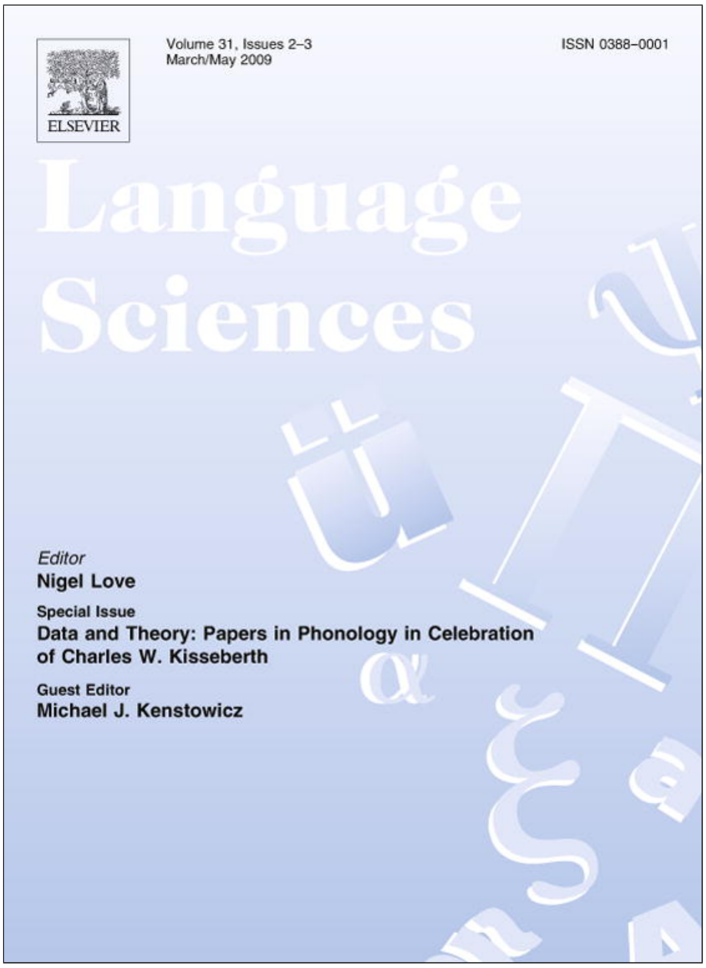

Volume 31, Issues 2–3
March/May 2009

ISSN 0388-0001

ELSEVIER

Language Sciences

Editor
Nigel Love

Special Issue
Data and Theory: Papers in Phonology in Celebration of Charles W. Kisseberth

Guest Editor
Michael J. Kenstowicz

Available online at www.sciencedirect.com
ScienceDirect
Language Sciences 31 (2009) 144–160
Language Sciences
www.elsevier.com/locate/langsci

Emergent feature structures: harmony systems in exemplar models of phonology

Jennifer Cole

Department of Linguistics, University of Illinois at Urbana-Champaign, 4080 Foreign Languages Building,
707 South Mathews, Urbana, IL 61801, USA

Abstract

In exemplar models of phonology, phonotactic constraints are modeled as emergent from patterns of high activation between units that co-occur with statistical regularity, or as patterns of low activation or inhibition between units that co-occur less frequently or not at all. Exemplar models posit no *a priori* formal or representational properties to the phonological units or sound patterns that emerge from the statistical regularities of speech, in contrast to analyses in the generative phonology tradition, including Optimality Theory, where sound patterns are determined by well-formedness constraints on phonological structures. This paper focuses on the analysis of long-distance assimilation, i.e., harmony systems, evaluating the predictions of generative analyses based on constraints on representation against typological and experimental evidence. Representational approaches model assimilation with constraints that favor extended feature structures. The question addressed here is whether and how the feature structures of harmony systems can be modeled as emergent structure. It is shown that an exemplar account can model the co-occurrence patterns of harmony systems in the transitional probabilities between segments that share the harmony feature, without invoking feature structure, but that the domain properties of harmony feature structures emerge due to the associations between phonological units (the harmonizing segments) and the morphological units that delimit harmony domains. This association grounds the sound pattern in the lexicon, and provides a "convergence of regularities" [Frisch, S., 2007. Levels of representation in acquisition (Commentary). In: Cole, J., Hualde, J.I. (Eds.), Laboratory Phonology, vol. 9. Mounton de Gruyter, Berlin, pp. 339–352] which facilitates learning.

Keywords: Exemplar models; Phonetics; Phonological processing; Phonology; Vowel harmony

---

1. The role of phonological structure: exemplar vs. generative models

In exemplar models of phonology, phonological structure emerges from generalizations over the phonetic detail of speech as experienced by a speaker-hearer (see Bybee and Hopper, 2001; Lacerda, 2003; Lindblom, 1999; Pierrehumbert, 2003). Phonological units, such as syllables, segments or even sub-segmental features, are categories that emerge from the shared phonetic properties of words (or larger utterances) that are stored in the mental lexicon. Viewed in this way, a phonological category is defined on the basis of

---

[*] Tel.: +1 217 244 3057; fax: +1 217 333 8430.
E-mail address: jscole@illinois.edu

0388-0001/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.langsci.2008.12.004

a statistical relationship between the discrete level and the parametric phonetic level (Pierrehumbert, 2003, p. 119). For example, the phonological category of a vowel (phoneme or positional allophone) is emergent from the clustering of individual instances of vowels plotted in a two-dimensional phonetic space of F_1 and F_2, as shown by Peterson and Barney (1952) or Hillenbrand et al. (1995).

Phonological units at "higher" levels, such as syllables, can be modeled on the basis of statistical patterns relating segment-sized units that are consecutive in the speech stream (Vitevitch et al., 1999; Vitevitch and Luce, 2004). For instance, syllables can be defined on the basis of the biphone probability for two consecutive phone-sized units. Biphone probability is an estimate of the probability that a particular biphone sequence (e.g., -*tr*-) will occur within a word, as calculated from the total number of words in the language that contain the sequence and the token frequency of those words (Vitevitch and Luce, 2004). There is a higher biphone probability between segments that belong to the same syllable (e.g., between C and V as onset and rime) than between consecutive segments that span a syllable boundary. Phonotactic constraints can also be modeled in this way; the biphone probability between a legal CC- onset cluster is much higher than the biphone probability between consonants in an illegal cluster (which may be very low or even zero). Exemplar phonology can be implemented with neural network models, in which phonotactic constraints are seen as emergent patterns of high activation between units that co-occur with statistical regularity, or as patterns of low activation or inhibition between units that co-occur less frequently or not at all (e.g., Dell, 1990).

A strict exemplar models posits no *a priori* formal or representational properties to the phonological units that emerge from the statistical regularities of speech as it is spoken or heard, or to the sound patterns that are defined over co-occurring units (Lindblom, 1999). Any regularities in the sound patterns that occur across languages should follow from similarities in the phonetic substance underlying the sound patterns, and/or in the mechanisms for speech processing that are common to all speakers. The exemplar model perspective is in stark contrast to much of the research in the generative phonology tradition, including its contemporary form in constraint-based Optimality Theory (OT). In OT, sound patterns are determined by well-formedness constraints on phonological structure, and therefore the nature of phonological structure is a key element of any analysis. For instance, ranked constraints on feature co-occurrence can model markedness patterns that are reflected in the composition of sound inventories (the combination [+back, +round] is more highly ranked than [−back, +round]), and can also account for sound patterns that restrict feature co-occurrence, e.g., in patterns of assimilation or dissimilation. Other constraints regulate phonotactics and stress patterns through the evaluation of syllable and metrical structures (e.g., Hammond, 1999).

This paper focuses on the analysis of vowel harmony as a type of long-distance assimilation, comparing the predictions of generative analyses based on constraints on feature structures with an exemplar model account. Three questions frame this comparison. The first asks about the role of constraints on feature structure in accounting for the sound patterns of harmony systems in generative phonology. The second question asks whether harmony processes can be modeled without recourse to *a priori* phonological structure, and the third question asks whether the structures that serve to characterize and constrain harmony processes are in any sense emergent from statistical regularities in the sound patterns of words in a harmony language. I will argue that representational accounts of harmony in generative phonology make faulty predictions about the uniformity of harmony processes for different features, and that observed non-uniformities have parallels in non-uniformities in speech processing and in phonotactic learning for harmony patterns involving different features. I claim that an evolutionary model, where harmony processes are viewed as developments from precursor patterns of phonetic variation (e.g., coarticulation), can provide a better account of attested harmony patterns, and is compatible with the theory of exemplar phonology. Furthermore, I argue that an exemplar approach can model the co-occurrence patterns of harmony systems in terms of the transitional probabilities between segments that share the harmony feature, without invoking feature structure, but that the domain properties of harmony feature structures (as in Cole and Kisseberth, 1995) emerge due to the associations between phonological units (the harmonizing segments) and the morphological units that delimit harmony domains.

