Clahsen, Wiese Richard, and Steven Pinker. 1995. German inflection: the exception that proves the rule. *Cognitive Psychology* 29:189–256.

McCarthy, John J. 1979. Formal problems in Semitic phonology and morphology. Doctoral Dissertation, MIT. Published by Garland Press, New York, 1985.

McCarthy, John J. 2002. *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, John J, Wendell Kimper, and Kevin Mullin. 2012. Reduplication in harmonic serialism. *Morphology* 22:173–232.

McCarthy, John J., and Alan Prince. 1986. Prosodic Morphology. Technical Report #32, Rutgers University Center for Cognitive Science, version of 1996.

McCarthy, John J., and Alan Prince. 1993a. Generalized Alignment. In *Yearbook of Morphology*, ed. Geert Booij and Jaap van Marle, 79–153. Dordrecht: Kluwer.

McCarthy, John J., and Alan Prince. 1993b. *Prosodic morphology I: Constraint interaction and satisfaction*. New Brunswick: Rutgers University Center for Cognitive Science. Available as ROA-482 on the Rutgers Optimality Archive, http://roa.rutgers.edu.

McCarthy, John J., and Alan Prince. 1994a. The emergence of the unmarked: Optimality in prosodic morphology. In *Proceedings of the North East Linguistic Society 24*, ed. Merce Gonzalez, 333–379. Amherst, MA: GLSA Publications.

McCarthy, John J., and Alan Prince. 1994b. Two lectures on Prosodic Morphology (Utrecht, 1994). Part I: Template form in Prosodic Morphology. Part II: Faithfulness and reduplicative identity. Available as ROA-59 on the Rutgers Optimality Archive, http://roa.rutgers.edu.

McCarthy, John J., and Alan Prince. 1995. Faithfulness and Reduplicative Identity. In *University of Massachusetts Occasional Papers in Linguistics 18*, ed. Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 249–384. Amherst, MA: GLSA Publications.

McQueen, James M. 1998. Segmentation of continuous speech using phonotactics. *Journal of memory and language* 39:21–46.

Mester, Armin. 1990. Patterns of truncation. *Linguistic Inquiry* 21:475–485.

Miyamoto, Tadao. 2000. *The light verb construction in Japanese: The role of the verbal noun*, volume 29. John Benjamins Publishing.

Mohanan, K. P. 1982. Lexical Phonology. Doctoral Dissertation, MIT, Cambridge, MA. Distributed by IULC Publications.

Moreton, Elliott, and Shigeaki Amano. 1999. Phonotactics in the perception of Japanese vowel length: Evidence for long-distance dependencies. In *Proceedings of the 6th European Conference on Speech Communication and Technology*, volume 99, 82. Budapest.

Newlin-Łukowicz, Luiza. 2012. Polish stress: looking for phonetic evidence of a bidirectional system. *Phonology* 29:271–329.

Parker, Steve, and David Weber. 1996. Glottalized and aspirated stops in Cuzco Quechua. *International Journal of American Linguistics* 62:70–85.

Pater, Joe. 2000. Nonuniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology* 17:237–274.

Pater, Joe. 2006. The locus of exceptionality: Morpheme-specific phonology as constraint indexation. In *Papers in Optimality Theory III*, ed. Leah Bateman, Michael O'Keefe, Ehren Reilly, and Adam Werle, 259–296. Amherst, MA: GLSA.

Pater, Joe. 2008. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In *Phonological argumentation: Essays on Evidence and motivation*, ed. Steve Parker, 123–154. London: Equinox.

Pesetsky, David. 1979. Russian morphology and lexical theory. Ms., MIT. Available at `http://web.mit.edu/linguistics/people/faculty/pesetsky/russmorph.pdf`.

Rubach, Jerzy. 2000. Glide and glottal stop insertion in Slavic languages: A DOT analysis. *Linguistic Inquiry* 31:271–317.

Rubach, Jerzy, and Geert Booij. 2001. Allomorphy in Optimality Theory: Polish iotation. *Language* 77:26–60.

