In Paul de Lacy (ed.), The Cambridge Handbook of Phonology, CUP.

16

Dissimilation in grammar and the lexicon

John D. Alderete
Stefan A. Frisch

16.1 Introduction

Dissimilation is the systematic avoidance of two similar sound structures in relatively close proximity to each other. It is exhibited in static generalizations over the lexicon, where combinations of similar sounds are systematically avoided in lexical items, like the avoidance of two homorganic consonants in Arabic roots (Greenberg 1950; McCarthy 1994). Dissimilation is also observed in phonological processes in which the target and trigger become less alike phonologically. In Tashlhiyt Berber, for example, two primary labial consonants in the same derived stem trigger a process of delabialization: /m-kaddab/ → [n-kaddab] 'consider a liar (reciprocal)' (Elmedlaoui 1992, Jebbour 1985).

Dissimilation has been an important empirical testing ground for many of the central research paradigms in modern linguistics. For example, dissimilatory phenomena have been crucial to the development of theories of feature geometry and feature specification in autosegmental phonology (McCarthy 1986, 1988, Padgett 1995, Yip 1989b). As the results emerging from this research were incorporated into constraint-based theories of phonology like Optimality Theory (Prince & Smolensky 2004), dissimilation became an important problem in the study of phonological markedness and constraint composition (Alderete 1997, Itô & Mester 2003). In a different line of research, dissimilation has been argued to have its seeds in the phonetics of sound change, restricted by the same vocal tract constraints involved in speech production and perception (Ohala 1981 et seq.). Finally, dissimilation has been shown to be statistically systematic when observed over lexical items, raising a host of questions about the interface between categorical and statistical information in grammatical models (Pierrehumbert 1993b, Frisch et al. 2004). The discussion below summarizes the results emerging from these separate theoretical enterprises, and it also strives to identify problems for future research that may require new models of grammar that mix symbolic, statistical, and phonetic information.

This chapter is structured as follows. Section 16.2 provides a basic toolbox for describing the ways in which dissimilation differs cross-linguistically. Section 16.3 introduces the hypercorrective theory of dissimilatory sound change proposed in Ohala (1981) and uses some of the background from Section 16.2 to assess this theory. In Section 16.4, generative approaches to dissimilation are discussed that make use of the principles of autosegmental phonology and Optimality Theory. Section 16.5 focuses on statistical patterns of dissimilation in the lexicon and lays out the analysis of these statistical patterns based on the scalar and quantitative property of similarity. Finally, the conclusion ties together some of the open issues arising from these discussions and clarifies some questions for future research.

16.2 Parameters of dissimilation

What are the 'parameters' of dissimilation? Which features are typically involved and what types of conditions may be placed on dissimilation rules? Direct answers to these questions are limited by the fact that, to date, there has not been a controlled study of the cross-linguistic variation in dissimilation. On the other hand, the wealth of examples available in the literature set the stage for further empirical investigation and assessment of scientific hypotheses.

One of the main ways dissimilation may differ cross-linguistically is whether it is observed in dynamic alternations or as static generalizations over the lexicon. An example of the former type is illustrated by the data from Tashlhiyt Berber below (Elmedlaoui 1992, Jebbour 1985, Selkirk 1993). Derived stems in Berber may only have one primary labial consonant, that is, one consonant from the set [b f m]. When a derivational prefix, like the reciprocal prefix /m(m)-/, attaches to a root that also contains a primary labial consonant, a process of labial dissimilation is triggered, causing the prefix to change to [n], as shown on the right. The left columns contain the surface forms, and the right columns the roots (marked by a √).

(1) *Primary labial dissimilation in Tashlhiyt Berber*

(a)	m-xazar	√xzr 'scowl'	(b) n-fara	√fra 'disentangle'
	m-saggal	√sggl 'look for'	n-ħaʃʃam	√ħʃʃm 'be shy'
	m-!ʃawar	√!ʃawr 'ask advice'	n-xalaf	√xalf 'place crosswise'
	mm-ʒla	√ʒla 'lose'	n-kaddab	√kddb 'consider a liar'

Dissimilation is also exhibited in static generalizations over the lexicon. That is, statistical analysis of the frequency of segment types in lexical items may reveal significant over- or under-representation of these types. Dissimilation in the lexicon involves significant under-representation of two similar segments in lexical items. A well-known example of this type is

Modern Arabic, where homorganic consonant pairs are significantly under-represented in roots (Greenberg 1950, McCarthy 1994). Table (2), taken from Frisch et al. (2004), illustrates the relative frequencies of adjacent consonants in Arabic verb roots (n = 2,674), sorted by place and manner of articulation classes. Each cell gives a measure of over- or under-representation as a value for O/E, or observed/expected (Pierrehumbert 1993b), where observed is the number of occurrences and expected is the number of occurrences that would be expected if the consonants combined at random. A value less than one therefore indicates that there are fewer observed consonant pairs than would be expected if they combined at random, as shown in the shaded portions below.

(2) *Co-occurrence of adjacent consonant pairs in Arabic*

	Lab	Cor Stop	Cor Fric	Dorsal	Uvular	Phar	Cor Son
Labial [b f m]	0.00	1.37	1.31	1.15	1.35	1.17	1.18
Cor Stop [t d tˤ dˤ]		0.14	0.52	0.80	1.43	1.25	1.23
Cor Fric [θ ð s z sˤ zˤ ʃ]			0.04	1.16	1.41	1.26	1.21
Dorsal [k g q]				0.02	0.07	1.04	1.48
Uvular [χ ʁ]					0.00	0.07	1.39
Pharyngeal [ħ ʕ h ʔ]						0.06	1.26
Cor Son [l r n]							0.06

The patterns of over- and under-representation above indicate that there is a systematic avoidance, though not an absolute avoidance, of adjacent homorganic consonants. For example, roots containing two labials (O/E = 0.00) or two dorsals (0.02) in a row are significantly under-represented, contrasting sharply with roots containing one labial and dorsal (1.15). These patterns also show the effect of manner features on coronals, as coronals co-occur with a relatively high frequency if they are not in the same manner classes, stop, fricative, and sonorant, but a sequence of two coronals within one of these classes is under-represented. Avoidance of two similar sounds may thus be observed in the statistical analysis of the frequency of segment co-occurrence in the lexicon.

In terms of the range of features referred to by dissimilation rules, there is no specific typological evidence that any particular feature is not active in dissimilation. However, a review of the commonly cited examples suggests that certain features are more frequent than others. In Suzuki (1998), 53 examples of dissimilation were amassed from some of the principal works on the subject (especially Odden 1994, Ohala 1981, Padgett 1995, Walsh Dickey 1997, and Yip 1989b). The list in (3) organizes these examples, from both alternations and lexical dissimilation, and sorts them by the principal feature or feature class that triggers dissimilation.