2. Predictions from representational models of harmony

Since the introduction of autosegmental feature structures (Goldsmith, 1976), vowel harmony has been modeled in generative phonology with rules and constraints that govern the phonological structures that encode harmony. In rule-based autosegmental accounts (e.g., Clements, 1976), harmony is modeled in the mapping from underlying to surface forms, with rules that introduce new associations between the harmony feature and segments in a directional process of feature spreading. For example, consider the well-studied vowel harmony system of Turkish (Clements and Sezer, 1982), where the backness feature of a suffix vowel assimilates to the backness feature of the vowel in the preceding syllable. In addition, a high suffix vowel also assimilates to the rounding feature of the preceding vowel. Examples are shown in (1), and the autosegmental feature structure that results from the backness spreading rule is illustrated in Fig. 1.

(1) Turkish vowel harmony (Data from Kenstowicz, 1994, p. 25)

Noun	plural	accusative	gloss
dal	dal-lar	dal-ɯ	'branch'
kol	kol-lar	kol-u	'arm'
kɯz	kɯz-lar	kɯz-ɯ	'daughter'
kul	kul-lar	kul-u	'slave'
yel	yel-ler	yel-i	'wind'
göl	göl-ler	göl-ü	'sea'
diʃ	diʃ-ler	diʃ-i	'tooth'
gül	gül-ler	gül-ü	'rose'

Later generative models stressed the role of constraints on the association between the harmony feature and the segments to which it spread, culminating in a number of similar proposals for the analysis of harmony in the constraint-based framework of Optimality Theory. In these non-derivational models of harmony there are no rules of feature spreading. The extended feature structures that characterize harmony are selected by markedness constraints that require feature structures (autosegments or featural domains) to align with the edges of a phonological word (e.g., Kirchner, 1993; Cole and Kisseberth, 1995), or to spread maximally within a prosodic domain (e.g., Kaun, 1995). Other analyses operate with markedness constraints that do not directly evaluate the extent or edge alignment of feature structures, but penalize forms in which there are multiple instances of the same feature structure (McCarthy, 2004), or in which there are adjacent non-agreeing feature specifications (Bakovic, 2000), while favoring forms in which a single feature specification characterizes all the target segments in a prosodic domain.

These constraint-based models of harmony differ somewhat in their predictions about the typology of occurring harmony processes (see McCarthy (2004) for comparison across models), but they share in common the characteristic that the mechanisms (rules or constraints) that give rise to harmony are formulated in very general terms. In principle, a harmony rule or constraint can be formulated to govern any phonological feature, with the ensuing prediction that all features may in principle participate in processes of long-distance assimilation, given the availability of compatible target phones in the phonological domain of assimilation (typically the word). Harmony rules or constraints are also uniform in their effect on targets in that the spreading harmony feature is expected to effect all eligible target phones within that domain. Some models of harmony will block feature spreading to a potential target if the target segment is ineligible due to incompatibility between the harmony feature and some other feature of the segment (e.g., a voiceless consonant may be

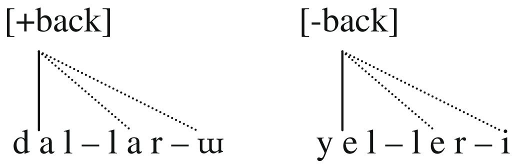

Fig. 1. Vowel harmony modeled through the spreading of the autosegmental feature [+/− back] from stem vowel to successive suffix vowels in the Turkish words *dal-lar-*ɯ 'branch, acc.pl.' and *yel-ler-i* 'wind, acc.pl.'.

ineligible as a target for the spread of [+nasal]), while other models simply require all segments in the harmony domain to take part in the assimilation (e.g., Ní Chiosáin and Padgett, 2001). Generative models of harmony are also similar in their treatment of the domain of harmony; assimilation spreads in an unbounded fashion from the source segment rightward or leftward to the end of the syllable, stress foot, or often the phonological word. Harmony is exhaustive within its prosodic domain, as stipulated by an unbounded rule of feature spreading, or through the alignment of the harmony feature structure with the edges of that domain.

These characteristics that are shared by all generative models of harmony yield several predictions about harmony systems across languages. Specifically, both the feature-spreading rules of earlier analyses and the constraints on extended feature structures of later analyses are sufficiently general that in the absence of additional machinery, they predict that:

1. Any feature can spread from any segment; *no asymmetries are predicted among features in their ability to spread and give rise to a harmony pattern.*
2. All segments are targeted, subject to feature compatibility; *no asymmetries are predicted among the segments that undergo harmony.*
3. The unmarked harmony process will spread the harmony feature (directionally) to the syllable/foot/word edge; *no restrictions are predicted on the extent of the harmony process across the word.*[1]

As discussed further below, these predictions are at odds with the observation of asymmetries in the segments that trigger or undergo harmony, and in the features that give rise to harmony systems. Some recent generative accounts seek to model the typological patterns through the incorporation of constraints that relate to the perception of phonological contrasts, or to constraints that relate to the position of a sound relative to word edges, which directly reflect perceptual and psycholinguistic mechanisms of speech processing (e.g., Kaun, 1995; Beckman, 1997; Walker, 2005). These accounts enrich the phonological model through the inclusion of additional constraints, while also providing grounding for phonological patterns in external (phonetic or psycholinguistic) factors. For instance, Walker (2005) discusses the harmony system in Veneto Italian dialects, and proposes an analysis based on constraints that serve to promote the perceptual salience of triggers. In Veneto Italian harmony, only [+high] vowels trigger harmony, spreading [+high] in one pattern to the preceding stressed vowel. The restriction to [+high] triggers is explained by appealing to perceptual salience: Due to their intrinsic short duration and lower amplitude, [+high] triggers are perceptually less salient than [−high] vowels in the same position, and the spreading of [+high] to the prominent stressed position effectively increases the perceptual salience of the [+high] feature. Constraint ranking ensures that priority is assigned to phonological structures that enhance the salience of the final vowel's [+high] feature through harmony, over alternative solutions such as lengthening the [+high] vowel, or shifting the location of stress to that vowel.

Walker's analysis, among others, demonstrates that generative models can be enhanced to provide better coverage of observed typological patterns, through the adoption of constraints that reflect limitations and biases of the speech processing mechanisms. This approach, in its reliance on factors external to phonology as a basis for phonological constraints, is compatible in spirit with the exemplar phonology model of harmony, introduced below. Both approaches locate the source of at least some typological asymmetries in factors external to phonology, and allow that these factors may shape sound patterns through a process of language evolution. The "enriched" models such as the OT model Walker proposes for Veneto Italian harmony differ from exemplar models in that they postulate grammatical devices to enforce the occurrence of sound structures that are favored by phonetic and other external factors. No such devices are invoked in an exemplar phonology approach.

The next section reviews evidence from a cross-linguistic typological study of vowel harmony alongside acoustic and behavioral evidence from recent experimental studies on the production of vowel harmony patterns. This evidence is evaluated against the three predictions introduced above from generative models of harmony that are based on the analysis of phonological feature structures.

---

[1] This prediction concerns only long-distance assimilation. Strictly local assimilation processes are bounded such that assimilation affects only a single target segment that is adjacent to the trigger (i.e., the source) of harmony.

3. Parallel observations of non-uniformity with harmony: typological and experimental evidence

3.1. Typological evidence for the non-uniformity of harmony processes

Comparative and typological studies of vowel harmony, such as Kaun's (1995) survey of labial harmonies, or Linebaugh's (2007) survey of lingual harmonies, reveal wide variation across attested harmony systems. Notably, there are very few harmony systems that instantiate the least complex type of system according to the formal models of harmony in generative phonology. That is, there are very few cases where the designated harmony feature (i) spreads directionally to the word edge, (ii) from every segment that is a source of the harmony feature, (iii) inducing assimilation of all target segments that intervene between the source and the word edge. I summarize the main findings from Linebaugh's survey of tongue body harmonies (backness and height) to illustrate.

