3 The Interaction Between Morphology and Phonology

SHARON INKELAS

1 Introduction

The morphology of a language concerns the generalizations about form and meaning that relate words to one another within that language. The phonology of a language concerns the generalizations about the sound patterns in that language. Morphology and phonology intersect insofar as the statement of morphological generalizations includes information about sound patterns, or insofar as the statement of phonological generalizations includes information about morphology.

2 When Morphology Affects Phonology: The Phonological Interpretation of Morphologically Complex Words

The earliest influential generative approaches to the intimate interaction between phonology and morphology (Chomsky and Halle 1968; Kiparsky 1982b; Mohanan 1986) focused on the phonological interpretation of morphologically complex words, and this is where we will begin our survey as well, although we will not restrict ourselves to the phenomena covered by any particular theory in the process.

2.1 Morphologically Conditioned Phonology

Phonological requirements in a language can alter the shape that individual morphemes take in different contexts, producing allomorphy. Sometimes these

The Handbook of Phonological Theory, Second Edition. Edited by John Goldsmith, Jason Riggle, and Alan C. L. Yu

alternation patterns are quite general in the language. In Turkish, for example, a very general rule of progressive vowel harmony determines the value of [back] for the vowels of most suffixes, which surface with front vowels following roots whose final vowel is front (e.g. *gyl-ler* 'rose-PL', *anne-ler* 'mother-PL') but with back vowels following stems whose final vowel is back (e.g. *ok-lar* 'arrow-PL', *elma-lar* 'apple-PL') (see, for example, Lewis 1967). Morphologically conditioned phonology arises when phonological alternations are *not* fully general in the language but are instead specific to particular morphological constructions, such as compounding, truncation, affixation, or reduplication (for overviews at a fairly theory-neutral level, see for example, Dressler 1985; Spencer 1998).

In Belhare, for example, intervocalic voicing occurs at stem-suffix boundaries (*lap* > *lab-u!* 'catch it!') but not at prefix-stem boundaries (*ka-pira!* 'give it to me!'), or in underived words (*pipisi* '(drinking) straw') (Bickel and Nichols 2007). In Turkish, the diminutive suffix *-cIk* triggers the deletion of stem-final *k* (Lewis 1967: 57): *bebek, bebe-cik* 'baby/baby-DIM', *köpek, köpe-cik* 'dog/dog-DIM'). No other consonant-initial suffix triggers this deletion (*bebek-çi* 'child care provider', *bebek-lik* 'infancy', *bebek-ten* 'baby-ABL', *bebek-ken* 'while a baby', etc.). In Dakota, a coronal → velar dissimilation rule targets coronal consonant clusters that straddle the juncture between the two copies in reduplication (/žat/ → *žag-žát-a* 'curved'), but not clusters arising in other morphological contexts, for example, compounding (*sdod* + *čhí-ya* 'know.I + you-cause = I know you'; Shaw 1985: 184).

Morphologically conditioned phonology can be segmental, as in the examples just cited, or can involve prosodic properties such as tone, stress or length. Very familiar examples include Indo-European accentuation (Kiparsky 1973b) and Japanese (McCawley 1968b; Poser 1984; Pierrehumbert and Beckman 1988; Alderete 1999, 2001). In Japanese, morphological constructions, which include prefixation, suffixation, zero-derivation and compounding, come in two essential varieties: those which preserve lexical stem accent and those which erase it. Poser (1984) terms the two types "recessive" and "dominant," respectively, building on terminology introduced in Kiparsky 1973b (see also Kiparsky and Halle 1977; Halle and Mohanan 1985). Japanese pitch-accent is subject to strict distributional regularities: each word has at most one accent, and in cases of conflict between two lexically accented morphemes in the same word, the general principle is that the leftmost accent wins (Poser 1984). Recessive suffixes, as shown in (1), behave according to the Leftmost Wins principle. An unaccented suffix, for example, past tense *-ta*, leaves stem accent unaffected (1a), while an accented recessive suffix, for example, conditional *-tára*, surfaces with its accent only if the stem is not already lexically accented (1b). Otherwise, Leftmost Wins results in the elimination of suffix accent (1c). Page numbers are from Poser 1984:

(1)	a.	/yob-ta/	→	yoNda	'called'	(49)
		/yóm-ta/	→	yóNda	'read'	(49)
	b.	/yob-tára/	→	yoNdára	'if he calls'	(48)
	c.	/yóm-tára/	→	yóNdara	'if he reads'	(48)

Other recessive suffixes are pre-accenting, depositing accent on the final syllable of unaccented stems but having no effect on lexically accented stems (2a). Still others are accent-shifting. Poser terms these "dependent"; they shift stem-accent, if any, to the stem-final syllable, but do not have any effect on lexically unaccented stems (2b):

(2)	a.	matumoto-si	→	matumotó-si	'Mr Matsumoto'	(54)
		áNdoo-si	→	áNdoo-si	'Mr Ando'	(54)
		nisímura-si	→	nisímura-si	'Mr Nishimura'	(54)
	b.	koná-ya	→	konáya	'flour seller'	(55)
		kúzu-ya	→	kuzúya	'junk man'	(55)
		kabu-ya	→	kabuya	'stockbroker'	(55)

In contrast to recessive affixes, dominant affixes trigger deletion of stem-accent. Accented dominant suffixes, like adjective-forming *-ppó*, erase stem accent and surface themselves as accented (3a). Unaccented dominant affixes produce completely unaccented outputs, like demonymic *-kko* (3b). Still other dominant suffixes place accent on the initial or final stem syllable, as illustrated by (most forms with) the "true" prefix *ma(C)-* (3c) and family naming *-ke* suffix (3d), or even on the stem-penultimate syllable, as with the girls' name-forming *-ko* (3e):

(3)	a.	abura	→	abura-ppó-i	'oil, fat/oily'	(49)
		yásu	→	yasu-ppó-i	'cheap/cheap, tawdry'	(49)
		adá	→	ada-ppó-i	'charming/coquettish'	(49)
	b.	kóobe	→	koobe-kko	'an indigené of Kobe'	(72)
		nágoya	→	nagoya-kko	'an indigené of Nagoya'	(72)
		nyuuyóoku	→	nyuuyooku-kko	'an indigené of New York'	(72)
	c.	futatu	→	map-pútatu	'two/exactly half'	(57)
		sáityuu	→	mas-sáityuu	'amidst/in the very midst of'	(57)
		syoozíki	→	mas-syóoziki	'honesty/downright honest'	(57)
	d.	nisímura	→	nisimurá-ke	'the Nishimura family'	(55)
		ono	→	onó-ke	'the Ono family'	(55)
		hára	→	hará-ke	'the Hara family'	(55)
	e.	haná	→	hána-ko	'flower/name'	(58)
		kaede	→	kaéde-ko	'maple/name'	(59)
		mídori	→	midóri-ko ~ midorí-ko	'green/name'	(59)

Thus for each affix, or more generally for each morphological construction, since zero-derivation and compounding are subject to similar accentual parameters, it is necessary to know which of several possible accent placement patterns the affix triggers (none, stem-initial, stem-final, stem-penultimate) and whether those patterns preserve or delete lexical stem accent (dominant vs. recessive).

A more unusual case occurs in the Mayan language Mam (England 1983; Willard 2004), in which vowel length is contrastive both in roots and in suffixes

and at most one long vowel is permitted per word. Suffixes divide into two types: those that trigger shortening of stem vowels, and those that do not. Willard terms these "dominant" and "recessive" suffixes, respectively, following the terminology used in the accentual literature. Vowel length of the suffix itself is not a predictor of vowel shortening, as shown in the table; neither is stress nor morphological function. Whether or not suffixation causes stem vowel shortening is an idiosyncratic property of each suffixation construction. Dominant suffixes are shown in (4a); recessive suffixes are shown in (4b):

(4)	a.	mool-	'burn'	mol-oloon	'easily wilted' (FACILITATIVE)
		juus-	'burn'	jus-b'een	'burned place' (RESULTANT LOCATIVE)
		jaaw-	'go up'	jaw-nax	'up' (DIRECTIONAL)
		yuup-	'put out fire'	yup-na	'put out' (PARTICIPIAL)
	b.	iil-	'sin'	iil-a	'scold' (INTRANSITIVE VERBALIZER)
		ooq'-	'cry'	ooq'-b'il	'something which causes crying' (INSTRUMENTAL)

While cases of stress and tone replacement are more common than vowel length manipulation, on the basis of current knowledge it seems reasonable to assume that any kind of phonological pattern, other than the most low-level allophonic alternations, can be restricted to a morphological context, in some language or another. Indeed most phonetically "unnatural" phonological alternations (see e.g. Anderson 1981; Buckley 2000, Hyman 2001a) are morphologically conditioned in just this way, maintaining their niche of productivity in specific morphological contexts (see e.g. Pierrehumbert 2006b).

How is morphologically conditioned phonology to be handled? Current thinking, building on ideas going back to the 1960s, offers two main options: cophonologies, which are co-existing sub-grammars within a single language, each indexed to a particular morphological construction or set of constructions (e.g. Orgun 1996; Anttila 2002a; Itô and Mester 1995; Inkelas and Zoll 2005); or indexed constraints, in which the language has just one phonological grammar, but particular constraints within it are indexed to specific morphemes or morphological constituents (e.g. McCarthy and Prince 1995; Itô and Mester 1999; Smith 1999; Alderete 2001; Pater 2009).

To handle Japanese accentuation, for example, a cophonological approach would subdivide the grammar into a number of closely related variants, and index each morphological construction to one of these variants (cophonologies). The "dominant" morphological constructions would be associated with cophonologies in which input stem accent is eliminated. Poser 1984, using a rule-based precursor to cophonologies, proposed indexing an accent deletion rule to each dominant affix. In an Optimality Theory (Prince and Smolensky 2004; McCarthy 2008) implementation of cophonology theory, the same goal would be accomplished by varying the ranking of the constraints characterizing a particular accentuation pattern either below or above the faithfulness constraint preserving stem accent. For example, suppose stem-final and stem-initial accent are imposed by the

alignment constraints ALIGN-RIGHT(accent, stem) and ALIGN-LEFT(accent, stem), respectively. The cophonology of a dominant suffix would rank its accent-placing constraints above MAX-accent, ensuring the deletion of stem accents that are in the wrong location. The cophonology of a recessive affix would rank MAX-accent highest, ensuring that the relevant accentuation pattern is imposed only as a default.

All affix cophonologies in Japanese share the constraint ranking ensuring that there is at most one accent possible in the output. Generally, in cophonological models, the great majority of constraint rankings are shared by all cophonologies in the language; Anttila (2002a) has modeled this sharing using an inheritance hierarchy, in which cophonologies are grouped together by the constraint rankings that unite them. The superordinate node in such a hierarchy, or what Inkelas and Zoll (2005) term the "master ranking," represents the unique genius of the language, a partial ranking of constraints to which every individual cophonology must conform. As Anttila's work makes clear, it is also possible to group smaller subsets of cophonologies under intermediate nodes to capture subregularities, for example, overall differences between nouns and verbs in Japanese (as documented by McCawley 1968b and Poser 1984) or between nouns and adjectives in Finnish (Anttila 2002a).

Constraint indexation is a different, contemporary approach to morphologically conditioned phonology, developed in the early days of Optimality Theory. The approach was originally morpheme-based, indexing constraints to particular (sets of) morphemes. For instance, Itô and Mester (1999) account for the resistance of recently borrowed roots in Japanese to native phonotactic restrictions such as NO-P, the ban on [p], by indexing a special, high-ranked faithfulness constraint to exactly the set of relevant roots: $\text{FAITH}_{\text{AssimilatdForeign}}$ >> No-p >> $\text{FAITH}_{\text{Yamato}}$. A native root with underlying illicit /p/ would have to get rid of /p/, but an assimilated foreign root, as in *pato-kā* 'patrol car' (p. 63), would preserve it. Constraint indexation has also been applied to derived stems, nearly merging the difference between cophonologies (indexed to stem-forming constructions) and indexed constraint theory. In his analysis of morphologically conditioned accentuation in Japanese, for example, Alderete (2001) differentiates dominant and recessive affixes by indexing anti-faithfulness constraints to stems derived by the former. The constraint $\neg\text{OO}_{\text{Dom}}$-MAX-ACCENT ("It is not the case that every accent in S1 has a correspondent in S2") specifies that in derived stems created by dominant affixes, an input stem accent is not preserved in output. In Alderete's model, if a dominant affix causes input stem accent to delete, then the default accentuation pattern of the language is imposed in its place; it is also possible for alignment constraints to locate accent at the boundaries between stems and specific affixes.