(3) *Examples from Suzuki 1998 (n = 53, of 21 features; 2.5 expected frequency)*

(a) Place, total ______ 15
- [place] (=homorganic) ______ 5
- [labial] ______ 7
- [coronal] ______ 2
- [dorsal] ______ 0
- [pharyngeal] ______ 1

(b) Laryngeal, total ______ 8
- [±voice] ______ 3
- [±constricted glottis] ______ 1
- [±spread glottis] ______ 1
- [tone] (H and L) ______ 3

(c) Liquids, total ______ 10
- [liquid] ______ 1
- [±lateral] ______ 4
- [rhotic]/[retroflex] ______ 5

(d) Major class, total ______ 0
- [±consonant] ______ 0
- [±sonorant] ______ 0
- [±approximant] ______ 0

(e) Vowel features, total ______ 8
- [±high] ______ 2
- [±low] ______ 4
- [±back] ______ 2

(f) Manner, total ______ 3
- [±continuant] ______ 2
- [±nasal] ______ 1
- [±strident] ______ 0

(g) Other, total ______ 9
- NC cluster ______ 3
- Length ______ 6

As a convenience-based sample, the dataset is not suitable for statistical analysis. However, the patterns in the first column (3a–c) suggest that dissimilation of place features, laryngeals (including tone), and liquids may have a greater than chance occurrence. When compared to the non-existent cases of dissimilation of the major class features, and the small number of cases with manner features, these cases stand out. Furthermore, there seem to be certain types of place restrictions that are more common than others, like those prohibiting a double occurrence of labials or any two homorganic consonants (i.e., identical [place] specifications).

With the exception of tone, the term dissimilation is more often applied to subsegmental attributes of segments, like the features listed above. However, the mode of analysis 'avoid similar elements' has also been applied with some success to prosodic features, such as vowel and consonant length, as well as complex phonological structures like pre-nasalized stops and nasal-obstruent clusters. For example, Japanese loans from English have geminate stops after lax vowels, as in *rakkii* 'lucky', but double geminate loans with two such lax vowel-stop clusters only allow a single geminate: *ma*[p]*etto* 'Muppet', **ma*[pp]*etto* (Iwai 1989, Itô & Mester 2003). Thus, two similar elements are avoided, though the prohibited structures are not subsegmental features of individual segments.

The validity of the dissimilation analysis for these non-segmental elements, however, is sometimes confounded by the availability of alternative analyses. For example, the resolution of sequences of NC clusters found in many aboriginal languages of Australia (e.g. Yinjibarndi, Gurindji, and Gooniyandi) seems to have a more natural analysis in terms of nasal spreading, because this resolution is blocked by intervening oral stops (see Odden 1994 for relevant examples and discussion). Likewise, the proposal to capture Clash Avoidance effects in stress systems as a consequence of a dissimilatory constraint against two adjacent stressed syllables (Yip 1988) must compete with the plausible alternative in which stress is the realization of the head of a metrical stress foot, which is rhythmically distributed because of robustly supported binarity requirements on stress feet.

Finally, there is a set of conditions on dissimilation that accounts for known cross-linguistic variation, defined and discussed below.

(4) *Conditions on dissimilation*
  - (a) <u>Root adjacency</u>: the target and trigger must be adjacent segments, i.e., adjacent root nodes
  - (b) <u>Syllable adjacency</u>: the target and trigger must be contained in adjacent syllables
  - (c) <u>Domains</u>: the target and trigger must fall within a specified domain, e.g., a morphological domain (root, stem, word, morpheme) or prosodic domain (e.g., syllable, foot, 'rime')
  - (d) <u>'No intervener'</u>: for target segment *x* and trigger *y*, there must be no segment *z*, of a specified type, that intervenes between *x* and *y*

One class of conditions involves requirements on the locality of the target and trigger of dissimilation, conditions that stem from a long line of research on locality in phonology (Archangeli & Pulleyblank 1987, Steriade 1987, Odden 1994). Root adjacency (4a) requires target and trigger to be adjacent segments, as illustrated by a second dissimilatory pattern in Berber. In the Tiznit variety of Tashlhiyt Berber reported in Jebbour (1985), rounded velar consonants unround when immediately preceded by a primary labial consonant, e.g. [gʷ]*rd* 'tame (preterit)', cf. *am*-[g]*rad, im*-[g]*rad* 'tame (agentive singular, plural)'; velars thus lose their rounding and become less like the adjacent labial segments. Locality requirements can also be somewhat looser, requiring only that target and trigger appear in adjacent syllables (4b), as exemplified by Dahl's Law in many Bantu languages. This process changes /k/ to [g] when another voiceless obstruent appears in an adjacent syllable, as shown by the following data from Kikuria (Odden 1994): /oko-téma/ → *o*[g]*o-téma* 'to hit', cf. *o*[k]*o-bára* 'to count' (no voiceless obstruent) and *o*[k]*o-mɛ́nɛ́nɛkánya* 'to make each other shine' (no voiceless obstruent in adjacent syllable).

While locality requirements such as these are rather common, perhaps due in part to their history (see Section 16.3), it is not uncommon for the target and trigger of dissimilation to be limited to some domain that is defined by a coherent linguistic unit. For example, Berber labial dissimilation described above in (1) is limited to the morphological stem, so the target and trigger must fall within this morphological domain. Other morphological domains, e.g., verb roots in Arabic (2), and even prosodic domains like syllables or feet may be placed as domain conditions on the application of dissimilation, as has been argued for syllable-bound glottal stop dissimilation in Seri (Marlett & Stemberger 1983, Yip 1988).

While the patterns of dissimilation in the lexicon in Arabic are limited to the root, there is a second sense in which locality is relevant to these patterns. It has long been known that the dissimilation patterns are stronger for consonant pairs that are adjacent in the root (e.g., forms with coronal stops like *dtC, O/E = 0.14) than for consonant pairs that are non-adjacent in the root (e.g., forms with coronal stops like *dCt, O/E = 0.38) (Greenberg 1950, McCarthy 1994, Pierrehumbert 1993b). In other words, the extent to which similar sounds are avoided in the roots of Arabic depends on their proximity to one another.

Finally, the parameters of dissimilation seem to also require a condition on the material intervening between target and trigger, namely that this intervening material does not contain a segment of a specified type. Latin lateral dissimilation is often described in this way, where a sequence of two laterals triggers dissimilation, /lun-alis/ → *lun-a*[r]*is* 'lunar', but this process is blocked if target and trigger are separated by a rhotic [r], e.g., /flor-alis/→ *flor-a*[l]*is* 'floral' (Walsh Dickey 1997).

In summary, cross-linguistic variation in dissimilation may be described using the distinction between dynamic processes and static lexical dissimilation, the specific features involved, and a set of conditions on the application of dissimilation. These parameters of dissimilation provide useful background for the understanding and assessment of the approaches to dissimilation discussed below.