Linebaugh's wide-ranging survey of tongue body harmony systems reviews the key features of the harmony systems from a geographically and genetically diverse pool of languages. Linebaugh's explicit goal is to compare vowel Backness Harmony (BH) and vowel Height Harmony (HH) for typological similarities and distinctions. His findings show greater variability for HH than for BH, and significant differences between BH as it occurs in Uralo-Altaic languages and BH elsewhere. BH in Uralic (Finnish, Hungarian, Samoyed) and Altaic (Turkic, Mongolian, Tungusic) exhibits the most robust, most general patterns of vowel harmony in any of the languages in Linebaugh's survey. In the Uralic and Altaic languages, all vowels participate in harmony,[2] harmony is bi-valued (involves both backing and fronting), harmony operates within morphemes, harmony extends throughout the prosodic word and is not restricted by morphological or prosodic limitations on triggers or targets. BH observed in other languages outside the Uralic and Altaic families (Chamorro, Kera, Tunica, Macuxi, Wikchamni) is more constrained. In those systems, only a subset of vowels participate, harmony is of the dominant/recessive type (targets are not typically also triggers), there are numerous, idiosyncratic morphological or featural limitations on the class of triggers and/or targets for harmony, and the vowel harmony pattern is not typically characteristic of all or most words. Linebaugh observes that HH is like BH outside of the Uralic and Altaic families in exhibiting many idiosyncratic differences across languages. He considers HH in Bantu, Romance, and other languages and finds that HH is rarely bi-valued (i.e., the same process rarely conditions both lowering and raising in the same system). Instead, there are many cases where HH works to lower target vowels, and others where HH raises target vowels, with some languages exhibiting both patterns in distinct HH phenomena. Some HH systems are asymmetrical in the class of triggers, distinguishing between front and back vowel triggers. Finally, HH does not generally extend throughout the phonological word but is often morphologically restricted (e.g., excluding prefixes as triggers)[3] or prosodically conditioned (e.g., dependent on stress conditions).

Linebaugh's survey reveals an impressive variety of vowel harmony patterns, with only Uralic and Altaic (and perhaps most famously, Turkish) coming close in exhibiting the kind of pervasive and robust pattern of harmony that can be modeled with the basic, unadorned harmony rule or constraint. In all other systems of vowel backness or height harmonies, the asymmetries that are observed are not predicted by formal models that invoke rules or constraints that optimize extended feature structures as the general case, at least not without added constraints or stipulation. In the next section, we consider parallel evidence of asymmetry between vowel height and vowel backness in speech production from experimental findings. The parallel occurrence of vowel backness and height asymmetries in typological studies and speech production experiments is significant as it suggests that production factors may play a role in determining the sound patterns of vowel harmony systems.

---

[2] Even the neutral vowels that exist in languages such as Hungarian participate as harmony triggers, and exhibit low-level coarticulatory variation (Gafos and Benus, 2007).

[3] Hyman (2002) also observes an asymmetry between prefixes and suffixes in their behavior as triggers in vowel height harmony systems, and suggests an explanation in terms of the phonetic basis for harmony in patterns of V-to-V coarticulation. He claims that effects of anticipatory coarticulation are stronger than perseveratory effects, which through the process of phonologization would result in fewer instances of prefix vowels as harmony triggers than suffix vowels. This line of reasoning, linking synchronic phonological patterns to their bases in phonetic patterns, is in lines with the proposal advanced in this paper.

3.2. Experimental evidence for backness/height asymmetries in speech production

Linguists have long noted that phonological patterns arising from diachronic sound change can often be characterized as facilitating speech production by reducing articulatory complexity (Grammont, 1933; Hock, 1991). Assimilatory sound change is one such example, since when one sound assimilates to another there is a potential overall reduction in the number of distinct articulatory gestures to be performed (Lindblom, 1983). To test the hypothesis that vowel harmony may facilitate speech production by reducing articulatory complexity, I have conducted a series of production studies that compare the speed and accuracy of production under conditions of vowel harmony and disharmony in tasks of speeded repetition, with speakers of Spanish (Linebaugh and Cole, 2005; Linebaugh, 2007), Korean (Oh and Cole, 2006), and English (Cole et al., 2002). In this section I describe the experimental design and materials for all the languages studied, and summarize the results of the experiment with speakers of Spanish.

The experimental hypothesis for all three production studies is that vowel harmony facilitates production through the priming of features between harmonizing elements, and that priming effects should be apparent in an increase in production speed and/or accuracy. Priming effects are expected for all harmonies, regardless of the identity of the harmonizing feature. In each experiment, subjects are asked to repeat a nonsense phrase consisting of two CVCV nonsense words in one of three harmony conditions: Back Harmonic (BH), Height Harmonic (HH), and back + height disharmonic ("doubly disharmonic", DD). The Spanish and Korean studies compared all three harmony conditions, while the English study only compared back harmony to back disharmony. The vowels in the words were selected from the set /i, e, o, u/ for all three languages, and the consonants were selected from the set {b, t, k} (Spanish), {b, g, k} (English), or the tense consonants {pp, tt, kk} (Korean). Example nonsense phrases from the Spanish experiment are shown in (2) along with the vowel sequences for each harmony condition.

(2) The harmony conditions and example nonsense phrases from the Spanish speeded repetition study

Height Harmonic (HH)	"*bibu la bobe*"
Back Harmonic (BH)	"*teti la tuto*"
Height/Back Disharmonic (DD)	"*koki la keku*"

HH: height harmonic, back disharmonic	i-u, u-i, e-o, o-e
BH: height disharmonic, back harmonic	i-e, e-i, u-o, o-u
DD: height disharmonic, back disharmonic	i-o, o-i, u-e, e-u

On each trial the subject read a single nonsense phrase on the computer screen, and at the prompt of a tone, initiated a speeded repetition of the phrase over a 4-s interval. Subjects were asked to speak as fast as possible without compromising accuracy. Subjects received no feedback on their speed or accuracy during the experiment. In the Spanish experiment 20 subjects produced 108 non-identical phrases, each phrase consisting of two words from the same harmony condition (HH, BH, DD) for a total of 108 distinct test phrases (12 possible vowel sequences for the first word, three possible vowel sequences for the second word, and three possible consonants yields $12 \times 3 \times 3 = 108$ phrases). Subjects were native speakers of the experiment language, and were students at the University of Illinois between the ages of 21 and 40, with roughly equal numbers of females and males. Subject responses to each trial were recorded onto digital tape using a desktop microphone, and later coded for rate (syll/sec) and errors (production of incorrect V). In each study, a subset of the data was coded by a second coder with inter-coder agreement above 90% (in the Spanish study, 400 trials from eight subjects were recoded, with 92.5% inter-coder agreement).

The results from Spanish speakers were analyzed using repeated-measures ANOVA with the independent factor of Harmony Condition (HH, BH, DD). There was a significant effect of harmony condition on the number of errors ($F(2, 19) = 3.67$, $p < .05$), but no significant effect on speech rate. Table 1 shows the mean, s.d.,

Table 1
Mean, s.d., and % of errors in vowel production from speeded repetition experiment with 20 Spanish speakers (pooled). Results are shown for the conditions of Back Harmony (BH), Height Harmony (HH), Back + Height Disharmonic (DD), and for the combined sets of back disharmonic data (HH + DD) and height disharmonic data (BH + DD).

	BH	HH	DD	Non-BH (HH + DD)	Non-HH (BH + DD)
Mean	3.0	3.7	3.9	3.8	3.4
s.d.	2.6	2.4	2.7	2.4	2.5
% Errors	8.2%	10.3%	10.8%	10.6%	9.5%

and % of errors in vowel production for the 20 Spanish speakers, pooled. Graphs of the error rates as a function of harmony condition are shown for each of the intervening consonants separately in Fig. 2.

When the intervening consonant is labial /b/ or velar /k/, there are lower error rates for BH than for HH or DH (the two non-BH conditions). This finding indicates that BH has a facilitative effect on production relative to non-BH forms. When the intervening consonant is dental /t/, the facilitation effect extends to the HH condition as well. The interaction between harmony condition and intervening consonant is a puzzling result in the Spanish data which is not replicated in the experiments with Korean (the English experiment did not include /t/). In each experiment we find marginal or non-significant differences in error rates as a function of intervening consonant, but there is no consistency to the pattern across the experimental languages. In the Korean data error rates were lowest for /pp/, while in the English data error rates were lowest for /k/.