With regard to the types of substantive differences that can exist between morpho-phonological patterns in the same language, the two approaches are very similar substantively, at least when implemented in Optimality Theory; each uses the same set of constraints and thus predicts the same range of possible

morphologically conditioned phonological effects. A more probative question is what degree of difference can exist across different morphologically conditioned patterns in the same language. Proponents of indexed constraints have suggested that the bulk of language-internal variation can be described in terms of relative faithfulness. Alderete (2001) has termed this "grammar dependence," claiming that each language has a single set of phonological restrictions (syllable structure, accentuation, segment inventory, etc.); individual morphological constructions differ only in the degree to which they are faithful to input structures which violate these restrictions. Thus, for example, in Japanese, stems created by recessive affixes are faithful to input accent, while stems created by dominant affixes are not faithful, and exhibit the default accentuation pattern of the language. The theoretical arguments on this point are subtle and complex; see for example, Itô and Mester (1999), Inkelas and Zoll (2007), Pater (2009). The empirical issue is whether any language ever imposes completely contradictory patterns in different morphological environments. Japanese accentuation is arguably a case of this kind, since even within the set of dominant affixes, at least four contradictory accentuation patterns are observed, as seen in (3).

A related substantive question about morphologically conditioned phonology is the number of variants (cophonologies, indexed constraints) a single language can permit, and the degree of differences among them. This question has been addressed explicitly in work by Anttila (2002a), whose hierarchical cophonological model predicts that every constraint ranking possibility not excluded in the "master ranking" of a language is expected to be instantiated in some cophonology. Neither cophonology theory nor indexed constraint theory addresses the question of how many different cophonologies are possible, or, really, to what degree they could potentially differ. As observed by Itô and Mester (1999), Inkelas and Zoll (2007), and Pater (2009), these issues may ultimately be laid at the feet of the historical origins of cophonological variation, which include language-internal factors like gramaticalization and analogy as well as external factors like lexical borrowing or more extreme language contact, as well as influences of language acquisition.

One way in which the cophonological and indexed constraint approaches clearly differ is in their ability to capture the interaction between different morphologically conditioned patterns in the same language, or "layering effects."

2.2 *Layering Effects*

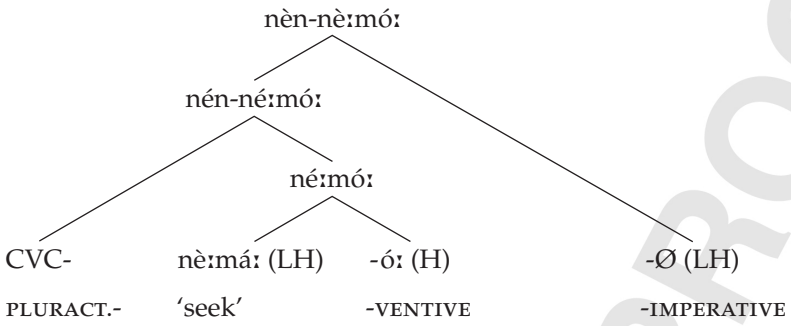
If two morphological constructions are present in the same word, and each is associated with its own phonological pattern, which pattern prevails, or if both do, how do they interact?

The evidence suggests that both patterns prevail, and that they are imposed in the order in which the associated morphological constructions are combined. This is perhaps easiest to illustrate using accentuation patterns that are incompatible,

such that when two morphological constructions affiliated with incompatible patterns co-occur in the same word, one must take precedence over the other. A good case study is Turkish, which resembles Japanese in some of its overall accentuation principles. The default position for stress in Turkish words is final (thus *arabá* 'car', *araba-lár* 'cars', *araba-lar-dán* 'from car's); there is exactly one stress per word, regardless of morphological complexity. A number of morphological constructions assign stress; these always override the default final stress pattern. (On Turkish stress, see e.g. Lewis 1967; Sezer 1981; Kabak and Vogel 2001; Inkelas and Orgun 2003.) In words with more than one stress-assigning morphological construction, order of morphological combination predicts the stress outcome. For example, Turkish has a productive zero-derivation construction forming place names out of words of any part of speech; the construction is marked by a distinctive stress pattern (Sezer 1981) which places stress on the penultimate or antepenultimate syllable, depending on syllable weight: *bak-acák* 'look-FUT' ~ *Bakácak* (place name), *torba-lí* 'bag-ASSOC' ~ *Tórbalı* (place name), , and so on. Turkish also has pre-stressing suffixes like past tense predicative -(y)DI (*torbá-ydı* 'it was a bag'), negative -mE (*gel-dí* 'came' vs. *gél-me-di* 'didn't come'), or mitigative -CE (*süt-lü-lér* 'milk-ASSOC-PL = the milky ones', vs. *süt-lű-ce* 'milk-ASSOC-MIT = kind of milky'). As documented in Inkelas 1999, Inkelas and Orgun 1998, Inkelas and Orgun 2003, the stress patterns of the language are all recessive in the sense that they are imposed only if the input stem lacks stress. In words like /torba-lI/, the stress outcome depends on whether an unstressed root, for example, /torba/ 'bag' is first converted to a stressed place name (*Tórba*) and then suffixed (→ *Tórba-lı*), retaining its place name stress, or first suffixed (*torbá-lı*) and then converted to a place name (*Torbálı*), retaining the stress assigned by the suffix instead of displaying the place name stress pattern. Like the Indo-European cases discussed by Kiparsky 1973b, Turkish respects a principle of "Innermost Wins" (Inkelas 1999; Inkelas and Orgun 2003).

Another useful illustration of layering can be found in Hausa, a lexical tone language whose morphological constructions either preserve stem tone (comparable to the "recessive" morphology of Japanese) or replace it with a new tone melody ("dominant") (Newman 1986; 2000; Inkelas 1998). The structure in (5) illustrates a verb root which combines with the dominant ventive suffix *-o:*, then undergoes pluractional reduplication, and is finally converted, via zero-derivation, to an imperative. Both the ventive and the imperative constructions are dominant. The ventive imposes an all-H melody (e.g. *fìtá:* (LH) 'go out' → *fít-ó:* (H) 'come out', *gángàrá:* (HLH) 'roll down' → *gángár-ó:* (H) 'roll down here', and so on (Newman 2000: 663). The imperative imposes a LH melody (e.g. *ká:mà:* (HL) → *kà:má:* (LH) 'catch!', *bíncìké:* (HLH) → *bìncìké:* (LH) 'investigate!'; *né:mó:* (H) → *nè:mó:* 'seek!', *nánné:mó:* (H) → *nànnè:mó:* (LH) 'seek repeatedly!'). In (5), the ventive occurs hierarchically inside the imperative. Predictably in Hausa, the outermost dominant construction is the one whose pattern surfaces; in this case the outermost construction is the imperative, and consequently the whole word surfaces LH. Zero-derivation constructions are represented by null suffixes for purely graphical convenience:

(5)



This kind of pattern is challenging for indexed constraint theory, in which all constraints, morphologically indexed and general, exist in one fixed ranking in the grammar of the language. In Turkish, the constraints that require place names to have the Sezer stress pattern must rank either below or above the constraints requiring stress to immediately precede suffixes like /-lI/. In a word containing both a zero-derived place name and a pre-stressing suffix, the higher-ranked pattern should always prevail, regardless of morphological structure. The problem is that both types of embedding can occur in Turkish, with different meanings and different stress outcomes corresponding to the two possible hierarchical structures (Inkelas and Orgun 1998). A single ranking, as in indexed constraint theory, can capture one but not the other, missing the connection between morphological embedding and constraint ranking. By contrast, in cophonology theories this connection is captured intrinsically (see e.g. Inkelas 1993; Orgun 1996 on "deriving cyclicity"); the hierarchical relationship between two constructions directly determines the input-output relationship between the associated cophonologies.

Some layering theories have bundled layering with additional claims, and have been weakened insofar as the additional claims have not held up. For example, the theory of Lexical Morphology and Phonology (Kiparsky 1982; Mohanan 1986) associated cyclicity (layering) with structure preservation and strict level ordering, to which subsequent literature has raised compelling empirical objections. Stratal Optimality Theory (Kiparsky 2000, 2008) limits the number of cophonologies (layer types) in any given language to three, which are strictly ordered. The virtue of limiting strata in this way is that it draws attention to general properties of stems, words, and phrases, but often at the expense of being able to describe more "minor rules." Close studies of strata in agglutinating languages, for example, have generally resulted in the postulation of more than three levels below the word level alone (see e.g. Hargus (1985) on Sekani, Mohanan (1986) on Malayalam, Buckley (1994) on Kashaya). Both Hargus and Mohanan, like Czaykowska-Higgins (1993, for Moses-Columbian Salish) and Inkelas and Orgun (1998, for Turkish) argue in addition that the strata necessitated to describe the morphophonological subgeneralizations in the languages in question cannot be crucially ordered in the way that level ordering theory would require. It is important, however, to emphasize that the essence of level ordering theories is the same as the essence of cophonology theory, namely the interleaving of

phonology and morphology is due to the association of morphological constructions with particular phonological patterns.

2.3 *Paradigm Uniformity*

A promising avenue of research on "optimal paradigms" seeks to examine whether paradigm-level considerations could motivate or even supplant cyclic cophonological models. This is especially promising in cases of recessive phonological alternations in which stem structure is preserved under subsequent affixation. The overall result is that paradigms are kept level, that is, with phonologically uniform stem shape. It has been proposed that rather than resulting from cyclicity, stem uniformity effects follow from paradigm uniformity constraints which keep the shared portions of morphologically related words phonologically identical; see for example, the Base-Identity constraints of for example, Kenstowicz 1996. When evaluated only with respect to the subconstituents of a single word, Base-Identity constraints function like high-ranked input-output faithfulness on a cophonology account, causing structure that is optimal for the innermost morphological constituent to persist even if outer layers of morphology render it phonologically opaque. This occurs in Turkish, as discussed earlier: lexically stressed roots (e.g. *lokánta* 'restaurant') and derived stressed stems (e.g. *süt-lǘ-ce* 'milk- ASSOC-MIT') keep their stress when they combine with would-be stress-assigning suffixes like pre-stressing predicative *-(i)di*, for example, *lokánta-ydı*, *süt-lǘ-ce-ydi*. The recessive character of stress-assigning suffixes can be attributed to paradigm uniformity: the derivational and inflectional paradigms of a lexically stressed noun like *lokánta* 'restaurant' all share an identically stressed root (*lokánta, lokánta-lar* (-PL), *lokánta-da* (-LOC), *lokánta-lar-da* (-PL-LOC), *lokánta-ydı* (-PRED), and so on.).

Of course, Base-Identity is not absolute in Turkish; it is only stressed roots whose phonological stress pattern is maintained across the paradigm. Lexically stressless roots alternate systematically, according to whether they combine with a stress-neutral suffix, for example, *araba-yá* 'car-DAT', or a stress-assigning suffix, for example, *arabá-yla* 'by/with car'. It is also important to note that the definition of "base" of a paradigm must be broadened to include not just roots but also complex stems. While the root *araba* is not inherently stressed, and therefore varies in shape depending on morphological context, a stressed stem like *arabá-yla* keeps its stress when suffixed, for example, *arabá-yla-m*$_1$ 'car-ASSOC-INTERROGATIVE = by/with car?'. Thus "base" is equivalent to "subconstituent" in a layering theory.

The predictions of paradigm constraints diverge from the predictions of cophonological layering models when applied to the shared stems of words neither of which is a subconstituent of the other. For example, Kenstowicz (2005) discusses the case of Spanish diminutives, formed by adding *-cito* [sito] (m.) /*-cita* [s*i*ta] (f.) when the base ends in [n] or [r] and by adding *-ito/-ita* when the base ends in a vowel.[1] Examples cited by Kenstowicz, using his phonemic transcription, include [limon] 'lemon (m.)' → [limon-sito], [barko] 'ship (m.)' → [bark-ito], [korona]

'crown (f.)' → [koron-ita]. For nouns that have feminine and masculine gender counterparts, like [raton] 'mouse (m.)', [raton-a] 'mouse (f.)', the surface conditions for attachment of the [-sita/-sito] diminutive formatives are met by the *n*-final masculine but not by the *a*-final feminine. On Kenstowicz's assumption that the form of the non-diminutive noun determines the dimunitive suffix that is added, the feminine diminutive of 'mouse' should be [raton-ita], based on [ratona], whereas the masculine diminutive of 'mouse' should be [raton-sito], based on [ratón]. In fact, however, both diminutives have the diminutive formative triggered by an *n*-final input: [ratonsito], [ratonsita]. Kenstowicz proposes a paradigm uniformity analysis, which he attributes to Aguero-Batista, on which masculine and feminine diminutives are required to have the same surface stem shape. Masculine [ratón] transparently selects [-sito] ([raton-sito]), and by paradigm uniformity, the feminine [ratona] is required to select the [-sita] allomorph as well. Paradigm uniformity favors [ratonsita], while transparency of suffix selection favors [ratonita]; paradigm uniformity wins out. (There is an alternative to invoking paradigm uniformity in this case, namely treating gender-unspecified [ratón] as the input both to [raton-sit-o] and [raton-sit-a]. The argument for paradigm uniformity as a constraint is only as strong as the argument that nouns are gender-marked in the input to diminutivization. Since the diminutive endings themselves encode gender, this assumption could be questioned.)