16.3 Diachronic seeds of dissimilation: hypercorrection

Considerable attention has been paid in recent years to how phonological processes come about historically, often with careful examination of the phonetic structures that give rise to synchronic phonological patterns (Hyman 1976, Ohala 1983a, Blevins 2004). One approach to dissimilation has been developed within this program that attributes dissimilation rules to hypercorrection, a kind of mental inversion of a perceived phonetic co-articulation (Ohala 1981, 1992, 1993). Rich in predictions, this theory clarifies a number of questions for future work, some of which have been taken up in the lines of research discussed below.

Ohala's hypercorrection theory is illustrated below with an example from Slavic (5). In this sound change, the low front vowel [a] is shifted to back [ɑ] after palatal and palatalized consonants, e.g., *stoj + /aː/ → *stoj*[ɑː] 'stand'. This change is a case of dissimilation for the feature

[±back]: palatals are [−back], and so are front vowels, so *ja > jɑ avoids two adjacent segments that are both [−back].

(5) *Scenario for hypercorrective sound change*

Speaker				*Listener*
/stoj + a:/				/stoj + ɑ:/
⇓				⇑
faithfully produces				listener corrects perceived fronting
⇓				⇑
[stoj-a:]	⇒	heard as	⇒	[stoj-a:]

According to the hypercorrection theory, the listener in the scenario above falsely attributes the frontness of [a] to the neighboring palatal, a plausible assumption, given the ubiquity of C-V co-articulation. Believing that [j-a] is co-articulated, the listener posits a mental representation that inverts the [−back] specification, and if the same action is copied by other listeners, a regular sound change may develop.

The attraction of the hypercorrective theory is that it draws on a number of parallels between known patterns of co-articulation in speech production and the observed patterns of dissimilatory sound change. These parallels make for a highly constrained theory of the historical source for dissimilation, as explicated in the list of predictions given below.

(6) *Predictions of hypercorrective theory*

(a) Locality: since hypercorrective dissimilation is due to perceived co-articulation, the target and trigger must be adjacent segments, unless the relevant phonetic feature is mediated by an intervening segment.

(b) Which features dissimilate: 'stretched out' features that have longer temporal intervals are more likely to dissimilate than others because they provide better cues that the listener can attribute to a false co-articulation.

Predicted to dissimilate	*Predicted not to dissimilate*
labialization, uvularization, pharyngealization, palatalization, retroflexion, place, glottalization, aspiration, laterals	continuancy (stop/fricative) complex segmenthood voice

(c) Persistent triggers: because hypercorrection depends on the existence of a triggering element, the triggering segment is never lost at the same time as a dissimilatory sound change; thus, e.g. *$b^h$and^h > *ban* is predicted not to exist as a single sound change

(d) No new segments: hypercorrective dissimilation involves a normalization, i.e., a reverting back to an assumed input; if the input is constructed from a common stock of sounds, no new segments can result from dissimilation.

How do these predictions measure up? Beyond Ohala (1981), we know of no works that have investigated these predictions systematically in dissimilatory sound change. However, it seems plausible to test some of them, albeit indirectly, using data from synchronic patterns, under the assumption that the sound changes that fall under the scope of this theory will give rise to lexical and morpho-phonemic patterns of dissimilation, and that many of these patterns inherit the primary characteristics of the original sound changes. These assumptions seem plausible, since lexical dissimilation is quite naturally analyzed as the accumulation of regular sound changes affecting lexical items, and an extension of the resulting lexical patterns is a likely source of synchronic alternations.

It seems fair, therefore, to consider the dataset in (3) in connection with the prediction that only features with long temporal intervals may dissimilate. In this light, the patterns in (3) are in large part consistent with this prediction: dissimilation of place, tone and laryngeal features are quite common, which clearly have a long temporal interval, as does retroflexion in liquid dissimilation. Laterals probably have prominent cues too, because they have long F2 and F3 transitions, which accounts for the large number of cases of dissimilation of [±lateral]. One sticking point is dissimilation of [voice], supported by three examples in (3), though at least one of these, Thurneysen's Law, has been reanalyzed with reference to higher-level prosodic structure (Ohala 1981; see also Ohala 1993).

The locality question is also interesting in connection with the observed conditions on dissimilation from Section 16.2. Since co-articulation has the greatest effect on adjacent segments, the prediction is that hypercorrective inversion of a perceived co-articulation will also be of local elements, relating naturally to patterns of root-adjacent dissimilation. Dissimilation of vowels in adjacent syllables is also a possibility, since it is known that vowels can influence each other across consonants, at least in VCV structures (Öhman 1966). As for dissimilation of consonant attributes across vowels, this too seems to be possible, because of the robust phonetic evidence that consonants may be co-articulated with neighboring vowels. For example, the F2 and F3 formant transitions that constitute the main cues for consonantal place of articulation are known to persist into a significant portion of neighboring vowels. If attested, this C-V co-articulation effectively makes the dissimilation between local segments because the vowel mediates the dissimilating feature (see Ohala 1981 for explicit discussion and examples). This analysis does not seem to work, however, in cases where the target and trigger of dissimilation are separated by more than a consonant or vowel. Labial dissimilation in Berber, for example, occurs over two heterosyllabic vowels and consonants, e.g. /m-kaddab/ → [n-kaddab] 'consider a liar (reciprocal)', making inversion of a phonetic assimilation highly implausible. Such long-distance dissimilations are not uncommon, but they are not predicted by the hypercorrective theory.

Another problem is the apparent asymmetries in the perceptual response to co-articulation. The crux of the hypercorrective theory is that when a surface form contains two neighboring elements that share a feature compatible with co-articulation, the listener may interpret this as the result of co-articulation and correct the structure. Consequently, the theory predicts, as emphasized in (Ohala 1981 et seq.), that there are two plausible responses to these forms: (i) hypercorrection, leading possibly to dissimilation, or (ii) hypocorrection, where the listener does nothing to the heard form and adopts the co-articulated structure at face value. Ohala's work enumerates dozens of very plausible examples of both types, but certain cases seem to favor one type of response over another. For example, co-articulation in nasal-obstruent clusters is extremely common cross-linguistically, and has led to countless examples of phonological nasal-place assimilation, i.e., hypocorrection caused by co-articulation of place features. But we know of no dissimilatory examples of this type, so hypercorrection seems impossible. A similar argument could be made for C-C [voice] assimilation, though this is confounded by the question of whether [voice] dissimilation arises from hypercorrective sound change. Conversely, liquid dissimilation is extremely common, second only to place dissimilation in (3), but liquid assimilations are vanishingly rare (though Palaun liquid assimilation may be a counterexample; see Josephs 1975). Why is hypercorrection favored here? It seems therefore that the conditions on the perceptual responses to co-articulation must be different in these examples, and future hypothesis-testing may profit from linking these conditions to whether or not the listener actively responds to co-articulation with hypercorrection.