While the consonantal effect on error rates does not generalize across experiments, the finding of an asymmetry between BH and HH does. The Korean data closely mirror the Spanish data, with lower error rates in the production of VCV sequences under the BH condition, compared to HH and DD (the non-BH conditions). The experiment with English speakers tested only for effects of BH, comparing back-harmonic and back-disharmonic forms, with no HH condition, and there were lower error rates in the BH condition with no effects on speech rate.

One obvious difference between BH and HH vowel patterns in our experimental conditions is that the BH condition pairs together vowels that share not only backness, but also roundness: {i, e} vs. {u, o}. Vowels paired in the HH condition share only height in common, and differ in roundness: {i, u} vs. {e, o}. It is possible that the asymmetrical facilitation effects of BH vs. HH reflect the greater similarity (as measured by shared features) between vowels in the BH condition, rather than any intrinsic difference between vowel backness and height in speech production. To test this hypothesis, Oh and Cole (2006) ran a second experiment with Korean subjects, using only the unrounded vowels {i, e, ɨ, ʌ}. This experiment replicated the first Korean experiment in every other detail, but by using all unrounded vowels, the BH sets {i, e} and {ɨ, ʌ} differ only in the feature [back], just as the HH sets (i, ɨ) and {e, ʌ} differ only in the feature [high]. The results were the same: production errors were fewer in BH words compared to HH words, while there was no effect of HH on errors (and no effects from any harmony condition on speech rate). This result provides clear evidence of a produc-

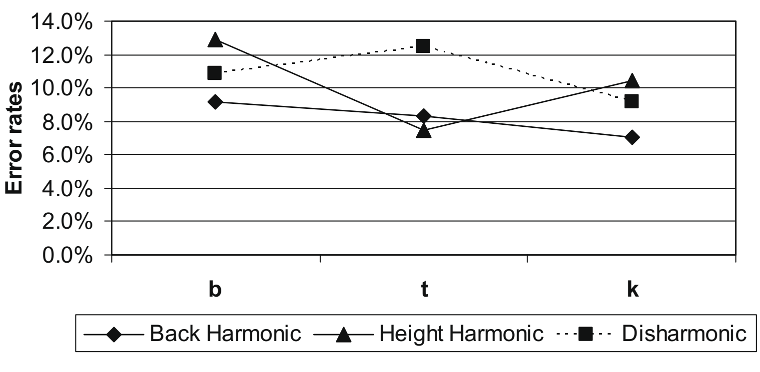

Fig. 2. Error rates for vowel production from the speeded repetition experiment with 20 Spanish speakers (pooled data), grouped by intervening C.

tion asymmetry between vowel backness and vowel height that is independent of the phonological similarity between vowels in a harmony condition.

In summary, the findings from speeded repetition production studies with speakers of three languages show a facilitative effect of back harmony on fast speech production, with no systematic corresponding effect from height harmony. Note that none of these languages have phonological rules or constraints on back harmony or height harmony, so we assume that the findings reflect properties of speech production that are independent of language-specific phonological properties. Rather, the back harmony effect on the accuracy of fast speech must reflect a facilitation effect in speech planning that arises in back harmony conditions, but not in height harmony conditions. Facilitation could occur at the level of phonological planning, articulatory (phonetic) planning, or both. At the phonological level, facilitation could arise from priming of the phonological features across syllables: utilizing the same feature in successive vowels in the harmony condition may reduce the complexity of constructing the phonological plan for the phrase. This characterization of priming is compatible with the representation of harmony in terms of extended feature structures. Yet, if the effect on speeded repetition is due to phonological priming, we expect the same effects to arise for height harmony as for back harmony, especially in the Korean experiment where rounding was not a factor, since there is a single shared phonological feature in both cases. Because the shared height feature in the height harmony condition fails to systematically influence speech production, we conclude that the facilitation effect of back harmony does not reflect priming at the phonological level alone, and must reflect differences in speech planning at the phonetic level.

It is beyond the scope of this paper to determine the source of the facilitation effect of back harmony on speech planning, but there are asymmetries in coarticulation and articulatory complexity that may relate to our production findings, and which should be investigated in future work. Since the work of Öhman (1966) it is known that vowels coarticulate across an intervening consonant, in VCV sequences. While coarticulation effects have been reported for both F1 and F2 measures, reflecting coarticulatory influences in the production of vowel backness and height, direct comparisons of coarticulation in these two dimensions are few. If coarticulation effects were found to be greater in the backness dimension compared to height, we may expect that the back harmony condition would yield a bigger benefit of articulatory stability than height harmony. Another difference relating to the production of vowel backness and height concerns the articulatory correlates of the phonological features. The phonological backness feature of a vowel is implemented primarily through activity of the genioglossus muscle, but there is no single articulator that implements vowel height. Vowel height targets are achieved through the coordination of tongue body and jaw gestures, and the contribution of these two articulators can vary substantially from vowel to vowel. Thus, if the stability of an articulatory gesture is a basis for the facilitation effects we observe with back harmony, such effects are less likely to extend to the multiple articulators required to achieve height targets.

Factors related to coarticulation and/or articulator stability may offer plausible phonetic explanations for the asymmetry between back and height harmonies in their effects on production. Moreover, these phonetic factors would be expected to influence typological patterns, if in fact harmony has a basis in the mechanisms of speech production. The observation of parallel asymmetries between height and backness harmony in typological studies and in experimental studies of speech production is compatible with an evolutionary account in which harmony develops from phonetic precursor patterns, which in turn are influenced by the constraints and biases of the speech production system. Specifically, height harmonies, where they arise, would not reflect the same production benefits we find with back harmony, and must therefore have their source in other phenomena, where different factors may play a role in shaping the harmony system.

4. The status of assimilation in phonotactic learning

One advantage of the representational account of harmony, i.e., one that invokes rules or constraints on feature structures, is that the extended feature structure that characterizes assimilation provides a direct means to account for the frequency of assimilation dependencies cross-linguistically: if extended feature structures are favored by grammatical rules and/or constraints, then assimilation is a favored outcome for any phonological system, since within the extended feature structure all (eligible) segments are associated with the same feature. Other kinds of long-distance dependencies can not be modeled with extended feature structures, and so would not be similarly favored. Moreton (2006, 2008) claims that the constraints and rules of Universal Grammar form a cognitive bias that makes assimilation patterns more learnable than other kinds of sound patterns, in some cases even when the non-assimilatory sound pattern has a stronger precursor in phonetic patterns, and that the better learnability of assimilation is at least one factor underlying the greater typological frequency of assimilation compared to other kinds of sound patterns. In this section I review findings from a series of phonotactic learning studies that demonstrate, contra Moreton's predictions, that adults can acquire novel patterns of co-occurrence between non-adjacent segments, with no differences between phonetically natural assimilation patterns and other patterns.

Phonotactic learning studies test the ability of subjects to acquire novel sound patterns through perception or production experience with an artificial language in the laboratory. Recent work by Pycha et al. (2003), Koo and Cole (2006) and Koo (2007) use the artificial language learning paradigm to investigate the role of phonetic naturalness and phonological complexity in phonotactic learning. Assimilation is recognized as a phonetically natural sound pattern, one that may reduce articulatory complexity or minimize perceptual confusability, and assimilation also has a relatively simple formal expression with extended feature structures. In comparison to assimilation, dissimilation is typologically less common, and does not share the same phonetic basis in that dissimilation does not in any obvious way reduce the complexity of speech production or perception.[4] The output of dissimilation is arguably more complex than assimilation, since it requires the inclusion of distinct features for the dissimilar sounds, and does not maximize the extent of any single feature. Nonetheless, dissimilation can still be expressed with feature notation, and is typically included in the repertoire of sound patterns that can be modeled by rules or constraints available in Universal Grammar (e.g., invoking constraints such as the Obligatory Contour Principle), within the framework of generative phonology. Based on the dual considerations of phonetic complexity and on the complexity of the phonological encoding, assimilatory sound patterns are predicted to be more readily or more effectively learned than dissimilatory sound patterns, and both assimilatory and dissimilatory sound patterns are predicted to more learnable than sound patterns that are phonetically arbitrary, and/or for which there is no phonological rule or constraint that favors such a pattern.