A particularly interesting set of examples of paradigm uniformity is cited by Downing (2005a: 24, 130 ff.), in a study of suffix doubling in Jita (Bantu). In Jita verbs, the causative suffix *-y* triggers mutation (spirantization) of any preceding /r/: /gur-a/ 'buy-FV' → [gura], vs. /gur-y-a/ 'buy-CAUS-FV' → [gusya]. Jita has at least two other derivational suffixes with which the causative can co-occur: applicative /-ir/ and reciprocal /-an/. When the causative co-occurs with either of these, it must double, occurring both directly after the root and directly after the other suffix, for example, /gur-y-ir-y-a/ 'run-CAUS-APPL-CAUS-FV' → [gusi:sya] or /gur-y-an-y-a/ 'run-CAUS-RECIP-CAUS-FV' → [gusyanya]. In verbs with causative, reciprocal *and* applicative suffixes, the causative must occur three times: /gur-y-ir-y-an-y-a/ 'run-CAUS-APPL-CAUS-RECIP-CAUS-FV' → [gusi:sya:nya]. Similar multiplication of the causative occurs in Kinande (Mutaka and Hyman 1990) and Cibemba, among other Bantu languages (Hyman 1994, 2003). According to Downing, the multiplication of the Jita suffix occurs under pressure from paradigm uniformity. Downing proposes that the causative suffix is always the one added morphologically first to the root (thus, for a verb with all three suffixes, the abstract underlying structure is /Root-CAUS-APPL-RECIP-/. Phonologically, however, the causative /-y/ is always required to be last in the stem, by a right-alignment constraint. Crucially on Downing's analysis, the phonological form of the (always innermost) Root-CAUS subconstituent is required to be uniform across all causative forms of a given stem. The only way to satisfy both the uniformity and the right-alignment requirement is to add the causative more than once. In the applicativized causative /gur-y-ir-y-a/, for example, the /gur-y- . . ./ portion satisfies stem uniformity while the /. . . -y-a/ portion satisfies rightward y-alignment. Downing argues (p. 128) against an alternative cyclic account of causative doubling facts,

such as the one proposed for parallel affix doubling facts in Cibemba and a number of other Bantu langauges by Hyman (1994, 2003), on the grounds that there is no other evidence for cyclicity in Jita.

2.4 *Paradigm Contrast*

Another manifestation of paradigmatic considerations is the morphological (or lexical) need to keep words or stems phonologically distinct from one another; this need for paradigm contrast has been argued to inhibit or trigger phonological effects.

For example, Crosswhite (1999) argues on the basis of evidence in the Trigrad dialect of Bulgarian that otherwise general rule of vowel reduction is blocked just in case it would cause the merger of two words in the same paradigm. In Trigrad Bulgarian, unstressed /o/ and /ɔ/ surface as [a], merging with underlying /a/:

(7)	/rog-ave/	['rogave]	'horns'
	/rog-ave-te/	[raga'vete]	'the horns'
	/sɔrp-ave/	['sɔrpave]	'sickles'
	/sɔrp-ave-te/	[sarpa'vete]	'the sickles'
		cf. [a'rala]	'plough'

Crosswhite observes that unstressed /o/ fails to reduce in a number of suffixes, for example, the *-o* ending on nominative masculine animate nouns: ['ago] 'older brother (nom.)', not *['aga]. Crosswhite observes that /o/ reduction fails precisely when, as in these cases, two distinct suffixes (one with /o/ and one with /a/) would merge if reduction applied. The accusative ending on masculine animate nouns is *-a*, as in ['aga] 'older brother (acc.)'. According to Crosswhite (and Kenstowicz 2005), vowel reduction is blocked when it would merge the nominative and accusative paradigm cells of masculine animate nouns. For a recent survey of these and other effects in which a neutralizing alternation is claimed to be blocked by a constraint against homophony, see Ichimura 2006.

According to Kurisu (2001), anti-homophony considerations can also trigger dissimilatory phonological alternations. Kurisu interprets a number of effects previously described as realizational morphology (see Section 5) as resulting from the requirement that input and output forms be distinct. On this view, process morphology is a repair of what would otherwise be the null realization of a morphological construction. Examples include the use of ablaut to mark plural in German (*Vater* ~ *Väter* 'father(s)', *Mutter* ~ *Mütter* 'mother(s)', p. 191), and the use of vowel deletion to derive deverbal nouns from infinitives in Icelandic (*klifra* 'climb-inf' → *klifr* 'climbing', *puukra* 'conceal (inf.)' → *puukr* 'concealment', p. 31, citing Orešnik 1978; Arnason 1980; Kiparsky 1984; Itô 1986; Benua 1995). Kurisu's analysis is that these constructions consist, morphologically, of zero-derivation, but that anti-homophony considerations compel the phonology to alter the output to avoid identity with the input. The fact that ablaut (in German) or vowel deletion (in Icelandic), are the preferred options, as opposed to any other imaginable

changes, follows, in Kurisu's account, from the ranking of faithfulness constraints penalizing deletion, insertion, and/or featural changes.

A challenge for Kurisu's view comes from cases of morphologically conditioned phonological effects applying alongside affixation, for example, German: *Gast* ~ *Gäst-e* 'guest(s)' or *Gaul* ~ *Gäul-e* 'pack horse(s)', with suffixation *and* ablaut (p. 191). Since affixation alone suffices to make two word-forms distinct in these cases, what motivates the accompanying ablaut effect? Kurisu's answer is that these cases are instances of double morphological exponence resulting from morphological opacity: the affixes in these examples are essentially invisible to the anti-homophony principle that requires the singular and plural cells of the paradigm to be distinct. The "first" layer of morphology is null, and phonology conspires to make the zero-marked plural stem (*Gäst*) distinct from the singular stem (*Gast*). The second layer of morphology, to which the phonology is blind, then double-marks the plural with a suffix: *Gäst-e*. Of course, double exponence is not limited to cases of this kind in which one exponent is arguably a phonological modification and the other is an overt affix; languages are known to use two or more overt affixes, or a suppletive stem plus overt affix(es), to mark a single category as well (e.g. Anderson 2001; Bobaljik 2000; Harris 2008a). Thus when ablaut is one of the two exponents of a morphological category, it could be analyzed, per Kurisu, as a phonological resolution to anti-homophony, or it could be attributed to whatever morphological factors are responsible for multiple exponence more generally.

Further afield, Wedel and Ussishkin (2002) have suggested that neutralizing phonological alternations can be inhibited if the words they would apply to exist in dense phonological lexical neighborhoods, that is, if there are high numbers of phonologically similar words in the lexicon. If this hypothesis is correct, contrast preservation might inhibit phonological alternations not only when the words in question are in the same paradigm, but even when they are morphologically unrelated. Dispersion might thus play an active role synchronically, not just the diachronic role suggested by Frisch, Pierrehumbert and Broe (2004) in their discussion of Arabic root consonants. Frisch, Pierrehumbert, and Broe show that the distribution of consonants in Semitic roots is skewed to favor triples of root consonants that are phonologically internally disparate over triples of root consonants that are internally similar. Frisch, Pierrehumbert, and Broe suggest a diachronic path by which dissimilatory phonological pressures affect the lexicon. Whether the pressures are purely diachronic or also synchronic is a question that future research is sure to focus on.

Whatever the nature of contrast preservation principles turns out to be, the principles clearly play a subordinate role in grammars. Phonological alternations and neutralizations are rampant, as is the creation of homophony in paradigms. Even setting aside all cases of systematic syncretism within paradigms (see e.g. Baerman 2005), we still find numerous situations in which phonological neutralizations create homophony. To take just one example, in Russian the neutralization of unstressed /a/ and /o/ produces homophony between nominative/accusative and genitive forms of neuter *o*-stems (Baerman 2005: 809). A desinence-stressed

stem, for example, 'wine', has distinct nominative/accusative (*vin[ó]*) and genitive (*vin[á]*) forms, but a root-stressed stem, for example, 'place', is identical in both contexts (*mést[ə]*), due to vowel reduction.

2.5 *Non-derived Environment Blocking (NDEB)*

It has been widely observed that neutralizing phonological alternations which are triggered at morpheme boundaries fail to apply when the same phonological environment occurs morpheme-internally. "Derived environment effects," or "non-derived environment blocking" (NDEB), has been generally attributed to contrast preservation pressures, although formal accounts of the phenomenon vary widely. The classic example of a derived environment effect occurs in Finnish: as noted by Kiparsky (e.g. 1993b), the neutralizing assibilation alternation converting /t/ to /s/ before /i/ and /e/ applies regularly at stem-suffix boundaries but does not affect morpheme-internal /ti/ sequences: *tilat-a* 'order-INFINITIVE' ~ *tilas-i* 'order-PAST', *but *silat-a, *silas-i.*

It was thought in the 1970s and 1980s that NDEB effects were associated with the class of cyclic, structure-changing rules; "Strict Cycle" principles proposed by Kiparsky (1982b) and Mascaró (1976) formalized this apparent correlation as part of the theory of Lexical Morphology and Phonology. However, subsequent findings (e.g. Hualde 1989a; Kiparsky 1993b) undermined the Strict Cyclicity correlation, showing that NDEB effects were not restricted to cyclic or to structure-changing rules and that not all cyclic or all structure-changing rules exhibit NDEB effects. A later wave of proposals, couched in Optimality Theory, focused on the tension between preserving input substrings from alteration, the idea being that morpheme-internal substrings (e.g. Finnish *ti*) would be preserved, but derived substrings (*t-i*) would be subject to alternation (e.g. Burzio 1997; Itô and Mester 1996b; McCarthy 2003a). A related approach is taken by Lubowicz (2002), who suggests that NDEB effects are those which apply only when input faithfulness has to be disrupted for some other reason, for example, resyllabification.

More recent work has gone back to the intuition that was first advanced by Kiparsky in the 1960s, namely that NDEB effects preserve contrast. Kiparsky's (1968a) Alternation Condition, though flawed in detail and later abandoned by Kiparsky (1982b) in favor of the Strict Cycle condition, captured the generalization that a given morpheme will undergo a neutralizing phonological alternation only if there is a contrast between morphological contexts in which the alternation is applicable to that morpheme and contexts in which the alternation is not applicable, making it possible for the underlying form of the morpheme to be recoverable by the learner.

For example, in Finnish, the initial *t* of *tilat* is always in the context of the Assibilation trigger *i*. By the Alternation Condition, it cannot alternate. The final *t* of *tilat*, however, sometimes occurs in an Assibilation context and sometimes does not. As a consequence it may alternate between *t* and *s* without obscuring the lexical contrast between stem-final /t/ anad /s/.

The Alternation Condition has found recent new life in work by Lubowicz (2003), who proposes that neutralizing alternations be constrained by a grammatical pressure to preserve contrast. Morphemes that contrast underlyingly should not be neutralized in every possible surface context in which they might occur.

As with anti-homophony, it is not clear that the effects labeled by various analysts as NDEB are all of the same type, functionally or formally. Some effects which could be classified under NDEB are more likely due to the restriction of the pattern in question to a particular cophonology within the language. In Japanese, for example, a condition of bimoraic minimality is imposed on affixed words, leading to vowel lengthening and/or inhibiting the degree to which suffixed stems can be truncated; but the requirement is not imposed on bare roots, even when used as words (Itô 1990). A similar minimality phenomenon in Turkish is documented in Itô and Hankamer 1989 and Inkelas and Orgun 1995. Though the specifics of their analyses differ, these authors essentially characterize the minimal size restrictions as properties of stems of a particular morphogical type. Roots are not stems of this type, and evade the minimal size condition by virtue of its never being imposed on them at all. This same sort of analysis is given by Yu (2000) to the phenomenon in Tohono O'odham whereby final secondary stress is prohibited except in morphologically complex words. Yu provides an explicitly cophonological account in which the assignment of secondary stress to final syllables is part of the cophonology of word-formation constructions, but not part of the cophonology applied to roots, even those used as words.