Selkirk, Elisabeth. 1995a. Sentence prosody: intonation, stress, and phrasing. In *The Handbook of Phonological Theory*, ed. John A. Goldsmith, 550–569. Cambridge, MA, and Oxford, UK: Blackwell.

Selkirk, Elisabeth. 1995b. The prosodic structure of function words. In *University of Massachusetts Occasional Papers: Papers in Optimality Theory*, ed. Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 439–470. Amherst, MA: GLSA Publications.

Siddiqi, Daniel. 2009. *Syntax within the word: Economy, allomorphy, and argument selection in distributed morphology*, volume 138. Amsterdam: John Benjamins Publishing Company.

Smith, Jennifer. 1998. Noun faithfulness: Evidence from accent in Japanese dialects. In *Japanese/Korean Linguistics, Volume 7*, ed. Noriko Akatsuka, Hajime Hoji, Shoichi Iwasaki, Sung-Ock Sohn, Susan Strauss, Noriko Akatsuka, Hajime Hoji, Shoichi Iwasaki, Sung-Ock Sohn, and Susan Strauss, 611–627. Stanford, CA: CSLI.

Smith, Jennifer L. 2011. Category-specific effects. In *The Blackwell Companion to Phonology*, ed. Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice, volume IV of *Blackwell Reference Online*. Wiley-Blackwell.

Steriade, Donca. 1982. Greek prosodies and the nature of syllabification. Doctoral Dissertation, MIT, Cambridge, MA.

Steriade, Donca. 1988. Reduplication and syllable transfer in Sanskrit and elsewhere. *Phonology* 5:73–155.

Steriopolo, Olga. 2008. Form and function of expressive morphology: a case study of Russian. Doctoral Dissertation, University of British Columbia, Vancouver, BC.

Suomi, K., J.M. McQueen, and A. Cutler. 1997. Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language* 36:422–444.

Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge: Cambridge University Press.

Trubetzkoy, N. S. 1939. *Grundzüge der Phonologie*. Prague: Travaux du cercle linguistique de Prague 7.

Urbanczyk, Suzanne. 1996. Patterns of Reduplication in Lushootseed. Doctoral Dissertation, University of Massachusetts, Amherst, MA.

Urbanczyk, Suzanne. 2006. Reduplicative Form and the Root-Affix Asymmetry. *Natural Language and Linguistic Theory* 24:179–240.

Wagner, Michael. 2010. Prosody and recursion in coordinate structures and beyond. *Natural Language & Linguistic Theory* 28:183–237.

Weeda, Donald. 1992. Word Truncation in Prosodic Morphology. Doctoral Dissertation, University of Texas, Austin.

Wilson, Colin, and Gillian Gallagher. to appear. Constraint complexity in surface-based phonotactics: A case study of South Bolivian Quechua. *Linguistic Inquiry* .

Wolf, Matthew. 2007. For an autosegmental theory of mutation. In *UMOP 32: Papers in Optimality Theory III*, ed. Leah Bateman, Michael O'Keefe, Ehren Reilly, and Adam Werle, volume 239–258. Amherst, MA: GLSA.

Wolf, Matthew. 2008. Optimal interleaving: Serial phonology-morphology interaction in a constraint-based model. Doctoral Dissertation, University of Massachusetts, Amherst, Amherst, MA.

Wolf, Matthew. 2011. Exceptionality. In *The Blackwell Companion to Phonology*, ed. Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice, volume IV, 2538–2559. Chichester, UK & Malden, MA: Wiley-Blackwell.

Wolf, Matthew. 2013. Candidate chains, unfaithful spellout, and outwards-looking phonologically-conditioned allomorphy. *Morphology* 23:145–178.

Zoll, Cheryl. 1998. Positional asymmetries and licensing. Ms., MIT. Available as ROA-282 on the Rutgers Optimality Archive, `http://roa.rutgers.edu`.

Zsiga, Elizabeth. 2011. External sandhi in a second language: The phonetics and phonology of obstruent nasalization in Korean-accented English. *Language* 87:289–345.

Zuraw, Kie. 2006. *Using the web as a phonological corpus: A case study from Tagalog*, 59–66.