Koo and Cole (2006) test the predictions of phonetic naturalness and phonological complexity in phonotactic learning using a paradigm of implicit learning. Adult subjects were exposed to nonce words in an artificial language with a phonotactic constraint between non-adjacent phones. Four experiments were run to compare learning of constraints enforcing assimilation between non-adjacent vowels or consonants ($CVCV_iCV_i$ or CVC_iVC_iV) with learning of constraints enforcing dissimilation between non-adjacent vowels or consonants ($CVCV_iCV_j$ or CVC_iVC_jV). The vowel dependency paired the vowels {i, u} in patterns of total featural assimilation or dissimilation, where /iCi/ and /uCu/ represent an assimilation sequence, and /iCu/ and /uCi/ represent a dissimilation sequence. The consonant dependency paired the liquid consonants {l, r} in similar patterns of assimilation or dissimilation. Comparisons were made between groups of subjects learning the four different phonotactic constraints (V-assimilation, V-dissimilation, C-assimilation, C-dissimilation). In the experiments, subjects listened to a nonce word and repeated it as quickly as possible, striving for accuracy. Subjects were not told to attend to any particular property of the words, nor were they told that there was any special phonological pattern in the words. After an initial familiarization phase, in which subjects heard and repeated only words that were legal with respect to the phonotactic constraint of the experimental condition, subjects were exposed to illegal words (i.e., words that violated the assimilation or dissimilation condition), which were interspersed with new tokens of legal words. If the experimental constraint is learned during the familiarization phase, through the acts of listening and repeating, then subjects should be quicker in responding to new legal words than to new illegal words presented in the test phase.

Measures of reaction times confirm this prediction for the consonant assimilation condition as well as for the consonant dissimilation condition. These findings demonstrate that implicit phonotactic learning can

---

[4] This is not to deny the possibility that dissimilation may arise due to phonetic factors. Ohala (1993) suggests that that dissimilation may arise due to the mis-parsing of phonetic cues, when a listener assumes (incorrectly) that a similarity between two nearby sounds is due to assimilation, and then reverses the assimilation to render the two sounds distinct. Under this view, dissimilation increases the phonetic distinctiveness between two sounds, but this does not by itself mean that the dissimilated sounds are easier to perceive than the corresponding sequence of similar sounds.

occur based on brief auditory and repetition experience (corroborating the findings from Onishi et al., 2002), and also demonstrate that dissimilation patterns can be learned as effectively as assimilation patterns, even when subjects are not aware they are learning anything at all. But the results from the vowel assimilation and dissimilation were not as expected; there was no evidence of learning for either of the vowel constraints (no significant difference in reaction times to legal and illegal words). This result was unexpected due to the greater typological frequency of non-adjacent vowel dependencies compared to non-adjacent consonant dependencies.

Koo and Cole suggest that the apparent lack of evidence for learning of vowel dependencies may be due to perceptual factors. Specifically, they note for English there is greater potential for perceptual confusion between the constrained liquid consonants /l, r/ than between the constrained vowels /i, u/, and that the effect of the phonotactic constraint restricting the liquid consonants is that the perceptual confusion is reduced. Once the subject has learned the constraint, the task of identifying a possible /l/ or /r/ in the third syllable position is simplified by the identification of an /l/ or /r/ in the preceding syllable, resulting in quicker identification of the word. The phonotactic constraint on vowels will not have such a beneficial effect on recognition if the constrained vowels are not highly confusable. The prediction is that there would be evidence of learning for a vowel phonotactic pattern if the constrained vowels were relatively confusable, and conversely, that there may be no evidence of learning for a consonant phonotactic pattern if the constrained consonants are relatively less confusable. These further predictions are tested and confirmed by Koo and Cole (2007) and Koo (2007) through additional experiments involving implicit learning. For instance, using the same auditory repetition procedure as in the first set of experiments, Koo and Cole (2007) find evidence that subjects learned a vowel harmony pattern (faster responses to legal novel words) when the constrained vowels were rendered perceptually confusable through the addition of noise in the acoustic signal. Or, in an experiment using an offline task where subjects produce grammaticality judgments for novel words, Koo (2007) finds evidence that subjects learned both vowel harmony and liquid harmony, with comparable ease and accuracy.

Summarizing over all the phonotactic learning experiments conducted by Koo and Cole, we find that subjects can (implicitly) learn a harmony constraint over liquids as readily as the corresponding disharmony constraint, and they can learn a constraint on vowel harmony as readily as a constraint on liquid harmony. The vowel harmony pattern was expected to be the most easily learned of all the experimental conditions, given the strong phonetic basis for phonological vowel harmony in V-to-V coarticulation, reflected also in the greater frequency of assimilatory vowel dependencies compared to dissimilatory vowel dependencies across languages. In all, the studies by Koo and Cole find no support for Moreton's hypothesis of a learning bias that favors assimilation over dissimilation, or of a bias that favors typologically common and/or phonetically grounded processes over rare or phonetically arbitrary processes.

In a different type of phonotactic learning experiment, Pycha et al. (2003) also test the hypothesis of a learning bias that favors phonetically natural processes, by comparing learning of vowel backness harmony as a phonetically natural pattern to backness disharmony as an "unnatural" pattern—a pattern of dissimilation which they claim cannot be accounted for by any 'strictly phonetic tendency' on its own (Pycha et al., 2003, p. 104). In addition to the harmony and disharmony patterns, Pycha and colleagues included a third condition in which subjects were tested on learning of a phonetically arbitrary sound pattern. Their study involved a test of explicit learning, as subjects were told to learn the stem-suffix pattern of the artificial language and apply it offline in judging new words. Ten subjects were assigned to learn each of the three phonotactic patterns (Harmony, Disharmony, Arbitrary), and were tested on 36 novel words, through a grammaticality judgment task. Subjects in the harmony and disharmony conditions were more accurate in their responses than were subjects who were exposed to the phonetically arbitrary pattern, based on results from ANOVA and post hoc analyses, and from the non-parametric Mann-Whitney U-test (Harmony vs. Arbitrary: $U = 6$, $p = .001$; Disharmony vs. Arbitrary: $U = 17$, $p = .013$; Harmony vs. Disharmony: $U = 31$, $p = .15$). These findings indicate that the harmony and disharmony conditions were more effectively learned than the arbitrary condition.

Pycha et al. argue that the fact that the phonetically arbitrary rule was not learned suggests that learning depends on the availability of a phonologically simple characterization of the sound pattern. Both the assimilation and dissimilation patterns have such a characterization (involving the single feature [back]), but the arbitrary pattern had no such simple characterization. This finding of a learning bias that favors sound patterns that are simple in terms of their featural definition is interesting, but is also somewhat at odds with findings from studies of implicit phonotactic learning, such as Dell et al. (2000) or Onishi et al. (2002) where subjects effectively learn patterns of distributional restriction that are phonetically arbitrary. The point of convergence between the studies by Koo and Cole (2006), Koo (2007) and Pycha et al. (2003) is in the finding that phonotactic patterns involving dissimilation (disharmony) can be learned as effectively as assimilation patterns (harmony).

The phonotactic learning studies reviewed here show that adults can learn novel sound patterns in an artificial language, and that in this situation there is no apparent advantage for sound patterns involving assimilation compared to patterns involving dissimilation. These findings present strong evidence against the claim that there is a cognitive bias for learning assimilation patterns over other patterns, and cast doubt on the role of phonetic naturalness in phonotactic learning. The learning mechanism is capable of formulating phonological dependencies between segments even without a structural encoding of their dependency, as in the form of an extended feature structure. I conclude that extended feature structures are not privileged in phonotactic learning.

5. Harmony in exemplar models of phonology

In the studies reviewed above we fail to find evidence for the representation of harmony using extended feature structures from phenomena related to speech production and phonotactic learning. Section 3.1 reviewed a typological study of vowel backness and height harmonies, and noted asymmetries between backness and height harmonies in the classes of trigger and target vowels that participate in harmony, and in the occurrence of morphological and positional conditions on harmony. Section 3.2 reviewed evidence from speech production studies that reveal a parallel asymmetry between backness and height harmony in the speeded production of nonsense VCV patterns. Backness harmony facilitates accurate production, while no systematic production advantage is found for height harmony. Phonological models of harmony employing extended feature structures do not predict the occurrence of these kinds of asymmetries between harmonies involving different features, at least not without the addition of constraints that encode phonetic and other external factors. Furthermore, evidence from phonotactic learning studies fails to confirm that assimilatory sound patterns are privileged in relation to learning. Assimilatory sound patterns are learned as effectively as dissimilatory sound patterns, once perceptual factors are held constant. The evidence from phonotactic learning offers no direct support for a model of phonology that gives special status to assimilation patterns relative to dissimilation patterns, for instance by invoking representational devices like extended feature structures, and constraints that optimize such structures.