Cophonological accounts such as these have little to say about local NDEB effects of the type seen in Finnish; conversely, accounts of local NDEB effects do not extend to the more global effects seen in Japanese, Turkish, and Tohono O'odham. The Alternation Condition, whether in its original form or in Lubowicz's more modern incarnation, is not applicable to Tohono O'odham secondary stress, which is not neutralizing.

It could well be that there are simply two types of NDEB effects, which cannot be merged: those involving neutralization, which are typically segmental and therefore typically local and for which a contrast preservation approach is appropriate; and those which involve prosody, which are not local and do not involve contrast neutralization, for which the cophonological accounts are suited (Inkelas 2000). The typology of NDEB effects is clearly an area of ongoing research.

2.6 *Locality and Bracket Erasure*

An important question for any model of the morphology-phonology interface is whether phonological patterns applying to one subconstituent of a word can make reference to properties of embedded structure.

The existence of NDEB effects suggests that phonology needs to distinguish complex from simplex structures. In Finnish it is necessary for the phonology, when applying to a form like /tilat-i/, to have access to the information that *tilat* and *-i* differ in their morphological status.

Beyond NDEB proper, many other interactions between morphology and phonology have been analyzed using phonological rules or constraints that directly reference morpheme boundaries or morpheme identity. One type of evidence that is frequently adduced is root prominence. In Turkish, for example, the presence of a lexically accented morpheme (root or affix) in a word overrides the default assignment of stress to the final syllable. However, when both a lexically accented root *and* a lexically accented affix combine in the same word, one must disappear, since Turkish words have only one stress each. In Turkish it is always affix stress which disappears, giving rise to the appearance of what McCarthy and Prince (1995) have characterized as a universal principle of root-faithfulness: grammar is always more faithful to root structure than to affix structure, in situations where it is necessary to choose. Alderete (1999, 2001) has analyzed similar root-prominence effects in Cupeño and Japanese, characterizing them in terms of root-faithfulness. It does not particularly matter, of course, whether the analytical tool is a root faithfulness constraint or something else; what matters is that the phonological grammar must be sensitive to the distinction between roots and affixes. In cases like these, some cophonological accounts have a different interpretation of what is going on. It is possible, in cophonology theory, to treat affixation like realizational morphology, the result of a morphologically-specific phonological mapping that takes a stem as input and produces an output that includes what in more traditional morpheme-based accounts would be called the affix (Orgun 1996; Inkelas 1998). On this implementation of cophonology theory, the phonological substance of the "affix" is not present in the input; only the phonological substance of the root is present. The asymmetry between root and affix on this account does not require reference to morpheme boundaries or to the identity of morpheme types; it only requires reference to input. The most extreme view of the relevance of morpheme boundaries to phonology, then, would be that the rules or constraints within a particular cophonology are completely insensitive to morphology, and that morphological sensivity arises only indirectly by means of the association of different cophonologies with different morphological word-building constructions. The most permissive view would grant phonology access to all kinds of morphological information. This was the original assumption in generative phonology (Chomsky and Halle 1968), in which all morpheme boundaries were visible to phonological rules, and is still prevalent in the Optimality Theory literature (e.g. McCarthy and Prince 1993).

A view that falls somewhere in between was developed in the 1980s in the general Lexical Morphology and Phonology framework (Kiparsky 1982b; Mohanan 1986), in which it was assumed that phonological rules applying on a particular cycle (or stratum, if non-cyclic) could see morpheme boundaries created on that cycle (or in that stratum), but that once such rules had applied, the internal morpheme boundaries would be "erased" or on some principle made invisible to rules applying on a subsequent cycle (stratum) of morphology. (A version of this principle of bracket erasure can be found in Chomsky and Halle 1968 as well.) The question of bracket erasure and the relevance of morpheme boundaries has drawn little direct attention since the rise of Optimality Theory, aside from a few

works such as Orgun and Inkelas (2002), Itô and Mester 2002; Shaw 2009). While many analyses in Optimality Theory allow phonological constraints to directly reference all embedded morphological structure, it is not always clear whether this follows from necessity or from convenience.

3 When Phonology Affects Morphology: Combinatorics

Thus far we have focused on cases in which phonological patterns differ across different morphological zones of complex words. In this section we examine a different kind of interface, in which word-formation possibilities can themselves be constrained by phonology, either because of phonological requirements on inputs to word formation or because of phonological requirements on the outputs of word formation. Constraints on word formation can result in the choice of one suppletive allomorph over another, or they can result in morphological gaps, where no output (or only a periphrastic output) is possible. There are even cases in which it appears that affix ordering is phonologically determined.

3.1 Suppletive Allomorphy

Thus far we have discussed interactions ascribable to grammar. Suppletive allomorphy is a type of morphology-phonology interface which involves the lexicon. Suppletive allomorphy is familiar to every beginning morphology student as the situation in which a given morphological category has two or more exponents which cannot be derived from a common phonological form but must be stored separately. Suppletive allomorphy enters the realm of the morphology-phonology interface when the choice between or among suppletive allomorphs is phonologically determined.

In a number of such cases, the distribution of suppletive allomorphs appears to resonate with phonological patterns in the language, suggesting that the phonological grammar could be responsible for handling the allomorphy. In Modern Western Armenian, for example, the definite article takes the shape *-n* following vowel-final nouns (e.g. *katu-n* 'cat-def') and *-ə* following consonant-final nouns (e.g. *hat-ə* 'piece'); Vaux 1998: 252.[2] Similar effects are familiar from Korean, in which several suffixes exhibit V- and C-initial suppletive allomorphs which occur after C- and V-final stems, respectively. Thus, the nominative, accusative and topic-marked forms of *param* 'wind' are *param-i, param-il* and *param-in*, vs. the corresponding forms of *pori* 'barley': *pori-ka, pori-ril, pori-nin* (Paster 2006: 67, citing Odden 1993: 133). As researchers such as Mester (1994), Kager (1996), Anttila (1997a) and others have observed, constraints optimizing syllable structure (e.g. NoCoda, or Onset) would automatically entail the selection of allomorphs which produce CV syllables over those resulting in heterosyllabic consonant clusters (e.g. **hat-n*, in Armenian) or vowel sequences (e.g. **katu-ə*).

In a broad cross-linguistic survey of suppletive allomorphy, Paster (2006) uncovered a continuum of cases: some suppletive allomorphy (especially cases conditioned by syllable or metrical structure) is easy to characterize as phonologically optimizing, while other cases of allomorphy seem arbitrary or even non-optimizing. Consider, for example, the case of Haitian Creole, in which a particular determiner takes the form *-a* following vowels (e.g. *panié-a* 'the basket', *trou-a* 'the hole') and *-la* following consonants (e.g. *pitit-la* 'the child', *madãm-lã* 'the house') (Paster 2006: 86, citing Hall 1953: 32, via Klein 2003). This is the exact opposite distribution from what is seen in, for example, Korean, yet overall the syllable structures of the two languages are similar. If one allomorphic distribution makes sense phonologically, the other cannot. Or take Armenian noun pluralization: according to Vaux (1998: 31), "monosyllabic nouns take the suffix *-er* . . . and polysyllabic nouns take the suffix *-ner*: *ʧaʃ, ʧaʃ-er* 'meal(s)', *dodo, dodoʃ-ner* 'toad(s).'"

For apparently arbitrary phonologically conditioned allomorphy of this kind, lexical subcategorization is a common approach (e.g. Kiparsky 1982b; Inkelas 1990; Booij 2001; Paster 2006). The lexical entry includes all suppletive allomorphs, some or all of which are listed with a selectional frame identifying the phonological environment. In the case of the Armenian noun plural, for example, at least one of the suffix allomorphs must stipulate the number of syllables that the base of affixation is required to have. The other one can be the elsewhere case, if desired: {[[σ] er], [[ ] ner]}.

Paster argues for a subcategorization approach in all cases of suppletive allomorphy, even those which the allomorph distribution could be attributed to grammar rather than the lexicon. Her argument is partly based on the fact that suppletive allomorphy is often opaque, conditioned by input factors which are obscured in the output by phonological alternations affecting the derived stem. In such cases, input conditioning is necessary even if the distribution of allomorphs makes phonological sense. Paster discusses the example of Turkish, in which the third-person possessive suffix has two suppletive allomorphs: *-I*, used after consonant-final stems (*ev-i* 'his/her/its house'), and *-sI*, used after vowel-final stems (*anne-si* 'his/her/its mother') (Lewis 1967; see also Paster 2006: 99). This distribution is rendered opaque when intervocalic velar deletion applies to a suffixed stem. In the third-person possessive, a velar-final word like *inek* 'cow' combines with the *-I* allomorph, as expected since *inek* is consonant-final. However, the result of velar deletion is [ine.i] (orthographic *ineği*), with the "wrong" surface allomorph. A surface optimization approach, given the choice between [ine.i] and [inek-si], would almost certainly be expected to pick [inek-si] (or even [ine-si]); CC clusters across morpheme boundaries, as would occur in [inek-si], are commonplace and never repaired by deletion or epenthesis, whereas VV clusters across morpheme boundaries are tolerated at no other stem-suffix junctures in the language. For these reasons, Paster analyzes this case not as output optimization but purely as input selection.

In one very interesting case of opaque allomorph selection in Polish, Lubowicz (2007) cites phonological contrast preservation as the motivation for the choice between suppletive allomorphs. The locative in Polish has two suppletive

allomorphs: *-e* and *-u*. Like other front suffix-initial vowels in Polish, *-e* triggers palatalization of a stem-final coronal consonant: *lis[t]* (nominative), *o liś[ć]-e* (locative) 'letter'. Exactly those stems whose final consonant is underlyingly palatal take *-u* instead: *liś[ć]* (nominative), *o li liś[ć]-u* (locative) 'leaf'. Lubowicz attributes the selection of the *-u* allomorph to contrast preservation. Exactly when *-e*, the preferred allomorph, would merge the contrast between underlyingly plain and underlying palatal root-final coronal consonants, *-u* is selected instead. It is important to note that the contrast being preserved here is an abstract phonological one. While "letter" and "leaf" form a minimal pair, the same distribution of *-e* and *-u* is found with roots that are independently distinct in other ways, for example, *łobu[z]* (nominative), *o łobu[ź]-e* (locative) 'troublemaker', but *pa[ź]* (nominative), *o pa[ź]-u* (locative) 'type of butterfly'.

3.2 Phonologically Motivated Morphological Gaps

In phonologically conditioned suppletive allomorphy, phonological grammatical constraints or the phonological requirements of individual affixes control which stems can combine with which affix allomorphs. Sometimes the phonological grammar, or the phonological selectional requirements of individual affixes, can be so strict as to block morphological combination altogether, resulting in phonologically driven morphological gaps. For many speakers of Turkish, suffixation is grammatical only if the resulting word is at least disyllabic (Ito and Hankamer 1989; Inkelas and Orgun 1995). In Dutch, the superlative ending *-st* cannot be added to adjectives ending in [is], [sk], [st]; thus *bruusk* 'sudden' has no lexical superlative counterpart (**bruusk-st* [bryskst]) but must enter into a periphrastic syntactic alternative: *meest bruusk* (Booij 2005). In Tagalog, infixation of the agentive focus marker *-um-* is impossible if the stem begins with /m/ or /w/, creating paradigm gaps for such words (Schachter and Otanes 1972; Orgun and Sprouse 1999). A number of similar cases are surveyed by Carstairs-McCarthy 1998. Phonologically conditioned gaps differ from suppletive allomorphy in that there is no "elsewhere" allomorph; without this alternative, the word simply cannot be formed, resulting in a gap. Often there is a syntactic alternative; for example, in English, for example, the comparative suffix *-er* attaches only to (loosely speaking) monosyllabic stems; thus *vast-er* but **gigantic-er*, forcing speakers to resort to the periphrastic comparative, for example, *more gigantic* (Poser 1992).