While there is a relatively good fit between many of the above predictions and observed dissimilatory sound changes, the limits of this theory suggest that there may be additional causes of dissimilation. One additional cause may follow from the difficulty in producing speech that contains repeated items (Dell et al. 1997). For example, tongue twisters containing repeated similar onset consonants (e.g. *sit zap zoo sip*) are more difficult to produce than those that do not contain repeated onsets (e.g. *sit shop zoo tip*). Tongue twisters containing repetition are analogous to polysyllabic words containing pairs of similar segments, i.e., the input to dissimilation rules. Thus, there may be an additional functional motivation for dissimilation in the difficulty in processing words containing repeated segments during speech production (Berg 1998, Frisch 2004).

Furthermore, though the hypercorrective theory is a predictive theory in terms of the types of dissimilatory patterns that are likely to be found, it is not a theory of the representation of dissimilation in synchronic grammars. Thus, the hypercorrective theory may explain sound changes that result in systematic changes to lexical entries, but when morpho-phonemic alternations are involved, formal mechanisms are needed to model dynamic processes of dissimilation. This limitation applies to any functional explanation of dissimilation. Therefore, we next consider generative approaches to synchronic dissimilation, as considerable progress has been made in modeling dissimilation as a dynamic process.

16.4 Generative approaches to dissimilation

Advancements in autosegmental and prosodic phonology in the 1980s provided many of the leading ideas behind the generative approach to dissimilation, including explicit formalisms for morpho-phonemic alternations and long-distance dissimilation. Indeed, dissimilatory phenomena constituted an important empirical testing ground for proposals concerning tier structure in phonology, feature geometry, and the nature of feature specification. In addition, these investigations of dissimilation clarified a number of problems with autosegmental approaches that were later addressed in constraint-based approaches to dissimilation like OT and probabilistic linguistics.

16.4.1 Tier phonology, the OCP, and feature specification

Generative approaches to dissimilation were born out of investigations of autosegmental tone. In this work (pioneered by Leben 1973 and Goldsmith 1976a), it is argued that many properties of tone systems can be explained with a phonological tier for tone that is distinct from the tier for units sponsoring tone. In addition, the tiers and association of tonal units to their sponsors is governed by a general set of principles and constraints. One set of conditions, dubbed the Association Conventions, constitutes an algorithm for linking up the elements on the two tiers, stipulating essentially that associations are made from left-to-right, one-to-one, until the end of the string has unassociated elements, in which case, the unassociated elements are linked with the rightmost element on the other tier (Goldsmith 1976a). These conventions for linking up structure on distinct tiers led to a parallel between tone and segmental phonology (McCarthy 1979a), illustrated in (7), which in turn engendered the autosegmental analysis of dissimilation.

(7) *Autosegmental association for tone and segmental features*

(a) Tone

σ σ σ

L H

(b) Consonants

C V C V C

s m

The shortage of tones in the LH melody in the tone example requires the last tone to link up with both the second and third syllable, which produces a low-high-high tone shape in the surface form. This analysis has the effect of systematically prohibiting words with a low-low-high tone shape, an important fact in tone systems like Mende (Leben 1973), because the Association Conventions do not allow double linking of a non-final L tone. In his study of root-and-pattern morphology in Arabic, McCarthy (1979a) found the same directionality pattern in root consonantism, accounting for Arabic roots like *s-m-m* but excluding non-existent **s-s-m* as a straightforward consequence of left-to-right association to the CV tier.

This analysis, however, depended crucially on the absence of LLH tone melodies, and for example *s-s-m* consonantal roots in Arabic; otherwise, LLH surface forms could be produced with simple one-to-one association. In autosegmental phonology, these structures are ruled out by the Obligatory Contour Principle (OCP).

(8) Obligatory Contour Principle *(Leben 1973, Goldsmith 1976a, McCarthy 1986)*
At the melodic level, adjacent identical elements are prohibited.

The OCP rules out LLH and *s-s-m* melodies because they contain two identical elements on the same tier. This insight behind the OCP, that adjacent identical elements are prohibited, has been extended to analyses of a host of dissimilation patterns, observed in both alternations (Itô & Mester 1986, Myers 1987a, Selkirk 1993, Yip 1988) and lexical distributions (Itô & Mester 1986, MacEachern 1999, McCarthy 1988, Mester 1986, Padgett 1995, Yip 1989b). The tier structures in (9) illustrate how the OCP motivates processes like long-distance labial dissimilation in Berber (1), as well as two important assumptions about feature structure and specification necessary to the analysis.

(9) *Autosegmental analysis of Berber labial dissimilation*

	Input violates OCP	Output satisfies OCP
	/ m + k a dd a b /	[N + k a dd a b]
		Ø
Dorsal	dor	dor
Coronal	cor	cor
Labial	lab lab	lab

The input contains two primary labial segments, which, because of the separation of tiers, are adjacent on the labial tier. Deletion of the [labial] feature in the output, the representation of delabialization, therefore satisfies the OCP. Crucially, the nasal prefix is not specified for place at the point at which the OCP applies; otherwise, the [coronal] feature would also violate the OCP. In cases like this, it is typically assumed that unspecified segments received a default feature specification, in this case [coronal], at a final step in the phonological derivation. In this way, appeal to the OCP in the analysis of dissimilation is often tied in explicit ways to assumptions about tier structure and feature specification, two points taken up below.

Part of the appeal of the autosegmental analysis of dissimilation is the application of the OCP to both lexical representations and alternations. When these two types of dissimilatory patterns are found in the same language, however, a problem arises, stemming more from the implementation of autosegmental ideas in derivational phonology than the content of the OCP itself. A well-known example is Lyman's Law in Japanese, a restriction against two voiced obstruents in a stem (Itô & Mester 1986; Itô et al. 1995). Lyman's Law applies both as a restriction on a lexical stratum, namely the Yamato stock of the Japanese lexicon, and to the output of phonological rules. In rule-based phonology, the static Lyman's Law amounts to a lexical redundancy rule, perhaps motivated by the OCP, acting on Yamato stems. However, this same restriction must be active in phonological derivations as well, because it blocks the application of Rendaku, a systematic morpho-phonemic process that voices stem-initial obstruents in the second member of a compound. The application and blocking of Rendaku is illustrated in (10).

(10) *Blocking of Rendaku sequential voicing by Lyman's Law*

(a)	/ori + kami/	→	[ori-gami]	'folding paper'
	/oo + sumoo/	→	[oo-zumoo]	'*sumo* tournament'
	/jama + tera/	→	[jama-dera]	'mountain temple'
	/mizu + hana/	→	[mizu-bana]	'running nose'
(b)	/kami + kaze/	→	[kami-kaze]	'divine wind'
	/ʃiro + tabi/	→	[ʃiro-tabi]	'white tabi'
	/mono + ʃizuka/	→	[mono-ʃizuka]	'tranquil'
	/maru + hadaka/	→	[maru-hadaka]	'completely naked'

The realization of the full potential of the autosegmental analysis is thus hampered by a technical problem, namely that restrictions on lexical items and constraints on the output of phonological rules are governed by different mechanisms. Though this issue is complex, and there are reasons to want to separate the analysis of lexical distributions from alternations, there is a consensus in the generative literature that this allocation of resources to two separate domains, dubbed the Duplication Problem, stands in the way of genuine explanation.