If the phonological model of harmony based on representations with extended feature structure is insufficient in its basic form to account for typological patterns, and if the phonological structures themselves are not supported through external evidence (from studies of speech production and phonotactic learning), we may consider whether there is a viable alternative model. One alternative, already mentioned, develops from the notion that sound patterns, including harmony, derive from phonetic precursors and as such directly reflect the factors that influence the planning, articulation and perception of speech. Under this view, exemplified in Blevins' (2004) Evolutionary Phonology, sound patterns do not result from, and are not explained by, properties of formal grammar. Rather, the phonological grammar of an individual language reflects the phonetic properties that arise within the lexicon of that language. But recognizing the source of harmony in precursor patterns does not in itself address the question of the mental encoding of a harmony pattern. Harmony systems clearly have a mental encoding. They comprise productive sound patterns that are transmitted across generations of speakers, and they can be learned by adults (as well as infants) even on the basis of brief exposure. In this section I sketch an account of harmony in terms of statistical regularities over non-adjacent segments in an exemplar model of phonology. Although the extended feature structures of representational models are not required in this account, something resembling the feature structure is an emergent property of the exemplar encoding.

Beginning with Öhman's influential study (1966), a wealth of experimental findings attest to robust and consistent patterns of vowel-to-vowel coarticulation in language. It is further known that listeners are sensitive to the effects of coarticulation, and though they compensate for coarticulatory effects in identifying a phone based on its acoustic properties, that compensation is not always complete (e.g., Beddor et al., 2002). In the residue of unparsed acoustic variability, there is the potential for harmony to develop through a process in which the listener misattributes the coarticulatory influence as an intrinsic feature of the perceptual target. In this way, harmony may develop as a phonotactic constraint from patterned coarticulation. With its source in the pervasive patterns of coarticulation found in every language, this model explains why feature assimilation is perhaps the most common type of phonological pattern. The generally local nature of harmony, by which eligible targets are not skipped, also follows directly from its basis in the language-specific, automatic processes of coarticulation. Furthermore, the fact that listener misperception is more likely when the coarticulating sounds share intrinsic (underlying) features in common also explains why harmony systems usually establish dependencies between triggers and targets that are featurally similar. Harmony patterns reduce perceptual confusability, which itself may be a factor that promotes rapid learning of the sound pattern, as suggested by the findings from Koo and Cole's (2007) study.

Through coarticulation patterns, vowels in adjacent syllables come to share phonetic properties in common. Under appropriate conditions, such as with a prosodically strong triggering segment, a prosodically weak target segment, and an intervening consonant that is not itself resistant to coarticulation, a target that is heavily influenced by a coarticulatory trigger may be reanalyzed by the listener as intrinsically specified for the coarticulating feature(s). When this occurs, it changes the patterns of V–V co-occurrence in the language, resulting in more instances of V–V sequences that share the features active in coarticulation, and fewer instances that do not share those features. These changing patterns of vowel co-occurrence can be modeled in an exemplar approach in terms of the transitional probabilities between vowels that share the harmony feature (Newport and Aslin, 2004). Fig. 3 illustrates the emergence of the harmonic classes of [+back] and [−back] vowels in Turkish through the greater transitional biphone probabilities that exist between vowels that co-occur within words (harmonizing vowels), and the lesser biphone transitional probabilities between vowels that co-occur more rarely (if at all). The harmonic classes are described here as emergent, because there is no *a priori* specification of the [back] harmony feature, nor is the class pre-defined in any sense. It emerges from the co-occurrence patterns of vowels in the lexicon.

A transitional probability in this model represents the probability that a vowel y will occur given the occurrence of a preceding vowel x, and is calculated based on the frequency of the sequence x–y in the language (i.e., type frequency, as calculated over lexical items), divided by the frequency of x in the language, as in (3).

$$(3)\qquad p(y|x) = \frac{p(xy)}{p(x)} \approx \frac{\text{freq}(xy)}{\text{freq}(x)}$$

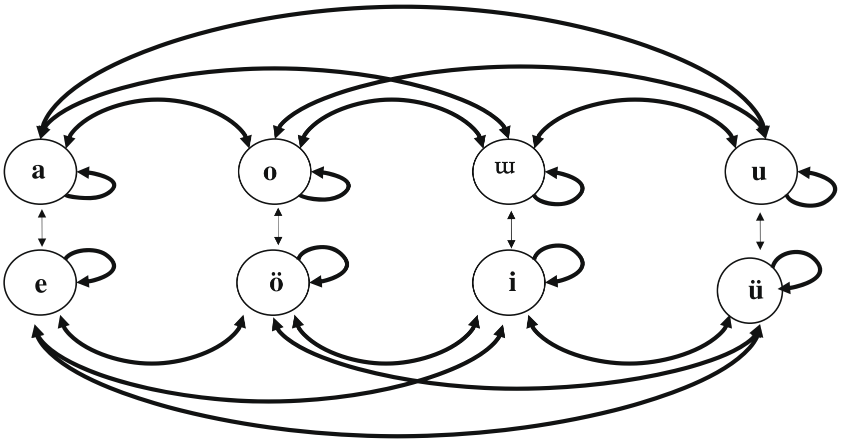

Fig. 3. Partial diagram of the transitional biphone probabilities between discrete vowels in Turkish. Thicker lines represent higher transitional probabilities, and model stronger co-occurrence relations between vowels that belong to the same harmonic class. Thinner lines (only four of 16 possible are shown) represent lower transitional probabilities between vowels that belong to different harmonic classes, but which may co-occur in disharmonic word forms.

Transitional probabilities can most easily be assessed for units that are adjacent in the phonological encoding. In the Articulatory Phonology model advanced by Browman and Goldstein (1992), articulatory gestures for vowels in successive syllables are adjacent in the gestural score. An onset consonant overlaps with the vowel of the same syllable, as shown in Fig. 4, where the vowels in successive syllables are strictly adjacent. But what about the dependencies involved in consonant harmony? Unless the consonants are string-adjacent (...CC...), they will not be adjacent in a gestural score (leaving aside the compressing effects of reduction on the realization of gestures), but transitional probabilities can still be modeled for non-adjacent elements. McClelland and Elman (1986) demonstrate how non-adjacent dependencies in syntax can be modeled with Simple Recurrent Networks, where "hidden" context units are used to encode relevant prior context at an arbitrary distance. The inclusion of hidden units comes at some computational cost, which may effectively limit the distance over which truly non-adjacent dependencies may be modeled. In harmony systems, dependencies that skip entire syllables are rare, but arise in systems that exhibit true transparency (possibly restricted to nasal harmony systems, see Ní Chiosáin and Padgett, 2001).

In addition to the high-probability transitions between vowels that share the same [back] feature, there may also be non-zero, low-probability transitions between vowels that differ in backness due to the presence of disharmonic words. These low-probability transitions are indicated with thin lines linking vowels in different harmonic classes in Fig. 3. In Turkish, disharmonic vowel sequences occur in some roots, though not all disharmonic sequences are attested with equal frequency (Clements and Sezer, 1982). The regularities within disharmony are also modeled as minor (low-frequency) phonotactic patterns in this approach.

The exemplar model of harmony sketched here does not rely on extended feature structures to capture the co-occurrence patterns of vowels within words, but a kind of multiply-linked feature structure emerges from the association between phonological and morpho-syntactic units, as follows. In an exemplar model, there are activation links between a word and the phone-size units that are part of the word. For a harmony language, there is a recurring pattern of activation between a harmonic class of vowels (phonological units) and words (morpho-syntactic units). Fig. 5 illustrates the flow of activation in the process of word recognition from speech input. The speech input causes activation of all phone-level units in the lexicon that match the input for some or all of its phonetic characteristics; the closer the match, the stronger the activation (represented in Fig. 5 by the thickness of the link between units).

Activation travels from the phone units "upward" to all word units which contain the phone, and "downward" from the word units to the phone units. This up-and-down pattern of activation between sound-level units (including phone and syllable) and word-level units is called *resonance* (Grossberg, 2003). A pattern of strong resonance between a specific word unit and a set of sound-level units results in the identification of the word, and the pattern of resonance between words and sound units derives harmony vowel classes in the lexicon, as illustrated in Fig. 6.