3.3 Haplology Effects

Menn and McWhinney (1984) draw attention to a common cross-linguistic pattern of prohibiting sequences of homophonous morphemes, which they term the Repeated Morph Constraint (RMC). A well-known example occurs in English, where the possessive ending /z/ is not added to – or at least not realized on – words ending in the homophonous plural suffix /z/; thus *dogs* and *dogs'* are pronounced identically ([dagz]) in the phrases *the dogs hate their collars* and *the*

dogs' collars drive them crazy. Irregular plurals take the possessive (*children's*) and so do words ending in the strings homophonous with allomorphs of the plural (i.e. [(ɨ)z] or [s]), for example, *Katz's* [kætsɨz] or *cats* [kæts]. The RMC may, according to Menn and McWhinney, result in in haplology, as in the English example of *cats'*, where a single phonological exponent [s] stands for what appear to be two morphemes. In other cases the RMC can also trigger suppletive allomorphy or avoidance, in which a periphrastic alternative is preferred. For example, the English adverbial *-ly* ending (*quick* (adj.), *quickly* (adv.)) cannot combine with those adjectives already ending in *-ly,* for example, *manly* or *heavenly*: **manlily, *heavenlily* (adv.). The lexical gap which **manly*, and so on. cannot fill must be approximated by a phrase like *in an manly fashion*. Since it affects word form and creates morphological paradigm gaps, the RMC generalization would seem to be a clear case of phonology interfering with morphology.

The RMC is, however, clearly not universal, even within a language. There exist many unperturbed sequences of homophonous morphs; there are also instances of suppletion and avoidance in which morph repetition is not an issue. English, for example, permits sequences of the plural or possessive followed by the homophonous reduced form of *is,* for example, *one of the cats's* [kætsɨz] *trapped in the closet!* or *Whose guacamole do you like best? John's is* [dʒɑnzɨz] *clearly the winner*. The RMC applies only to sequences of plural and possessive, not to all sequences of /z/ morphemes. Another example occurs in Turkish, which uses the same suffix (/-I/~/-sI/) both to mark third-person singular possessors (*aile* 'family', *aile-si* 'his/her/its family'; *araba* 'car', *araba-sı* 'his/her/its car') and also as a marker at the end of head-modifier compounds, in which the possession relation, if any, is quite abstract (Lewis 1967: 42): *aile araba-sı* 'family car'. The possessive suffix cannot occur twice in succession; therefore, in isolation, compounds like *aile araba-sı* are actually ambiguous between a possessed (e.g. 'his/her/its family car') and unpossessed reading (Lewis 1967: 46). The ungrammaticality of a doubly affixed possessive compound, for example, **aile araba-sı-sı*, cannot, however, simply be attributed to the RMC. Other possessive suffixes, for example, first person possessive /-m/, "associative" -/lI/ and "occupational" /-CI/, are also in complementary distribution with the compound-marking possessive suffix even though they are not homophonous with it (Lewis 1967: 49–50): *aile araba-m* 'my family car', not **aile araba-sı-m*, etc.) Thus even this apparently transparent case of a repeated morph prohibition on possessive /-I/~/-sI/ turns out to be part of a more general pattern of morpheme co-occurrence. How, then, are we to know whether affix co-occurrence restrictions between homophonous affixes are a distinct subtype of affix co-occurrence restrictions generally? Further research is required in this area, but answers are likely to be of two types. One is statistical: if homophonous affix pairs form a larger than expected subset of the class of morpheme pairs that cannot occur next to each other, the RMC would be supported, though this task would be hard to accomplish given current data. A second possible answer would be to show that the lexical gaps or lexical ambiguities resulting from RMC effects pattern differently from those resulting from other, more arbitrary morpheme co-occurrence restrictions.

3.4 Linear Order

A number of cases have been described in which phonology constrains the linear order of morphemes. Mortensen (2006) has collected a very interesting set of examples in which constituents in coordinate compounding are ordered according to their phonological properties, principally vowel quality and tone. In one dramatic case from Jingpho, which Mortensen draws from a 1990 monograph in Chinese by Qingzia Dai, the order of elements in compounds with coordinate semantics follows from the height of the tonic (root) vowels: the stem with the higher vowel always precedes the stem with the lower vowel. Thus *lùʔ-ʃá* 'drink-eat = food' is a grammatical compound, while **ʃá-lùʔ*, with the same presumed meaning, would be ungrammatical (Mortensen 2006: 222–223). Mortensen documents many such compounding cases, mainly involving vowel quality and/or tone, in which the order of elements follows a scale. Sometimes the scale is phonetically transparent, as in the Jingpho case of vowel height, and sometimes not, when historical changes have obscured the original phonetic or phonological basis for the scale.

Another area in which phonology determines linear order is found with "mobile affixes," discussed by Fulmer 1991, Noyer 1994, and Kim 2008. These vary freely between prefixal and suffixal attachment, with phonological considerations being the deciding factors. In the San Francisco del Mar dialect of Huave, for example, the subordinate marker *m* attaches as a prefix to vowel-initial bases (*m-[u-ty]* 'SB-TV-eat = (that) s/he eats') but as a suffix to consonant-initial bases (*[mojk-o]-m* 'face.down-V-SB = (that) s/he lies face down'). Similar behavior is exhibited by other affixes, including the stative *n: n-[a-kants]* 'ST-TV-red = red' vs. *[pal-a]-n* 'close-V-ST = closed' (Kim 2008: 332).

As proposed by Kim 2008 and, for similar facts in San Mateo Huave, by Noyer 1994, such cases can be modeled in Optimality Theory by the general schema proposed by McCarthy and Prince (1994a) in which phonological considerations ("P") outrank morphological considerations, for example, affix ordering ("M"). In Huave, according to Kim, mobile affixes are preferentially suffixing (the "M" condition), but will prefix if suffixation would produce consonant clusters that would require epenthesis (the "P" condition) (pp. 340–341). Thus for a base like [a-rang] 'TV-do', *m-* prefixation (*m-a-rang*) is preferred over m-suffixation (**a-rang-m*, **a-rang-am*), since the latter would produce an unsyllabifiable cluster requiring repair. In cases where both prefixation and suffixation options would require epenthesis, suffixation is preferred: first-person *s* combines as a suffix with base *t-a-rang* 'CP-TV-do = did (it)' to yield *t-a-rang-as*, with epenthesis, rather than as a prefix (**s-tarang* or **sa-tarang*) (pp. 340, 342).

In general, however, the effect that phonology has been argued to play in the ordering of morphological elements is fairly limited. The great bulk of affix ordering is determined by the morphology, not by the phonology. To take a very simple example from Turkish, consider the interaction [FIX IPA] of the "occupational" suffix /-CI/ and case endings, for example, the dative /-E/. Both can attach to roots. Turkish epenthesizes vowels to break up triconsonantal

clusters, and epenthesizes glides to break up vowel-vowel sequences at stem-suffix boundaries. Thus we find alternations like these: /jeni-CI/ → [jenidʒi] 'new-PROF' (*yenici*), /film-CI/ → [filɨmdʒi] 'film-PROF = film-maker' (*filimci*); /jeni-E/ → [jenije] 'new-DAT' (*yeniye*), /film-E/ → [filme] 'film-DAT'. When both /-CI/ and a case suffix occur in the same word, affix order is fixed. /-CI/, as a derivational suffix, always precedes case: /jeni-CI-E/ → [jenidʒije] 'new-PROF-DAT' (*yeniciye*; /film-CI-E/ → [filɨmdʒije] 'film-maker (dative)' (*filimciye*). In the latter example, two epenthesis operations are required to bring the syllable structure of the resulting word into conformity with Turkish requirements. By contrast, the alternative affix ordering would produce perfectly well-formed syllables with no need for epenthesis: /film-E-CI/ → [fil.me. dʒi] (**filmeci*). But **filmeci* is completely impossible in Turkish; phonological considerations do not trump morphological constraints on relative affix order.

Paster (2005) explores one well-known apparent case, from the Fuuta Tooro dialect of Pulaar (Fula), in which phonology has been claimed to order affixes. In a study of the Gombe dialect, Arnott (1970) observed that a number of C or CV suffixes in the same general "zone" of the word appear to occur in a phonologically determined order: all suffixes with "t" precede all suffixes with "d," which precede all suffixes with "n," which precede all suffixes with "r." Paster cites similar examples of this "TDNR" template from Fuuta Tooro Pulaar, for example, *jaɓ-t-id-ir-an-ii* 'take-INTENSIVE-COMPREHENSIVE-MODAL-DATIVE-PAST' and *yam-d-it-in-ir-ii* 'healthy-DENOMINATIVE-REPETITIVE-CAUSATIVE-MODAL-PAST' (Paster 2005: 164). As Paster argues, however, the order of affixes in both dialects of Pulaar conforms to semantic ordering principles of the kind articulated by Bybee (1985) and Rice (2000); there is no case in which the phonological TDNR template contravenes an ordering that one might otherwise expect on morphological grounds, and thus no clear evidence that phonology is interfering with morphology. Paster also observes that the TDNR template, in which consonant sonority increases from left to right, is not completely convincing as a phonological phenomenon, since in actual Pulaar words, vowels typically separate the consonants which correspond to the elements of the TDNR template. A sonority-based template like TDNR would make more sense, Paster argues, if consonants were being ordered by sonority in order to fit into a single syllable onset or coda, but this is not the case in Pulaar. In sum, Paster concludes, the Pulaar pattern is significant for coming closer than any other example to being a case of phonologically-driven affix sequencing but still not fully meeting that description.

The most substantial influence of phonological considerations on the linearization of morphemes is found in infixation, which is generally viewed as being just like affixation except that the affix is phonologically positioned within the stem instead of peripheral to it (see e.g. Moravcik 1977, 2000; McCarthy and Prince 1993; Yu 2007). The interest of infixation for the phonology-morphology interface lies in phonological generalizations about where in a word an infix can appear and about what, if anything, motivates infixation synchronically.

Surveys of infixation from Moravcsik 1977 to Yu 2007 have found a small and principled set of recurring sites for segmental infixes: next to the initial or final

consonant or vowel, or (in lexical stress languages) next to a metrical prominence. As Yu observes, these sites are defined in terms of types of elements that all words in the relevant language contain. All words contain consonants, vowels, and (in stress languages) stress. By contrast, there is apparently no evidence of infixation to syllables with particular tones, or to syllables containing particular types of segments (e.g. fricatives or ejectives) or even to heavy syllables (e.g. those with long vowels or consonant clusters). These are elements that languages do not typically require all words to possess.

Only dependent morphemes (affixes, and rarely clitics; Harris 2000) have ever reliably been shown to infix; infixation is apparently not a possible property of compounding or phrasal combination, aside from the suggestive example of expletive infixation before a stress foot in English words (*amálgamàted* → *amálga-bloody-màted, Kàlamazóo* → *Kàlama-fuckin-zóo,* and so on.; McCarthy 1982).

A stimulating theory of infixation was introduced in the early 1990s, within the framework of Optimality Theory, by McCarthy and Prince (1993), who observed that many cases of infixation could be interpreted as improving the prosodic structure of the derived word in comparison to the structure the word would have if the infixed element were instead adfixed. Infixation was a key motivator in McCarthy and Prince's proposal that at least some phonological constraints "P" can outrank morphological constraints "M," particularly those having to do with the edge-alignment of affixes.

The most convincing examples brought forth for this view, termed the "Phonological Readjustment" view in Yu 2007, involve syllable structure. McCarthy and Prince's original example concerns the agentive focus marker, *-um-*, in Tagalog, which precedes the initial vowel: *bilih* 'buy' → *b-um-ilih*, *gradwet* 'graduate' → *gr-um-adwet*, and so on. According to McCarthy and Prince, *-um-* is a prefix, subject to the "M" constraint Align-L(um-, Stem), which is outranked by a "P" constraint banning closed syllables (NoCoda). In considering the possible locations for *-um-* in case of *gradwet*, the prefixing candidate *um.grad.wet* has two closed syllables, while the infixing candidate *gr-u.m-ad.wet* has only one closed syllable, satisifying NoCoda better. Infixation is preferred because of the P >> M ranking.

By contrast to numerous examples in which infixation can be interpreted as improving syllable structure, very few examples have been found in which infixation can be construed as improving segment structure (see e.g. Yu 2007: Chapter 6). This asymmetry raises doubts about the generality of the potential for "P" constraints to outrank morphological alignment.

A second major issue confronted by the P >> M model is locality. If infixation is misalignment with the aim of avoiding bad structures, the P >> M model predicts what Yu calls "hyperinfixation" or unbounded infixation, a point made also by Orgun and Sprouse (1999) and McCarthy (2003b). In Section 3.2 we mentioned that the Tagalog *-um-* infix is prohibited from combining with *m-* or *w-* initial stems. Modeling this prohibition as a phonological constraint (e.g. OCP, following Orgun and Sprouse), a P >> M ranking would predict that *-um-* could infix further into the word than usual to avoid the undesired *mu* or *wu* sequence. This does

not happen in Tagalog, nor does the comparable situation appear to arise in any other language: infixes are tightly restricted to appear near edges. McCarthy (2003b) addresses this problem for P >> M by modifying alignment constraints so that they are categorical; Tagalog *um* is subject both to a violable constraint forbidding it from being separated by a segment from the beginning of the word and to an inviolable constraint forbidding it from being separated by a syllable (or more) from the beginning of the word (p. 95 ff.).