The derivational implementation of the autosegmental analysis commonplace in the 1980s, together with the necessary assumptions tied to feature specification, leads to another problem. A fundamental assumption in the autosegmental analysis is that the features that are active in dissimilation are specified, and those that are not active are not specified at the time at which the OCP is applied. It is necessary to assume, for example, [coronal] is unspecified at the derivational instant at which the OCP is applied in the Berber example illustrated above. It turns out that the assumption 'active is specified' leads to serious empirical problems, because dissimilation often requires specificational assumptions that are not consistent with other facts. For example, in the Japanese example above, it is assumed that redundant sonorant voicing is not specified in examples like *ori-gami*; otherwise, voicing in [m] would block Rendaku, just like voiced obstruents do. However, derived voicing in post-nasal obstruents also blocks Rendaku, as exemplified by examples like /ʃirooto-kaŋkae/ → *ʃirooto-*[k]*aŋgae* 'layman's idea'. A [voice] specification is therefore necessary prior to the application of Rendaku in order to give a natural analysis of voice assimilation in nasal-obstruent clusters, an assumption that is inconsistent with [voice] specification in *ori-gami* (see Itô et al. 1995 for explicit derivations and discussion, and Alderete 1997 for a parallel example in Berber). This problem, like the Duplication Problem, stems directly from the assumption that phonological generalizations are expressed as rules that are serialized in a phonological derivation.

16.4.2 Cumulative markedness in Optimality Theory

One of the core assumptions of Optimality Theory (OT – Prince & Smolensky 2004) is that phonological activity is driven by markedness (see Smolensky 1995, Rice (Ch.4) for cogent argumentation). The motivation for phonological processes in OT is to satisfy markedness constraints, a family of well-formedness constraints that prohibit cross-linguistically marked structure. Building on this insight, Alderete (1997) and Itô & Mester (2003) construct an account of dissimilatory phenomena that derives from the cumulative effects of markedness constraints. In this approach, complex markedness constraints are generated from simple ones through the operation of local self-conjunction (Smolensky 1995). The effect of this conjunction is a set of OCP-like markedness constraints that specifically ban multiple instances of marked structure. This type of analysis gives a natural account to many types of dissimilation processes, because they specifically prohibit multiple instances of marked structures. Lyman's Law is a case in point. Voiced obstruents are cross-linguistically marked. This is supported by the implicational relation in segment inventories where the presence of a voiced obstruent series requires a voiceless obstruent series, as well as abundant evidence from alternations, like coda devoicing of obstruents. These facts entail a constraint against marked obstruent voicing, *VoicedObstruent. In cumulative markedness theory, this constraint can be doubled via local self-conjunction to produce a constraint that specifically prohibits two voiced obstruents (i.e., *VoicedObstr^2, the exact restriction in Lyman's Law). This result is illustrated in a standard OT tableau (11).

(11) *Lexical dissimilation: Lyman's Law in Japanese*

/D ... D/	*VoicedObstr^2	Faithfulness	*VoicedObstr
☞ T ... D		*	*
D ... D	*!		**

Importantly, the same self-conjoined markedness constraint has a role in dynamic alternations, blocking the effects of a regular pattern of sequential voicing, Rendaku, as illustrated in (12).

(12) *Dynamic effects: blocking of Rendaku sequential voicing*

(a)

/kami+kaze/	*VOICEDOBSTR²	RENDAKU	FAITHFULNESS
☞ kami-[k]aze		*	
kami-[g]aze	*!		*

(b)

/ori+kami/	*VOICEDOBSTR²	RENDAKU	FAITHFULNESS
☞ ori-[g]ami			*
ori-[k]ami		*!	

The analysis of the inventory limitation on more than one voiced obstruent is unified with the blocking effect in dynamic alternations — they are due to the same constraint, applied consistently to pick the winning output form.

The activity of obstruent voicing is also evident in stems with derived obstruent voicing, the context that leads to the ordering paradox for the derivational analysis discussed above. In the OT analysis, however, the blocking effect is simply a matter of markedness, not the specific instant at which [voice] is specified. With an independently motivated constraint requiring voicing in post-nasal obstruents (Pater 1999) ranked above RENDAKU, the blocking effect is explained with the same ranking employed above (13).

(13) *Phonological activity of derived obstruent voicing in Japanese*

/ʃirooto-kaŋkae/	*VOICEDOBSTR²	*NC̥	RENDAKU	FAITHFULNESS
☞ ʃirooto-[k]aŋgae			*	*
ʃirooto-[g]aŋgae	*!			**
ʃirooto-[g]aŋkae		*!		*

The solution to both the Duplication Problem and the ordering paradoxes with feature specification therefore draws directly on one of the basic premises of OT, namely that constraints apply in parallel to output forms.

16.5 Probabilistic approaches to dissimilation

The generative approaches to dissimilation have focused on pin-pointing specific features or feature combinations that trigger dissimilation. An alternative approach has also been proposed in which dissimilation is grounded in the scalar and quantitative property of similarity (Pierrehumbert 1993b, Frisch et al. 2004). In this alternative approach, patterns of dissimilation have been analyzed statistically in an attempt to demonstrate that dissimilation is gradient and predictable from probabilistic functions that refer to categorical targets only as endpoints on a continuum. Statistical studies of dissimilation have focused on phonotactic patterns in the lexical items of particular languages. Such patterns have been documented in a wide variety of languages, including Arabic (see Section 16.2), English (Berkley 1994), Russian (Padgett 1995), Italian, and Thai (Frisch et al. 2004); in fact every language that has been studied statistically has revealed such patterns. This section will review the methods and findings that are representative of these studies, and present an argument that theories of grammar must account for gradient phonological patterns like that found in dissimilation.

16.5.1 Probability in the lexicon

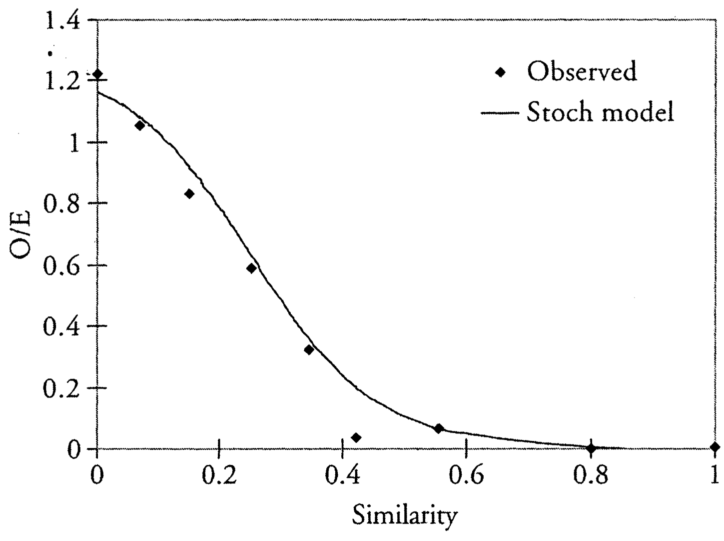
It is well documented that speakers of a language are sensitive to the statistical distribution of phonological and morphological forms in the language (e.g., Frisch et al. 2000, Hay et al. 2004, Kessler & Treiman 1997, Zuraw 2000). In other words, speakers can make judgements not just about the possible and impossible forms in a language, but also differentiate degrees of acceptability and wordlikeness within the groups of acceptable and unacceptable forms. In general, more wordlike forms provide a closer match to the statistical patterns of the language than less wordlike forms (though see Crosswhite et al. 2003, Moreton 2002 for potential counterexamples).