Harmony induces an alignment of the harmony feature with the edges of a morpho-syntactic domain, which is modeled in generative approaches by domain restrictions on feature spreading, or more directly through constraints on the alignment of feature domains with higher prosodic or morpho-syntactic domains (Cole and Kisseberth, 1995). In the exemplar model sketched here, the harmony domain is another instance of structure that emerges from the resonance between words and sound units. The sound units cluster into harmonic classes based on resonance patterns with word units, as sketched in Fig. 6, so the generalization about co-occurrence patterns of phones arises only in the mapping between phones and words. It is in this sense that we can describe the word as the domain of harmony. Furthermore, the strong activation patterns between the

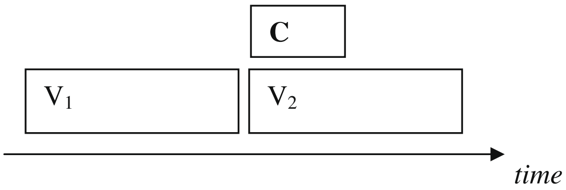

Fig. 4. A schematic gestural score of a $V_1 \cdot CV_2$ sequences showing the overlap of an onset consonant and its tautosyllabic vowel (V_2), and the adjacency of vowel gestures in successive syllables.

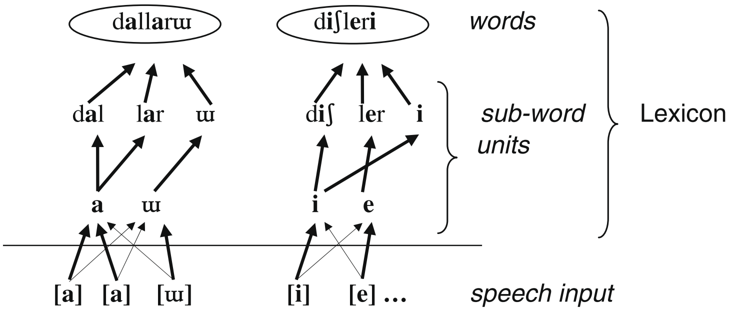

Fig. 5. Partial diagram showing activation spreading from units recognized in the speech input to units in the lexical encoding. Thicker lines link the speech units to the most closely matching phone-sized sub-word unit, and thinner lines represent weaker activations of phonetically similar but less well-matched phone units. Activation spreads ‘upward’ to syllable and word units (only the strongest activations are shown at these levels).

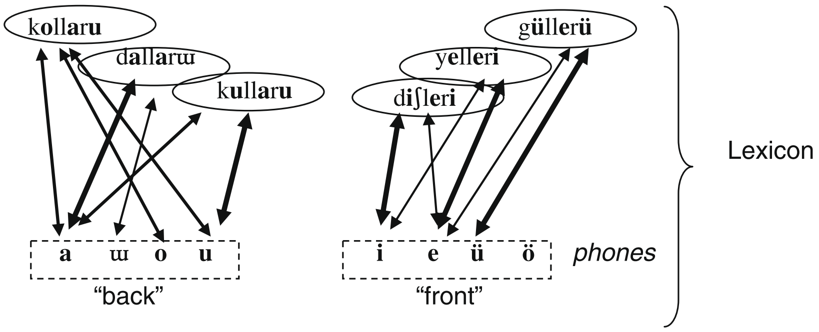

Fig. 6. Diagram illustrating *resonance* in the upward-downward activation between a word and its sub-word (phone) units, from which harmonic vowel classes (“back” and “front”) emerge.

word as a morpho-syntactic unit and a harmonic sound category can be considered as the analog of multiple association in representational models of harmony, also illustrated in Fig. 6. Here, the multiple association is between multiple word-level units and all the vowels that constitute a harmonic class. It is a pattern that is defined over the entire lexicon, not over any individual word. The resonance pattern that defines harmony is a particularly strong pattern since it links a set of sound units that are phonetically similar to word units that encode meaning at higher levels. This pattern of resonance grounds the harmony sound pattern in the lexicon, and provides a “convergence of regularities” that may facilitate learning (Frisch, 2007).

We have seen that harmony is encoded in exemplar models through strong patterns of association, or resonance, between phone-sized units (e.g., vowels) and words, and that these patterns of co-occurrence give rise to harmonic classes of vowels and derivatively, even the features we assign to them (e.g., [+back]). But if the exemplar model is to be considered a model of phonological encoding, we may ask how it fares against the typological and experimental evidence reviewed in Section 3. Can an exemplar model provide an explanation for the typological finding that harmony processes, and especially height harmonies, are complex and subject to myriad restrictions on trigger and target elements? Can it account for the asymmetrical effects of backness and height harmonies on speech production? The answer to these questions is both positive and negative. Exemplar models encode the phonetic detail of speech forms, and so any phonetic factors that influence the acoustic realization of a word (through production) or its auditory encoding (through perception) will be captured in the phonetic encoding of the exemplars. For example, if production factors exert a bias favoring the transmission of back harmonic vowel sequences over other kinds of vowel sequences, then there will be more instances of such sequences encoded as exemplars, and the greater frequency of that pattern may come to define a phonotactic regularity for the language. The source of the sound pattern is not explained through the exemplar model *per se,* but rather through its sensitivity to the phonetic detail that is the source of the pattern.

Similarly, exemplar models do not in themselves explain typological patterns, but rather typological patterns emerge from the phonetic (or other external) factors that shape the sound patterns of individual languages. The occurrence of strong phonetic biases favoring certain sound patterns, such as the co-occurrence of backness with rounding, or the reduction or devoicing of high vowels, means that those patterns are expected to occur frequently within and across languages, and to be widely evidenced in sound change. This is an evolutionary account of typology, as advocated by Blevins (2004), which is compatible with exemplar models, but which is also available with other models of phonology. For instance, Walker (2005) claims that the perceptual factor that motivates the restrictions on harmony targets and triggers in Veneto Italian dialects 'does not represent an intention on the speaker's part but rather it exerts influence on language change and shapes certain synchronic phonological processes through phonetically grounded constraints.' (p. 931). An important difference between the exemplar model of harmony introduced here and Walker's OT model is that the latter invokes a representational device (extended feature structures) and formal constraints on phonological representation to model harmony, while no such *a priori* devices are posited in the exemplar model account. The effect of phonetic and other external factors in shaping sound patterns is direct, and not mediated by an independent, universal grammar.

Continuing with the evaluation of the exemplar model of harmony, we saw in Section 4 that phonotactic learning does not appear to favor phonetically grounded sound patterns, such as assimilation, over other less-grounded patterns. The prediction from exemplar models would seem to be that any pattern that can be encoded in perceptual units, which themselves may be influenced by the prior linguistic experience of the individual, is in principle learnable. Sound patterns defined over adjacent units should be easier to learn than patterns defined over non-adjacent units, on the basis of computational considerations alone, when modeled in terms of neural networks, and that is consistent with the overwhelming evidence for locality as a constraining principle in phonology. Beyond locality, no other constraints on learnability are predicted from an unenriched exemplar model. Future research on phonological learning will be crucial in determining if there are in fact any substantive constraints on learning that are not due to limitations on phonetic processing.

6. Conclusion

Phonological theories that use rules or constraints on feature structures to model vowel harmony predict greater simplicity and uniformity in harmony processes across languages than is observed. Specifically, without additional enrichment from phonetically grounded constraints, such theories do not predict the asymmetries between height and backness harmonies observed in Linebaugh's (2007) typological study. Parallel asymmetries are observed in speech production experiments with harmony sequences, suggesting that at least some typological findings may have their source in properties of speech processing. With rules or constraints that favor extended feature structures, representational theories of phonology also predict that harmony patterns will be more learnable compared to other kinds of sound patterns, yet evidence from phonotactic learning studies shows no special advantage for harmony systems. An alternative to the representational models of generative phonology is introduced, where harmony is understood as developing from phonetic precursors in an exemplar account. In the exemplar model, lexical encoding includes phonetic detail and harmony patterns are modeled through the transitional probabilities between vowels in successive syllables in a word. Neither features nor extended feature structures are stipulated in phonological encoding. Rather, the patterns of activation, or resonance, between sound units and word units in the lexicon give rise to harmonic classes of vowels as emergent structure. Harmony is defined in terms of the links between sound units (phones) and meaning units (words), from which we obtain the word as the domain of harmony. And while there is no explicit model of feature spreading, patterns of multiple association emerge over the lexicon as a whole, as vowels from a harmonic class link individually and in combination with one another to words.