A radically different approach to infixation is offered by Yu (2007), who observes that many cases of infixation fall outside the P >> M model in the sense of not being prosodically improving in any discernible way, yet still conform to the locality generalizations that were potentially problematic for the P >> M model. Yu points out that some cases of infixation neither improve nor worsen syllable structure For example, the Hua negative infix *-ʔa-*, which is CV in shape, infixes before the final syllable: *harupo* → *haru-ʔa-po* '(not) slip', *zgavo* → *zga-ʔa-vo* '(not) embrace', even though adfixing (*ʔa-harupo, harupo-ʔa*) would have produced equally good syllables (Yu 2007: 30, citing Haiman 1980). Other cases of infixation arguably make syllable structure worse. For example, the nominalizing *-ni-* infix in Leti follows the first consonant (e.g. *kaati* 'to carve' → *k-ni-aati* 'carving', *polu* 'to call' → *p-ni-olu* 'act of calling, call'), producing marked consonant clusters and vowel sequences that would be avoided by simple adfixation (e.g. *ni-polu, polu-ni*) or infixation to a different position (e.g. *po-ni-lu*) (Yu 2007: 28, citing Blevins 1999).

Yu concludes that locality and generality, rather than phonological optimization, are the main generalizations that a synchronic model of infixation should capture, and proposes a lexical subcategorization approach building on, for example, Broselow and McCarthy 1983, McCarthy and Prince 1986, Inkelas 1990. On this approach, each infix is associated with a lexical statement defining its position relative to one or both edges of the stem it combines with. Phonological entities to which such statements are permitted to refer come from a small list of "pivots" that cross-linguistically are shown to separate infixes from stem edges: segments, syllables, and stressed constituents (Yu 2007: 52).

4 When Phonology Affects Morphology: Form

Some morphological constructions are phonologically compositional, in the sense that the morphology combines two or more elements with fixed phonological shapes and the "regular" rules of the phonology apply to give the combination its surface phonological form, which varies with the shapes of the input morphemes. For example, prefixation of *pre-* in English (*pre-register*, *pre-ordained*) is a simple matter of concatenating the fixed string [pri] with a base. But in some morphological constructions, extrinsic considerations constrain or determine output phonological shape, with input morphemes conforming to output shape requirements instead of determining output shape themselves. We will survey two such phenomena here: templatic morphology and reduplication. These are commonly termed "prosodic morphology," because in each case a morpheme is

expressed phonologically in a way that is not constant across the set of stems formed from that construction but is predictable from construction-specific metrical or syllabic constraints on the phonological shape of the complex stem. Prosodic morphology is often described as a trio, with infixation as the third member; however, the considerations driving infixation are rather different, as seen above.

4.1 Templates

Templates are morphological constructions, typically associated with specific derivational or inflectional morphological categories, which directly constrain the phonological shape of the derived stem. McCarthy (1979a, 1981) broke new ground by analyzing the fixed shape of specific derivational subtypes of Arabic verbs as composed of templates consisting of CV timing units. These templates, each expressing a specific morphological category, combine with other morphemes which consist of consonants, and with others consisting of vowels, to form complex words. For example, the consonantal root /ktb/ 'write' combines with the "perfective passive" vocalic morpheme /ui/ and the "causative" template CVC-CVC to form *kuttib.* In their seminal 1986 paper, McCarthy and Prince showed the role of prosodic structure in defining the various templates not just in the Arabic root and pattern morphological system but more generally in prosodic morphology cross-linguistically. According to McCarthy and Prince, templates are always defined in terms of the universally accessible units of mora, syllable, and foot, rather than in terms of the C, V, or X timing units proposed in earlier work by McCarthy (1979a, 1981), Leben (1980), Hyman (1985), and others.

Sometimes templates constrain the shape of stems or words without contributing any particular semantic or syntactic function of their own. A simple case of this occurs with minimal word size constraints, which can compel epenthesis or other phonological augmentation strategies in short words. In Swati, as in many other Bantu languages, a disyllabic minimality requirement on words compels the use of a dummy suffix *-ni* in verbs that would otherwise be monosyllabic, a situation which arises in imperatives, the one morphological environment with no prefixes. Thus, while the infinitive of the stem /dlá/ 'eat' is *kû-dlá*, disyllabic by virtue of containing the infinitive prefix, the unprefixed singular imperative is *dlá-ni*, with augmentation that is not required for verbs formed from longer stems, such as /bóna/ 'see': *kû-bóna* 'INF-see = to see', *bóna* 'see (singular imperative)' (Downing 2006: 3). In Lardil, uninflected nouns are subject to apocope, seen in alternations like *wiwala-n* 'bush mango-NONFUT.ACC', *wiwala-ṛ* 'bush mango-FUT.ACC', but *wiwal* 'bush mango, from /wiwala/; *karikari-n* 'butterfish-NONFUT.ACC', *karikari-wuṛ* 'butterfish-FUT.ACC', but *karikar* 'butterfish' (from /karikari/) (Hale 1973: 424). Apocope is blocked when the result would have only one short vowel, for example, *kela-n* 'beach-NONFUT.ACC', *kela-ṛ* 'beach-FUT.ACC', *kela (*kel)* 'beach', from /kela/ (Hale 1973: 421). Uninflected nouns with only one short vowel are even subject to augmentation, so that they achieve bimoraic size: *ṭer-in* 'thigh-NONFUT.ACC', *ṭer-uṛ* 'thigh-FUT.ACC', but *ṭera* 'thigh', from /ṭer/ (Hale 1973: 427). Uninflected nouns in Lardil are clearly affected by the prosodic limitation

on word size. (For further discussion of Lardil, see for example, Kenstowicz and Kisseberth 1979, Itô 1986, Blevins 1997, and Bye 2006).

A more involved case, in which prosodic templates constrain stem shape, occurs in Yawelmani (Archangeli 1983, 1991, based on Newman 1944). In the verbal system, each root and affix is lexically associated with one of three prosodic templates: a light syllable, a heavy syllable, and an iambic foot consisting of a light syllable followed by a heavy syllable. These templates determine the form of the root. When, in the same word, an affix and root are associated with conflicting templates, the one associated with the affix prevails, leading to root alternations in related stems. For example, the root "walk" has a default iambic template, as in *hiweet-en* 'will walk', but shortens to a heavy syllable and to a light syllable when combining with suffixes associated with the corresponding templates, as in *hewtihni* 'one who is roaming' (< *hiiwt-(ʔ)ihni*) and *hiwtiṅay* 'while walking' (< *hiwt-(ʔ)iṅay*) (Archangeli 1991: 232, citing Newman 1944: 101, 110, 136).

Prosodic templates can even specify segmental content. In Tiene verbs, derivational stems are constrained by a CV(C)VC- template whose consonants are subject to two major restrictions (Hyman and Inkelas 1997 and Hyman 2006a, based on Ellington 1977). In CVCVC- stems, the middle consonant (C_{med}) must be coronal and the final consonant (C_{fin}) must be non-coronal, that is, labial or velar. These restrictions can force the choice of infixal allomorphs of suffixes such as the stative, which has both an infixal allomorph (with coronal /l/) and a suffixal allomorph (with non-coronal /k/): *kab-* 'be divided' ~ *ka-la-b-* (stative); *yat-* 'be split' ~ *yat-ak-* (stative); *sɔn-* *'write'* ~ *sɔn-ɔŋ-* (stative). C2 and C3 must also agree in nasality, leading to nasal~oral alternations: *vwuɲ-* *'be mixed'* + *-ek-* *(stative)* → *vwuɲ-eŋ-*; *dim-* 'become extinguished' + *-se-* (causative) → *di-se-b-*.

In recent work in Optimality Theory, starting with McCarthy and Prince 1994a, researchers have argued that templates are emergent artifacts of constraint interaction, rather than abstract structures manipulated by grammar. In Arabic and Tiene, for example, it might be possible, instead of stipulating baldly that the (derivational) verb stem must be CV(C)VC- in shape, or even simply that it must be bimoraic, to let this profile emerge from constraints like Ft-Bin (feet are binary) and All-Ft-Left (every foot must be initial), ranked high in the morphological environment of the verb stem. The emergent template approach has been applied fruitfully to many cases of reduplication by McCarthy and Prince (1995) as well as Gafos (1998a) and Hendricks (2001), among others, and has been extended beyond reduplication by Downing (2006). The motivation for deriving rather than stipulating templates is two-fold: first, deriving templates from independently needed markedness constraints should yield a more limited, principled set of possible templates than what it is possible to stipulate, and second, templates constrain form in ways other than simple prosodic size. Markedness constraints can constrain segmental form as well. The flexibility of emergent templates is useful in characterizing cases like Tiene, in which the restriction about consonantal place of articulation in verbs cannot be expressed by annotating particular prosodic positions for segmental features. C_{fin} is unrestricted in CVVC stems; C_{fin} is constrained, by dissimilatory principles, only if C_{med} is present. This kind of contingent restriction is better suited to constraints of the sort posited by Hyman

and Inkelas, in which C_{med} and C_{fin} must differ in place of articulation (and C_{med} must be coronal).

4.2 Reduplication

Reduplication is the doubling of some part of a morphological constituent (root, stem, word) for some morphological purpose. Total reduplication duplicates the entire constituent. It is often, though nowhere near always, semantically iconic, as in the duplication of nouns with human reference to form plurals in Warlpiri (*kurdu* 'child', *kurdu-kurdu* 'children'; *wirriya* 'boy', *wirriya-wirriya* 'boys') (Nash 1986: 130).

Partial reduplication, which exhibits a very wide range of meanings, usually involves a prosodically characterized template for the reduplicant. For example, McCarthy and Prince (1986) analyze the reduplicating prefix marking progressive aspect in Mokilese as a bimoraic syllable: *poadok* [pɔdok] 'to plant' → *poad-poadok* [pɔdpɔdok] ~ *poah-poadok* [pɔɔpɔdok] 'to be planting'; *piload* [pilɔd] 'to pick breadfruit' → *pil-piload* [pilpilɔd] 'to be picking breadfruit'; *kohkoa* [kookɔ] 'to grind coconut' → *koh-kohkoa* [kookookɔ] 'to be grinding coconut' (Harrison and Albert 1976: 60, 220). Typically the base exceeds the reduplicant template in size and thus the resulting reduplication is, as in these Mokilese data cited, partial. Occasionally, a reduplicant template will be bigger than the base; in which case reduplicant augmentation occurs. For example, the Mokilese form *pa* 'weave', which is monomoraic, reduplicates as *pah-pa* [paapa], with a bimoraic reduplicant (Harrison and Albert 1976: 60).

While typical reduplicant shapes are described in the prosodic units of mora, syllable and foot, Moravcsik (1977) is credited with the observation that reduplication rarely unambiguously copies an *existing* mora, syllable or foot from the base. Rather, as modeled by theories like Prosodic Morphology (McCarthy and Prince 1986) and, subsequently, approaches to reduplication within Optimality Theory (McCarthy and Prince 1994a, 1995), templatic requirements seem to be output requirements on the reduplicant. In Mokilese, what copies is enough base material to flesh out a heavy syllable reduplicant, even if the corresponding string is not itself a syllable in the base, (e.g. reduplicant [pil], from base [pilɔd]). In Optimality Theory this output orientation can be modeled by stating reduplicant shape as the output requirement RED = $\sigma_{\mu\mu}$. McCarthy and Prince (1994a) and Urbanczyk (1996), working in Generalized Template Theory, and, from a different angle, Downing 2006, have pursued the goal of deriving templates rather than stipulating them. In different ways, these researchers propose that reduplicants assume the canonical phonological form of whatever morphological constituent type (affix, stem, morphologically complex stem) they instantiate. This form does not have to be stipulated specially for the reduplicant but is motivated more generally for the language, or even cross-linguistically. Recent literature has suggested that some reduplication, particularly when limited to consonants, may not have even an indirect prosodic templatic characterization at all. Hendricks (1999, 2001) points to cases such as expressive reduplication in Semai, which copies the first and last consonant of the base: *pŋ-payaŋ* 'appearance of large stomach constantly bulging out', *cw-cruha:w* 'sound of waterfall, monsoon rain' (Diffloth 1976).