So far, two types of statistical pattern in phonotactics have been examined and shown to have been acquired by experiment participants: constituent probabilities and constituent combination probabilities (Frisch et al. 2000). Constituent probabilities refer to the distribution of segments within the phonological constituents of words, such as syllable onsets and rimes. Constituent combination probabilities refer to the probability of the co-occurrence of two constituents, relative to their chance rate of occurrence. The product of constituent probabilities for two (or more) different constituents in a word provides a statistical estimate of the likelihood that these two constituents would be found together at random (or by chance). If the observed constituent combination probability is significantly below the probability of co-occurrence by chance, then there is statistical evidence in the lexicon for a phonotactic constraint between the constituents. This is found in cases of dissimilation in the lexicon. Similar consonant pairs that are subject to a dissimilatory constraint are found together less frequently than would be expected by chance, like lexical dissimilation in Arabic. The fact that knowledge of these statistical constraints has been acquired by native speakers is established by experiments asking for wordlikeness or acceptability judgments for novel forms containing combinations that have different frequencies of occurrence (Frisch et al. 2000, Frisch and Zawaydeh 2001).

16.5.2 Probabilistic dissimilation in the lexicon

Statistical analyses of lexical patterns of dissimilation have revealed that similarity between consonants along any dimension is avoided. Thus, the pattern appears to be one of dissimilation in its truest sense, dubbed 'similarity avoidance' by Frisch et al. (2004). Consonants that share place of articulation only, but differ in manner and voicing (e.g., [t] and [n]) have the weakest restriction. Consonants that share major place and manner features (e.g., [t] and [d] or [n] and [r]) are much more strongly restricted. Figure (14) shows the co-occurrence of adjacent consonant pairs in Arabic, based on the data in (2), but as a function of similarity. For these data, similarity was computed using a metric of shared natural classes between segments (Frisch et al. 2004) and groups of consonants were aggregated together into similarity intervals of 0.1. In this figure, similarity 0 consonant pairs (non-homorganic) and similarity 1 consonant pairs (identical) are presented in distinct categories (note also that there happened to be no consonant pairs with similarity that fell into the range 0.6–0.7 or 0.9–1). Figure (14) also shows a curve fit to the Arabic data based on a stochastic model of co-occurrence where rate of occurrence (O/E) is a decreasing function of similarity.

(14) *Consonant co-occurrence in Arabic by similarity*



In addition to the quantitative pattern presented above, it can also be argued that co-occurrence is similarity based by considering co-occurrence in specific natural classes that share or do not share features. The co-occurrence restrictions in Arabic have been shown to be sensitive to secondary place of articulation features (Frisch et al. 2004), the manner of articulation of the consonants (McCarthy 1994, Padgett 1995), and the voicing of the consonants (Frisch et al. 1997). In short, any co-occurrence of features appears to influence the probabilistic likelihood of co-occurrence. Thus, it has been proposed that the categorical co-occurrence restrictions in the Arabic lexicon are merely the extreme end of a range of statistical co-occurrence restrictions. The effect in Arabic is strongest along the place of articulation dimension, weaker for manner of articulation, and weakest of all for voicing. It has also been argued that the number of segments that intervene between the consonants influences co-occurrence probabilistically (Berkley 1994, Buckley 1997, Pierrehumbert 1993b).

These sorts of statistical observations were addressed by generative models of Arabic consonant co-occurrence. In Padgett (1995), an autosegmental account is proposed in which the OCP is sensitive to features that are subsidiary to place of articulation features in feature geometry, like the manner features [±continuant] and [±sonorant] for Arabic and Russian. While Padgett's account correctly identifies the classes over which co-occurrence constraints are applied, the generative implementation does not address the statistical nature of these patterns. In particular, Frisch et al. (2004) argue that these patterns are best explained probabilistically, because solutions based on categorical rules and exceptions miss significant generalizations. In particular, they argue that an analysis that rules out consonant pairs categorically based on feature classes cannot simultaneously explain exceptions to the rule, and the low rate of occurrence of consonant pairs that do not violate the rule.

The table in (15) presents various feature class analyses, and their implications in terms of lexical statistics. Feature classes based only on very general features have large numbers of exceptions, like the 'Place only' class, which would rule out any word containing a consonant pair sharing place of articulation, such as [t] and [n]. More specific feature classes can reduce the number of exceptions, but combinations outside of these classes should co-occur freely. For example, in the 'Place & [son] & [cont]' class, combinations of [t], [n], and [s] should be found freely. If combinations outside of these classes co-occur freely, then an estimate of the number of combinations that should be found in the lexicon can be made based on the rate of co-occurrence of non-homorganic consonant combinations. Thus, for any particular definition of classes, all under-representation outside of those classes is unexplained. So, for example, while the 'Place & [son] & [cont]' co-occurrence classes have few exceptions (36), they also predict many combinations of consonants that should be found in the lexicon but are not (estimated 430.5).

(15) *Trade-off with the rule plus exceptions approach*

Definition of Classes	Exceptions	Unexplained Under-representation
Place only	816	–
Place & [son]	123	160.8
Place & [son] & [cont]	36	312.7
Enumerated pairs	–	430.5

In summary, it appears that statistical phonotactic dissimilation patterns in the lexicon are grounded in similarity, as the degree to which co-occurrence is avoided is straightforwardly reflected in the degree to which consonants are similar (Pierrehumbert 1993b, Frisch et al. 2004). These differences in degree of co-occurrence suggest that grammar itself includes probabilistic information. Otherwise, the grammar will either contain systematic exceptions, or be unable to explain systematic non-occurrence. The question of how to interface categorical and probabilistic patterns is raised below, but it appears to be one of the real challenges of contemporary linguistics.