Acknowledgements

This work has benefited from the research and critical input of my collaborators: Gary Dell, Cindy Fisher, José I. Hualde, Alina Khasanova, Hahn Koo, Gary Linebaugh, Young-Il Oh, and other members of the

Phonotactic Learning group at the Beckman Institute at the University of Illinois. This work was supported by Grant NIH HD044458. The author is solely responsible for any errors or omissions.

References

Bakovic, E., 2000. Harmony, Dominance and Control. Doctoral Dissertation, Rutgers University, New Brunswick, NJ. ROA-360, Rutgers Optimality Archive, <http://roa.rutgers.edu/>.

Beckman, J., 1997. Positional faithfulness, positional neutralisation and Shona vowel harmony. Phonology 14, 1–46.

Beddor, P.S., Harnsberger, J.D., Lindemann, S., 2002. Language-specific patterns of vowel to-vowel coarticulation: acoustic structures and their perceptual correlates. Journal of Phonetics 30, 591–627.

Blevins, J., 2004. Evolutionary Phonology. Cambridge University Press, Cambridge.

Browman, C.P., Goldstein, L., 1992. Articulatory phonology: an overview. Phonetica 49, 155–180.

Bybee, J., Hopper, P., 2001. Frequency and the Emergence of Language Structure. John Benjamins, Amsterdam.

Clements, G.N., 1976. Vowel Harmony in Nonlinear Generative Phonology: An Autosegmental Model. Indiana University Linguistics Club, Bloomington, Indiana.

Clements, G.N., Sezer, E., 1982. Vowel and consonant disharmony in Turkish. In: Hulst, H., van der, Smith, N. (Eds.). The Structure of Phonological Representations, Part 2, pp. 213–255.

Cole, J., Kisseberth, C.W., 1995. An Optimal Domains Theory of vowel harmony. Studies in the Linguistic Sciences 24, 101–114.

Cole, J., Dell, G., Khasanova, A., 2002. Evidence of a production basis for front/back harmony. Poster Presentation at the 8th Conference on Laboratory Phonology, New Haven, Connecticut.

Dell, G.S., 1990. Effects of frequency and vocabulary type on phonological speech errors. Language and Cognitive Processes 5, 313–349.

Dell, G.S., Reed, K.D., Adams, D.R., Meyer, A.S., 2000. Speech errors, phonotactic constraints, and implicit learning: a study of experience in language production. Journal of Experimental Psychology: Learning, Memory, and Cognition 26, 1355–1367.

Frisch, S., 2007. Levels of representation in acquisition (Commentary). In: Cole, J., Hualde, J.I. (Eds.), Laboratory Phonology, vol. 9. Mouton de Gruyter, Berlin, pp. 339–352.

Gafos, A., Benus, S., 2007. Articulatory characteristics of Hungarian transparent vowels. Journal of Phonetics 35 (3), 271–444.

Goldsmith, J., 1976. Autosegmental Phonology. Ph.D. Dissertation, MIT. (Published 1979, Garland, New York).

Grammont, M., 1933. Traité de Phonétique. Delagrave, Paris.

Grossberg, S., 2003. Resonant neural dynamics of speech perception. Journal of Phonetics 31, 423–445.

Hammond, M., 1999. English Phonology. Oxford University Press, Oxford.

Hillenbrand, J., Getty, L.A., Clark, M.J., Wheeler, K., 1995. Acoustic characteristics of American English vowels. Journal of the Acoustical Society of America 97 (5), 3099–3111.

Hock, H.H., 1991. Principles of Historical Linguistics. Mouton de Gruyter, Berlin and New York.

Hyman, L., 2002. Is there a right-to-left bias in vowel harmony? In: Presented at the 9th International Phonology Meeting, Vienna. (Downloaded 3/15/08 from <http://www.linguistics.berkeley.edu/~hyman/Hyman-Vienna-VH-paper-forma.pdf>).

Kaun, A. 1995. The Typology of Rounding Harmony: An Optimality Theoretic Approach, Doctoral Dissertation, UCLA.

Kenstowicz, M., 1994. Phonology in Generative Grammar. Blackwell, Cambridge, MA.

Kirchner, R., 1993. Turkish vowel harmony and disharmony: an optimality theoretic account. In: Presented at Rutgers Optimality Workshop I http://roa.rutgers.edu/files/4-0000/roa-4-kirchner-3.pdf.

Koo, H., 2007. Change in the Adult Phonological Processing System by Learning Nonadjacent Phonotactic Constraints from Brief Experience: an Experimental and Computational Study. Doctoral Dissertation, University of Illinois at Urbana-Champaign.

Koo, H., Cole, J., 2006. On learnability and naturalness as constraints on phonological grammar. In: Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics. Athens, Greece, pp. 165–168.

Koo, H., Cole, J., 2007. Gradient perceptual facilitation from phonotactic knowledge. In: Proceedings of the 43rd Annual Meeting of the Chicago Linguistic Society.

Lacerda, F., 2003. Phonology: an emergent consequence of memory constraints and sensory input. Reading and Writing 16 (1–2), 41–59.

Lindblom, B., 1983. Economy of speech gestures. In: MacNeilage, P. (Ed.), The Production of Speech. Springer-Verlag, New York, pp. 217–245.

Lindblom, B., 1999. Emergent phonology. Proceedings of the 25th Meeting of the Berkeley Linguistic Society. University of California, Berkeley.

Linebaugh, G., 2007. Phonetic Grounding and Phonology: Vowel Backness Harmony and Vowel Height Harmony. Doctoral Dissertation, University of Illinois at Urbana-Champaign.

Linebaugh, G., Cole, J., 2005. Does vowel harmony facilitate speech production? In: Paper Presented at the 11th Mid-Continental Workshop on Phonology. University of Michigan, Ann Arbor.

McCarthy, J., 2004. Headed spans and autosegmental spreading. Unpublished manuscript, University of Massachusetts http://roa.rutgers.edu/files/685-0904/685-0904-0-0.PDF.

McClelland, J.L., Elman, J.L., 1986. The TRACE model of speech perception. Cognitive Psychology 18, 1–86.

Moreton, E., 2006. Phonotactic learning and phonological typology. In: Presentation at the 37th Meeting of the Northeast Linguistics Society, Urbana, Illinois.

Moreton, E., 2008. Analytic bias and phonological typology. Phonology 25, 83–127.

Newport, E.L., Aslin, R.L., 2004. Learning at a distance I. Statistical learning of non-adjacent dependencies. Cognitive Psychology 48, 127–162.

Ní Chiosáin, M., Padgett, J., 2001. Markedness, segment realization, and locality in spreading. In: Lombardi, L. (Ed.), Segmental Phonology in Optimality Theory. Cambridge University Press, Cambridge, pp. 118–156.

Oh, Y.I., Cole, J., 2006. A biased vowel harmony effect on speech production is not related to the number of features shared. In: Paper Presented at the 12th Annual Mid-Continent Workshop on Phonology. University of Iowa, Iowa City.

Ohala, J.J., 1993. The phonetics of sound change. In: Jones, C. (Ed.), Historical Linguistics: Problems and Perspectives. Longman, London, pp. 237–278.

Öhman, S.E.G., 1966. Coarticulation in VCV utterances: spectrographic measurements. Journal of the Acoustical Society of America 39, 151–168.

Onishi, K.H., Chambers, K.E., Fisher, C., 2002. Learning phonotactic constraints from brief auditory experience. Cognition 83, B13–B23.

Peterson, G.E., Barney, H.L., 1952. Control methods used in a study of the vowels. Journal of the Acoustical Society of America 24, 175–184.

Pierrehumbert, J., 2003. Phonetic diversity, statistical learning, and acquisition of phonology. Language and Speech 46, 115–154.

Pycha, A., Nowak, P., Shin, E., Shosted, R., 2003. Phonological rule-learning and its implications for a theory of vowel harmony. In: Tsujimura, M., Garding, G. (Eds.), Proceedings of the West Coast Conference on Formal Linguistics, vol. 22. Cascadilla Press, Somerville, MA, pp. 101–114.

Vitevitch, M.S., Luce, P.A., Pisoni, D.B., Auer, E.T., 1999. Phonotactics, neighborhood activation, and lexical access for spoken words. Brain and Language 68, 306–311.

Vitevitch, M.S., Luce, P.A., 2004. A Web-based interface to calculate phonotactic probability for words and nonwords in English. Behavior Research Methods, Instruments, and Computers 36 (3), 481–487.

Walker, R., 2005. Weak triggers in vowel harmony. Natural Language and Linguistic Theory 23, 917–989.