Partial reduplication usually duplicates that edge of the stem to which the reduplicant is closest, but opposite-edge reduplication (not including the dual-edge version found in Semai) occurs as well. A dozen or so cases are documented in surveys by Nelson 2003, 2004; Kennedy 2003; and Riggle 2003; all target the beginning portion of a base, for example, Koryak CVC reduplication marking absolute case: mɪtqa 'oil' → *mɪtqa-mɪt; qanga* 'fire' → *qanga-qan.*

Partial reduplication is also commonly infixing, as in Chamorro (Topping 1973: 183), where habitual/continuative CV reduplication targets stressed syllables (*hátsa* 'lift' → *há-ha-tsa*, *hugándo* 'play' → *hugá-ga-ndo*) and intensifying CV reduplication targets the final syllable (*métgot* 'strong' → *métgo-go-t* 'very strong', *ñálang* 'hungry' → *ñála-la-ng)*; see for example, Broselow and McCarthy 1983: 55–56). Internal reduplication usually duplicates adjacent material, as in these examples, but there are some exceptions to this. In Washo, for example, plural reduplication infixes a mora in the vicinity of the stressed syllable. In case the stressed syllable is closed, as in *nén.t'uš* 'old woman (nom.)' or *ʔéw.ʃiʔ* 'father's brother', reduplication copies a non-adjacent CV string: *ne.t'ún.t'uš-u* 'old women (nom.)', *ʔe.ʃíw.ʃiʔ* 'father's brothers' (Yu 2005: 440, citing Jacobsen 1964). Creek plurals are formed by infixing a copy of the stem-*initial* CV before the stem-*final* consonant (Riggle 2003, citing Booker 1980; Haas 1977 and Martin and Mauldin 2000): *holwak-íː* 'ugly, naughty' → *holwaː-ho-k-íː*, *falápk-iː* 'crooked' → *falap-fa-k-íː.*

4.2.1 Identity Effects in Reduplication: Over and Underapplication Since Wilbur's influential (1973) dissertation, researchers have paid special attention to phonological opacity arising in reduplication constructions. For example, consider Javanese total reduplication, which has pluralizing semantics and can apply to verbs and adjectives. When suffixed, for example, by demonstrative *-e,* reduplicated forms exhibit "overapplication" of intervocalic *h*-deletion and underapplication of closed syllable laxing and stem-final consonant devoicing (Inkelas and Zoll 2005: 146, 148, citing Dudas 1976: 207–208):

(7)

Gloss	*Stem*	*-demonstrative*	
'broken'	bəḍah	bəḍa-e	normal application of intervocalic h-deletion
	bəḍah-bəḍah	bəḍa-bəḍa-e	"overapplication" of h-deletion in first stem copy (demonstrative)
'cylindrical'	gilɪk	gilig-e	normal application of intervocalic voicing
	gilɪk-gilɪk	gilig-gilig-e	"underapplication" of stem-final devoicing in first stem copy (demonstrative)

Although opacity occurs outside of reduplication as well, its appearance in reduplicative examples like these is interpreted by Wilbur (1973), who posits a reduplicative Identity Principle, and in Base-Reduplicant Correspondence Theory, which posits Base-Reduplicant faithfulness constraints, as driven by the functional need to keep the two parts in reduplication – base and copy – segmentally identical. This imperative, while apparently obeyed in Javanese, is not satisfied in every case. For example, Urbanczyk (1996) and Struijke (2000) draw attention to a reduplicative construction in Lushootseed, illustrated by examples like *wális* 'type of frog' → *wá-w'lis* 'little frog', *caq'(a)* 'spear' → *ca-cq'* 'act of spearing big game on water', and so on. (Urbanczyk 1996: 167). Urbanczyk and Struijke analyze this pattern as CV prefixing reduplication accompanied by syncope in the base. On this account, underapplication of syncope would better preserve reduplicant-base identity (*wá-walis*); however, syncope applies transparently anyway without impedance from base-reduplicant identity constraints.

Some cases of reduplicative opacity can be attributed to layering or stratal aspects of the phonology-morphology interaction. With regard to the Javanese case above, Inkelas and Zoll (2005: Chapter 5) argue that demonstrative suffixation occurs prior to reduplication, triggering *h*-deletion and consonant voicing and preventing closed-syllable laxing from occurring; the suffixed stem (for example, *bəḍa-e* or *gilig-e*) is then input to reduplication, which copies the root as is, preserving the effects of the stem-level phonological alternations. What portion of reduplicative opacity will yield to layering accounts as proposed by Inkelas and Zoll 2005 and Kiparsky 2010, and what portion requires identity principles, is still an open question. It may be important, in deciding this question, to factor apart morphologically driven reduplication, such as the Javanese example, from phonologically-driven segment duplication. The latter clearly requires phonological identity principles (copying or correspondence, as appropriate to the theoretical framework in use). See Hendricks (1999, 2001), Yu (2005), Riggle (2006), Inkelas (2008a), and Pulleyblank (2009) for discussion relevant to the distinction between morphological reduplication and phonological copying.

4.2.2 Fixed Segmentism in Reduplication It is often the case that one of the two copies in morphological reduplication contains some fixed material which either co-occurs with or supplants material that would otherwise be expected to copy. An example of the former occurs in Khasi, where iterative verb reduplication connects the two copies of the verb with the linker *ši*, for example, *iaid-ši-iaid* 'to go on walking', *leh-ši-leh* 'keep repeating', *kren-ši-kren* 'keep talking' (Abbi 1991: 130, cited in Inkelas and Zoll 2005: 36). An example of the latter occurs in English, where an ironic/derisive total reduplication construction assigns "shm" to be the onset of the second copy, replacing an existing onset, if any: *fancy-shmancy*, *handsome-schmandsome*, *OT-shmOT,* and so on; see for example, Alderete *et al.* (1999). This phenomenon has been termed "Melodic Overwriting" (McCarthy and Prince 1986; Yip 1992; Alderete *et al.* 1999).

One functional motivation that has been offered by Yip (1997) for Melodic Overwriting is that it makes the two copies different. Support for this interpretation

is found in (a) the predominance in of Melodic Overwriting in total, as opposed to partial, reduplication and (b) the fact that constructions involving Melodic Overwriting sometimes block when the two copies would be identical phonologically, or, perhaps more commonly, exhibit dissimilatory allomorphy which guarantees that the copy with the fixed material is different from the intact copy. In Turkish, as described by Lewis (1967: 237–238), a construction meaning "and so on, and suchlike" doubles a word and imposes the onset *m* on the second copy, replacing an existing onset if there is one (e.g. *dergi* 'journals', *dergi mergi* 'journals or periodicals or magazines'). According to Lewis, this construction cannot be used if the word begins with *m* already (e.g. *müfettişler* 'inspectors'), and a periphrastic construction with *falan* or *filân* is used in its place (*müfettişler falan* 'inspectors and all that lot'). In Abkhaz, suppletive allomorphy comes to the rescue in the comparable situation. Bruening (1997), citing Vaux (1996), describes an Abkhaz echo-word construction which replaces the onset of the second copy with /m/ (*gazá-k'* 'fool' → *gazák'-mazák'*); however, if the word already begins with /m/, /č'/ is used instead (*gazá-k'* 'secret' → *mazá-k'-č'azá-k'*). This kind of required dissimilation seen in Melodic Overwriting is in some ways reminiscent of the anti-homophony morphological effects described in Section 2.4, which require inputs and outputs, or members of the same paradigm, to differ. Yip (1997, 1998) likens dissimilatory Melodic Overwriting to the kind of conventional poetic rhyme in which identity is required in one prosodic location (e.g. the syllable rhyme) but non-identity is required elsewhere (e.g. the onset of that same syllable); thus *rhyme-time* is a good rhyme but *rhyme-rhyme* is not.

Alderete *et al.* (1999) have argued that some cases of fixed segmentism are, rather than instances of Melodic Overwriting, instead the result of reduction driven by emergent unmarkedness, a phenomenon observed by Steriade 1988 to characterize partial reduplication. In Nupe gerundives (Downing 2004: 90, citing Akinlabi 1997; Smith 1969), an initial CV reduplicant has a fixed [+high] vowel and mid tone, regardless of what is found in the base: *kpi-kpà* 'drizzling', *ji-jákpe* 'stooping', *bi-bé* 'coming'. Insofar as [+high] and mid are the unmarked values for vowels and tone in Nupe, as argued by Akinlabi, the fixed values in the reduplicant can be derived, rather than stipulated. Reduction in partial reduplication is consistent with the hypothesis that partial reduplication typically derives historically from erosion of total reduplication (e.g. Niepokuj 1997). Total reduplication, however, virtually never displays phonological reduction in one copy, as observed in Inkelas (2008a: 379–380).

4.2.3 Morphological Character of Reduplicant It is tempting, based on form, to characterize partial reduplication as affixation and total reduplication as compounding. However, there is little morphological evidence for this distinction. Indeed, some phonologists have recruited the affixation/compounding distinction to account for phonological size differences *within* partial reduplication, terming reduplicants which are syllable-sized or smaller "affixes" and those which are foot-sized "roots" (Generalized Template Theory; for example, McCarthy and Prince 1994, Urbanczyk 1996). This distinction is generally motivated not by

morphological criteria but by the desire to avoid morpheme-specific reference to prosodic templates. Taking a different view, Inkelas and Zoll (2005) observe that reduplication constructions do sometimes impose morphological restrictions that are independent of prosodic ones, arguing against conflating the two types of restriction; they point (in Chapter 2) to the distinction between constructions which specifically double affixes, regardless of size, vs. those that double roots or stems as evidence that reduplication targets morphologically defined constituents and then imposes phonological shape requirements on the output of doubling.

5 When Phonology is Morphology: Realizational Morphology and Morphologically Conditioned Phonology

Morphologically conditioned phonology overlaps significantly with what has been called "realizational" or "process" morphology, an observation made by, among others, Ford and Singh (1983, 1985), Poser (1984), Dressler (1985), Singh (1996), S. Anderson (1992), Bochner (1992), Orgun (1996), Inkelas (2008b).

To illustrate these phenomena and their overlap, consider two cases of final consonant deletion. The first is a well-known process of subtractive morphology in Tohono O'odham, discussed by S. Anderson (1992), citing Zepeda 1983, and Yu (2000), citing Zepeda 1984. In Tohono O'odham, perfective verbs are derived from imperfectives through the deletion of a final segment (*síkon* 'hoe object' → *síko* (-PERF); *híwa* 'rub against object' → *híw* (-PERF) (Yu 2000: 129–130). This fits the standard description of realizational morphology because there is no other morphological exponent of the perfective.

Now consider the diminutive suffix /-CIk/ in Turkish, discussed in Section 2.1, which triggers an optional process of stem-final velar deletion (/bebek-CIk/ → *bebecik* 'baby-DIM', /köpek-CIk/ → *köpecik* 'dog-DIM') (Lewis 1967: 57) that applies before no other similar suffix. This would standardly be described as a morphologically conditioned phonological rule, because the morphological category of diminutive is marked overtly by the suffix. The operative intuition is that the suffix *-CIk* marks diminutive morphology, while the consonant deletion is just incidental.

This practical distinction between phonology as primary exponent and phonology as secondary concomitant does not always hold up. Sometimes it is difficult, even unproductive, given several exponents of a given morphological construction, to identify which is the primary (morphological) exponent and which are the phonological accompaniments. In Hausa, for example, tone melody replacement can serve as the sole mark of a morphological construction (8a), and so can overt affixation (8b). When both co-occur (8c), is tone melody replacement considered realizational morphology, such that the words in (8c) exhibit two morphological exponents of pluralization, or is tone melody replacement subjugated in (8c) to morphologically conditioned phonology? Page numbers refer to Newman 2000:

(8) a. No affixation; tone replacement (imperative formation)

ká꞉mà꞉	→	kà꞉má꞉	'catch (!)'	(267)
bíncìké꞉	→	bìncìké꞉	'investigate (!)'	(267)
nánné꞉mó꞉	→	nànnè꞉mó꞉	'seek repeatedly (!)' (< *né꞉mó꞉* 'seek')	(263)

b. Overt suffixation, no tone replacement (various)

dáfà꞉	→	dáfà꞉-wá	'cook-VBL.N' -LH	(699)
gàjé꞉ré꞉	→	gàjé꞉r-ìyá꞉	'short-FEM' -LH	(212)
hù꞉lá꞉	→	hù꞉lâ-r	'hat-DEF' -L	(144)

c. Overt affixation *and* tone replacement (various plural classes)

má꞉làm	→	mà꞉làm-ái	'teacher-PL.' -LH	(434)
hù꞉lá꞉	→	hú꞉l-únà꞉	'cap-PL.' -HL	(444)
tàmbáyà꞉	→	támbáy-ó꞉yí꞉	'question-PL.' -H	(432)

One possible way to avoid analytical ambiguity in the case of (8c) would be to reduce all phonological effects other than overt segmental affixation to morphologically conditioned phonology, reanalyzing apparent cases of realizational morphology as zero derivation accompanied by morphologically conditioned phonology. Alternatively, one could try to reduce all morphologically-specific phonological effects to realizational morphology, analyzing the data in (8c) as instances of "extended exponence," the multiple marking of a morphological category (for example, Matthews 1972; Stump 1991). Multiple exponence of overt morphology is a common enough phenomenon; in Hausa, for example, the formation of class 13 noun plurals involves suffixation (of *-e*), (LH) tone replacement, *and* reduplication, for example, *kwánà꞉* 'corner, curve' → *kwàné-kwàné* (PL) (Newman 2000: 458). Harris (2002, 2008b) has argued that circumfixes result diachronically from an earlier stage of multiple affixation, or multiple exponence.