16.6 Conclusion and issues for further research

Generative linguistics proposes to restrict the range of possible grammars in several ways by hypothesizing that speakers construct a grammar by selecting rules or constraint rankings from an inventory available in Universal Grammar (e.g., Chomsky & Halle 1968, Prince & Smolensky 2004). Universal Grammar (UG) accounts for parallels in the grammatical patterns of unrelated languages. UG also provides a limited set of possible grammars, and predicts that some linguistic patterns cannot occur because they cannot be described by one of the possible grammars. In the same way, UG can also explain why the same phonological patterns are seen in different domains of grammar. As mentioned in Section 16.4, dissimilation is observed in both segmental and suprasegmental phonology, and exhibited in both static lexical distributions and dynamic alternations. These different types of dissimilation can be unified in the grammar under certain guiding principles, like the OCP and parallel evaluation of markedness constraints.

The unification of facts about lexical distributions and facts about morpho-phonological alternations is desirable when it can be achieved, because the resulting theory of phonological competence is simpler than a theory that proposes distinct mechanisms. This is the analytical reasoning supporting the Duplication Problem. However, their unification rests on the conclusion that the patterns are indeed the same. There is growing evidence that phonotactic constraints, including dissimilation, can be gradient and statistical in nature, and that native speakers have very fine-tuned intuitions about phonotactic distributions. Since statistical dissimilation patterns in the lexicon can be learned and generalized probabilistically, based on lexical data, reference to specifically linguistic constraints or learning mechanisms is not necessary to account for them. (Though specifically linguistic representations of segments or other prosodic constituents are still necessary in any analysis, as these categories are the foundation upon which the statistics are computed.) Morpho-phonological processes, on the other hand, may be more readily captured by a non-statistical constraint-based analysis and may be more naturally encoded in specifically linguistic ways in a grammar. No case studies of dissimilation that we are aware of, however, have actually empirically examined the distribution of similar elements in the lexicon and the occurrence of phonological rules that resolve them in alternations. Thus, it is not clear at this time whether the two types of generalizations should be accounted for using the same mechanisms.

Though outside the domain of dissimilation, one study that shows that these two types of generalization should in fact be due to different mechanisms is Crosswhite et al. (2003). Based on a *wug*-test for stress (Berko 1958), this study showed that native speakers of Russian have a categorical rule assigning default stress to stem-final syllables in novel nouns. Examination of the lexical statistics of Russian noun stress, however, showed that, while stem-final stress is the most common pattern of stem stress, its rate of occurrence far undershoots the rate of stem-final stress in the experimental data: stem-final stress was assigned to 80–90% of all experimental items, but this pattern is observed in only 30–60% of stem-stressed nouns in the Russian lexicon (with a higher percentage for low frequency items). Stress assignment in the novel words, therefore, does not directly match lexical frequencies for stress, but could perhaps be derived from a more intelligent algorithm that works from statistically systematic patterns to derive categorical ones.

If it is indeed the case that the language learner acquires both statistical phonotactic patterns and categorical morpho-phonological patterns, then a significant challenge for this line of linguistic research lies in resolving how these different types of generalization interface with one another. The integration of statistical and symbolic generalizations has been at the core of phonetics–phonology interface research for many years, since at least the departure from the interface model of Pierrehumbert (1990). Within the generative community, considerable attention has been given to producing statistical outputs from variable rankings of categorical constraints in OT (Albright & Hayes 2003, Anttila 1997, Hayes 2000). Outside of the generative community, the approach has been to take the gradient and statistical patterns to be the norm, and to attempt to derive the categorical patterns through some type of abstraction or generalization over the statistical data (Bybee 2001, Pierrehumbert 2001, 2003b, Skousen 1989).

On the other hand, even if the two types of dissimilation might not be formally unified within the grammar, they could still be unified in their functional origin. For example, Ohala's hypercorrection model of dissimilation could be applied to either type of data, and the fact that one is statistical while the other is categorical may be derived from the difference in kind between statistical lexical distributions and symbolic morpho-phonological processes. The lexicon is a large database of relatively static forms, so statistical generalization over this database is a natural process. Morpho-phonological processes, on the other hand, are symbolically formalized within a limited inventory of phonological structures (e.g. features, segments, etc.), leading more naturally to categorical generalization. The difference in kind between inventory restrictions and phonological processes is not specific to dissimilation, and could be applied to other cases of phonological processes that have a potential functional motivation.

Finally, it is important to note that the probabilistic account crucially depends upon linguistic categories, as these are the categories over which probabilities are determined. In an exemplar model, for example, each instance in which a lexical item is encountered is a distinct item encoded in the model, as no two experiences with a lexical item are physically exactly identical. Without some degree of generalization, there are no groups of patterns over which to encode probabilities. Thus, even the most gradient and statistical model requires some sort of categorical underpinning to account for phonological generalizations.

Note

We thank Paul de Lacy, Scott Myers, Jaye Padgett, and Kie Zuraw for their insightful comments.

~~References omitted because not separate for this article.~~

References

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

Alderete, John. 1997. Dissimilation as Local Conjunction. In *Proceedings of twenty seventh annual meeting of the North East Linguistics Society*, ed. Kiyomi Kusumoto, 17-31. Amherst, MA: Graduate Linguistic Student Association.

Anttila, Arto. 1997. Deriving variation from grammar. In *Variation, change, and phonological theory*, eds. Frans Hinskens, Roeland van Hout and W. Leo Wetzels. Amsterdam: Benjamins.

Archangeli, Diana, and Pulleyblank, Douglas. 1987. Minimal and maximal rules: Effects of tier scansion. In *Proceedings of NELS 17*, eds. Joyce McDough and B Plunkett, 16-35.

Berg, Thomas. 1998. *Linguistic structure and change: An explanation from language processing*. Oxford: Clarendon Press.

Berkley, Deborah. 1994. Variability in Obligatory Contour Principle Effects. In *Papers from the 30th Regional Meeting of the Chicago Linguistic Society, Part 1*.

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14:150-177.

Blevins, Juilette. 2004. *Evolutionary Phonology: the emergence of sound patterns*. Cambridge: Cambridge University Press.

Buckley, Eugene. 1997. Tigrinya root consonants and the OCP. *Penn Working Papers in Linguistics* 4:19-51.

Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Chomsky, Noam, and Halle, Morris. 1968. *The sound pattern of English*: Studies in language. New York: Harper & Row.

Crosswhite, Catherine, Alderete, John, Beasley, Tim, and Markman, Vita. 2003. Morphological effects on default stress placement in novel Russian words. In *Proceedings of the West Coast Conference on Formal Linguistics 22*, eds. Gina Garding and Mimu Tsujimura, 151-164. Somerville, MA: Cascadilla Press.

Dell, Gary S., Burger, Lisa K., and Svec, William R. 1997. Language production and serial order: A functional analysis and a model. *Psychological Review* 104:123-147.

Elmedlaoui, Mohamed. 1992. *Aspects des representations phonologiques dan certaines langues chamito-semitiques*. Thèse de doctorat d'état, Université Mohamed V, Rabat.

Frisch, Stefan, Broe, Michael, and Pierrehumbert, Janet. 1997. Similarity and phonotactics in Arabic. Ms. Northwestern University, Evanston, Illinois.

Frisch, Stefan, and Zawaydeh, Bushra. 2001. The psychological reality of OCP-Place in Arabic. *Language* 77: 91-106.

Frisch, Stefan A. 2004. Language processing and OCP effects. In *Phonetically-based phonology*, eds. Bruce Hayes, Robert Kirchner and Donca Steriade, 346-371. Cambridge: Cambridge University Press.

Frisch, Stefan A., Large, Nathan R, and Pisoni, David S. 2000. Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language* 42:481-496.

Frisch, Stefan A., Pierrehumbert, Janet, and Broe, Michael B. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22:179-228.
Goldsmith, John. 1976. Autosegmental phonology, MIT: Doctoral dissertation.
Greenberg, Joseph. 1950. The patterning of root morphemes in Semitic. *Word* 6:162-181.
Hay, Jennifer, Pierrehumber, Janet, and Beckman, Mary. 2004. Speech perception, well-formedness and the statistics of the lexicon. In *Phonetic Interpretation: Papers in Laboratory Phonology VI*, eds. John Local, Richard Ogden and Rosalind Temple, 58-74. Cambridge: Cambridge University Press.
Hayes, Bruce. 2000. Gradient well-formedness in Optimality Theory. In *Optimality Theory: Phonology, syntax, and acquisition*, eds. Joost Dekkers, Frank van der Leeuw and Jeroen van de Weijer, 88-120: Oxford University Press.
Hyman, Larry M. 1976. Phonologization. In *Linguistic studies offered to Joseph Greenberg on the occasion of his sixtieth birthday. Vol. 2: Phonology*, ed. Alphonse Juilland, 407–418. Saratoga: Anma Libri.
Itô, Junko, and Mester, Armin. 1986. The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17:49-73.
Itô, Junko, Mester, Armin, and Padgett, Jaye. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26:571-614.
Itô, Junko, and Mester, Armin. 2003. *Markedness and word structure: OCP effects in Japanese*. Cambridge, MA: MIT Press.
Iwai, Melissa. 1989. A prosodic analysis of Japanese loanwords, University of California, Santa Cruz: BA thesis.
Jebbour, Abdelkrim. 1985. Dialecte Tachelhit (Parler de Tiznit), Université Mahammed V, Rabat: Mémoire de Phonologie.
Josephs, Lewis. 1975. *Palauan reference grammar*. Honolulu: University Press of Hawaii.
Kessler, Brett, and Treiman, Rebecca. 1997. Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language* 37:295-311.
Leben, Will. 1973. Suprasegmental Phonology, MIT: Doctoral dissertation.
MacEachern, Margaret. 1999. *Laryngeal cooccurrence restrictions*. New York: Garland.
Marlett, Sephen A., and Stemberger, Joseph. 1983. Empty consonants in Seri. *Linguistic Inquiry* 5:617-639.
McCarthy, John J. 1979. Formal problems in Semitic phonology and morphology, MIT: Doctoral dissertation.
McCarthy, John J. 1986. OCP Effects: Gemination and antigemination. *Linguistic Inquiry* 17:207-263.
McCarthy, John J. 1988. Feature geometry and dependency: A review. *Phonetica* 43:84-108.
McCarthy, John J. 1994. The phonetics and phonology of Semitic pharyngeals. In *Papers in Laboratory Phonology III*, ed. Patricia A. Keating, 191-233. Cambridge: Cambridge University Press.
Mester, Ralf-Armin. 1986. Studies in tier structure, University of Massachusetts, Amherst: Doctoral dissertation.

Moreton, Elliott. 2002. Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84:55-71.

Myers, Scott. 1987. Tone and the structure of words in Shona, University of Massachusetts, Amherst: Doctoral dissertation.

Odden, David. 1994. Adjacency parameters in phonology. *Language* 70:289-330.

Ohala, John J. 1981. The listener as a source of sound change. In *Chicago Linguistic Society: papers from the parasession on language and behavior*, eds. Carrie S. Masek, Roberta A. Hendrick and Mary Frances Miller, 178-203. Chicago, IL: Chicago Linguistic Society.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In *The Production of Speech*, ed. P. MacNeilage, 189-216. New York: Springer-Verlag.

Ohala, John J. 1992. What's cognitive, what's not, in sound change. In *Diachrony within Synchrony: Language History and Cognition*, eds. G. Kellerman and M. D. Morrissey, 309-355. Frankfurt: Peter Lang Verlag.

Ohala, John J. 1993. The phonetics of sound change. In *Historical Linguistics: Problems and Perspectives*, ed. C. Jones, 237-278. London: Longman Academic.

Öhman, Sven E.G. 1966. Coarticulation in VCV utturances: spectrographic measurements. *Journal of the acoustical society of america* 39:151-168.

Padgett, Jaye. 1995. *Stricture in feature geometry*. Stanford, CA: CSLI Publications.

Pater, Joe. 1999. Austronesian nasal substitution and other NC effects. In *The prosody morphology interface*, eds. René Kager, Harry van der Hulst and Wim Zonneveld, 310-343. Cambridge, MA: Cambridge University Press.

Pierrehumbert, Janet. 1990. Phonological and Phonetic Representation. *Journal of Phonetics* 18:375-394.

Pierrehumbert, Janet. 1993. Dissimilarity in the Arabic verbal roots. In *NELS 23*, 367-381.

Pierrehumbert, Janet. 2003. Probabilistic phonology: Discrimation and robustness. In *Probability theory in linguistics*, eds. Rens Bod, Jennifer Hay and Stefanie Jannedy, 177-228. Cambridge, MA: The MIT Press.

Prince, Alan, and Smolensky, Paul. 1993. Optimality theory: constraint interaction in generative grammar. Piscataway, NJ: Rutgers Center for Cognitive Science.

Selkirk, Elisabeth. 1993. [Labial] relations: University of Massachusetts, Amherst.

Skousen, Royal. 1989. *Analogical modeling of language*. Dordrecht: Kluwer.

Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. Ms. Handout presented at the Rutgers University Optimality Workshop, New Brunswick, NJ.

Smolensky, Paul. 1995. On the internal structure of the constraint component *Con* of UG: Handout of talk given at University of Arizona.

Steriade, Donca. 1987. Locality conditions and feature geometry. In *Papers of NELS 17*, 595-617. Amherst, MA: Graduate Linguistic Student Association.

Suzuki, Keiichiro. 1998. A Typological Investigation of Dissimilation, University of Arizona: Doctoral dissertation.

Walsh Dickey, Laura. 1997. *The phonology of liquids*. Amherst, MA: Ph.D. dissertation, University of Massachusetts.

Yip, Moira. 1988. The Obligatory Contour Principle and phonological rules: A loss of identity. *Linguistic Inquiry* 19:65-100.
Yip, Moira. 1989. Feature geometry and cooccurrence restrictions. *Phonology* 6:349-374.
Zuraw, Kie. 2000. Patterned exceptions in phonology, UCLA: Doctoral dissertation.