5.1 *Theoretical Approaches to Realizational Morphology and Morphologically Conditioned Phonology*

The literature on morphologically conditioned phonology, primarily represented by item-based approaches, has had little to say about realizational morphology, despite the obvious formal similarities between the two phenomena. One reason for this is that much of the most influential literature on morphologically conditioned phonology, going back to Kiparsky's (1982b) theory of Lexical Morphology and Phonology (LMP; see also Kiparsky 1984; Mohanan 1986), focuses on phonological patterns common to the morphology of a certain stratum. Both LMP and its successor, Stratal Optimality Theory (Kiparsky 2000, 2008a, to appear), assume a grammatical architecture in which the morphological constructions of each language cluster into a small, possibly universally fixed number of sets ("levels," "strata"), each internally uniform in its phonological patterning, which are totally ordered. In stratal theories like these it is necessary to know only the stratum to which a morphological construction belongs to predict which phonological patterns it will conform to; which stratum a construction belongs to is predictable

from whether its place in the morphology, that is, whether it is an "early" or "inner" affix as opposed to a "late" or "outer" one.

Because stratal theories focus on commonalities, they are not suited to the description of phonology which is unique to a particular morphological category; very narrowly conditioned phonological effects have to be set aside and treated as exceptions within a stratum, rather than constituting their own individual stratum. For example, in English both *-ible* and *-ive* trigger spirantization on a preceding consonant, an unambiguously stratum 1 effect (*divide, divis-ible, divis-ive*). However, *-ible* triggers voicing while *-ive* does not. This distinction cannot be captured by stratal assignment but must be tied to individual suffixes using exception features or other mechanisms besides strata.

Because, by its nature, realizational morphology is also narrowly tied to individual morphological contexts, stratal ordering theories do not lend themselves to the description of realizational morphology any more than they are suited to capturing idiosyncratic morphophonology.

A middle ground which can capture those morphophonological generalizations sought by stratal theories but also describe highly morphologically-specific phonological patterns is represented by cophonological models. These, as discussed in Section 2, associate each morphological construction (affixation, compounding, zero-derivation) with its own phonological mapping. The cophonological approach eliminates the "too many analyses" problem by using exactly the same mechanism to handle realizational morphology and morphologically conditioned phonology. A phonological alternation specific to a particular affix is included in the cophonology that is unique to that affix. Realizational morphology is accomplished by the cophonology of what might otherwise be described as phonologically null morphological constructions. Whether a construction is "null" or not, that is, whether or not it is associated with an overt affix, is in cophonology theory almost incidental. In this way cophonology theory resembles the approach of Bochner (1992), in which phonological patterns are part and parcel of the description of the rules relating words in a paradigm. Cophonology theory is not limited to enumerating idiosyncracies; as Anttila (1997a, 2002a) has demonstrated, its inheritance architecture also gives it the ability to posit meta-constructions like "word," "stem," or "stratum," with associated cophonological restrictions inherited by the member constructions, to capture generalizations holding across all constituents of a certain type.

6 When Phonology and Morphology Diverge: Nonparallelism Between Phonological and Morphological Structure

The domains of word-internal phonological patterns are generally coextensive with the morphological sub-constituents of a word; for this reason, phonology provides strong evidence about the morphological structure of a word. However,

there can be mismatches, that is, situations in which phonological domains are not matched with morphological sub-constituents. In some cases the phonological domain – prosodic root, or stem, or word – is a sub-portion of a word (see e.g. Booij 1984; Sproat 1986; Inkelas 1990; Booij and Lieber 1993, among many others). Three situations stand out in this regard: compounding, the distinction between cohering and non-cohering affixes, and reduplication of an internal prosodic stem.

The literature on the phonology of compounding constructions has often drawn attention to a distinction between compounds that behave phonologically like one word and those that behave phonologically like two words. In the 1980s this difference was attributed to prosodic structure which is loosely related to but exists independently of morphological and syntactic structure. Nespor and Vogel (1986) proposed that while Greek compounds form a single prosodic word and thereby receive one stress, for example, *kúkla* 'doll' + *spíti* 'house' → [*kuklóspito*] 'doll's house' (p. 112), the members of Hungarian compounds form separate prosodic words and retain their own lexical stresses: [könyv]_ [tár]_ 'book collection' (p. 123). In Malayalam, simple sub(ordinate) compounds, with head-modifier semantics, form a single domain for accentuation, whereas the members of simple co(ordinate) compounds, with coordination semantics, form separate domains for accentuation. Sproat (1986) and Inkelas (1990) proposed that this difference in behavior could be attributed to different prosodic structure, though Mohanan (1995) later countered this argument with evidence from complex compounds with three or more members. In a detailed study of Indonesian, Cohn (1989) documents a stress difference between two constructions that concatenate stems. Head-modifier compounds impose stress reduction on one member (*polúsi* 'pollution' + *udára* 'air' = *polùsi udára* 'air pollution', p. 188), suggesting that they are competing for prominence within a single phonological word, while total reduplication constructions maintain two equal stresses (*minúman* 'drink', *minúman-minúman* 'drinks', p. 184). Cohn attributes the latter pattern to the fact that total reduplication consists of two prosodic words. Itô and Mester (1996a) point to a similar distinction in Japanese, in which stem-stem compounds form one prosodic domain, word-word compounds form two domains, and stem-word compounds "type-shift," by means of a principle of Prosodic Homogeneity (p. 38), to pattern like word-word compounds (see also Han 1994 on Korean).

Perhaps even more interesting than prosodic differences across types of compounds are comparable differences in affixed words. Booij (1984) was one of the first to highlight the distinction between "cohering" and "non-cohering" affixes and to model the distinction using prosodic structure: cohering affixes form a single prosodic word with the base of affixation, while non-cohering affixes form a separate prosodic domain. In Dutch, for example, suffixes are either cohering, meaning they syllabify with and join into a prosodic domain with the stem they combine with, or non-cohering, meaning they create a separate prosodic domain. Nonnative suffixes in Dutch are all of the non-cohering type. The difference between the two types of suffix is illustrated with this minimal pair of suffixes both of which are equivalent to English "-ish": *rood-achtig* [roːt.ɑx.təx] and *rod-ig*

[roɪ.dəx] 'reddish' (Booij 1984 :152). As Booij (1984, 2001, 2002: 172) observes, non-cohering Dutch suffixes like *-achtig* allow deletion under identity: if two words ending in *-achtig* are coordinated, the suffix is omissible from the first conjunct: *storm-achtig en regen-achtig* 'stormy and rainy' can also be realized as *storm en regen-achtig* (1984: 151). Deletion under identity is also found in compounding constructions: *wespen-steken en bije-steken ~ wespen en bije-steken* 'wasp and bee stings' (1984: 146). Non-cohering affixes and members of compounds form individual prosodic words, explaining their parallel behavior. By contrast, cohering suffixes like *-ig* cannot be omitted under identity: *blau-ig en rod-ig* 'blueish and reddish', but **blau en rod-ig* (1984: 149). On deletion under morphological identity in other languages, see for example, Vigário and Frota (2002), Orgun (1996).

Evidence for the accessibility to "later" processes of word-internal prosodic stems is found in reduplication. In a number of cases, a late morphological process of reduplication targets the root, even if the root has already undergone significant affixation. Aronoff (1988) refers to these as "head operations," and Booij and Lieber (1993) propose that they involve reference to a prosodic stem, which corresponds closely if not exactly to the morphological root. Inkelas and Zoll (2005) cite the example of Chumash, which has what Applegate (1972: 383–384) characterizes as a very late process of reduplication, conferring the meaning of a repetitive, distributive, intensive, or continuative force. Chumash reduplication targets a sub-constituent of the word which Inkelas and Zoll term the prosodic stem. The prosodic stem always contains the root, along with any preceding prefixes of the type Applegate (1972) identifies as reduplicating, and which Inkelas and Zoll (2005) analyze, in Booij's terms, as cohering. For example, the root-adjacent prefix in *k-sili-*{*pil-*<u>*wayan*</u>} 'I want to swing', in which curly brackets demarcate the prosodic stem and the root is underlined, is cohering and participates in reduplication: *ksili*{*piw-piwayan*} (Applegate 1972: 387). By contrast, the prefixes in *s-am-ti-*{<u>*lok'in*</u>} 'they cut it off' are non-cohering and do not reduplicate: *samti*{*lok-lok'in*} (Applegate 1972: 387). Evidence that what reduplicates is a prosodic stem, occupied by the root and joined by cohering prefixes, is that the prosodic stem is subject to a typical stem-shape constraint; it must be consonant-initial. Onset consonants are not required of Chumash roots or prefixes, many of which are vowel-initial. But prosodic stems must be consonant-initial. As a result, even an otherwise non-cohering prefix will contribute its final consonant to a following prosodic stem, as shown by reduplicated forms such as *s-iy-ak*{*t-aqu-*<u>*s*</u>*mon*} → *siyak*{*taq-taqusmon*} 'they come to gather it' (Applegate 1972: 388). Parallel phenomena, documented in Inkelas and Zoll (2005), occur in Tagalog (see also Booij and Lieber 1993) and Eastern Kadazan (Hurlbut 1988). An alternative analysis of the Chumash and Tagalog phenomena is offered within Base-Reduplicant Correspondence Theory (BRCT) by McCarthy and Prince (1995), who propose that the reduplicant is not infixing but is instead prefixed directly to the material that is copied. On their account, the copying of the final consonant of a prefix *preceding* the reduplicant is the result of morphological fusion between the prefix consonant and the VC reduplicant and "back-copying" of the result to the base of reduplication: *s-iš-*RED*-expeč* → $s\text{-}i\text{-}šex_{RED}\text{-}šexpeč_{BASE}$ (with backcopying of the *iš* -final *š* to the base).

Inkelas and Zoll (2005) argue against this account of Chumash, in particular, on language-internal morphological grounds; McCarthy and Prince (1995) have, however, identified other apparent cases of backcopying in other languages, and backcopying in general remains a viable analysis within BRCT.

Returning to mismatches between morphological constituent structure and prosodic structure, there is also strong evidence that word-sized prosodic domains can include material outside of the morphological or lexical word, clitics being the most obvious example. It has been widely argued that clitics are phonologically defective syntactic terminal elements, having to join with another (non-clitic) syntactic terminal element to form a single prosodic word (e.g. Inkelas 1990; Halpern 1992; Booij 1996). A question of considerable current interest is whether prosodic word structure can be recursive; see Peperkamp (1996), Itô and Mester (2003a), and Kabak and Reviathidou (2009), among others.

7 Summary

The phonology-morphology interaction sheds light on word-internal structure and on the ability for relatively unnatural phonological alternations to be productive, at least within a given morphological niche. Both realizational morphology and morphologically conditioned phonology operate in the same domains and manipulate the same structural elements. The many related phenomena constituting the phonology-morphology interface are central to word-formation in virtually all languages, and must therefore be taken seriously by morphologists and phonologists, especially those seeking to reduce synchronic morphological patterns to syntax, or synchronic phonological patterns to universal phonetic motivations.

NOTES

1 This is the approximate generalization, as stated by Kenstowicz; the actual picture is more detailed, in ways not material to the point made here. See for example, Butt and Benjamin 2008.

2 Vaux (p. 252) analyzes the definite suffix as underlyingly /-n/ and attributes the schwa allomorph to rules of epenthesis and consonant deletion; however, as the n~ə alternation is specific to the definite, and most researchers would probably classify this as suppletive allomorphy.