On measuring phonetic precursor robustness: A response to Moreton 2008*

Alan C. L. Yu

July 6, 2011

1 Introduction

What factors shape the synchronic typology of sound patterns and how should they be detected? To the extent that the synchronic typology of sound patterns follows from the results of language change, it is commonly, if not implicitly, assumed that analytic and channel biases are two major factors involved in shaping phonological typology (Zuraw 2007; Wilson 2006; Moreton 2008, 2010). Analytic biases are limitations in computation or markedness relations and constraints imposed by the Universal Grammar. An analytic bias might render certain patterns difficult to acquire even from perfect learning data. Channel bias refers to the relative likelihood of phonetic precursor to sound change becoming phonologized into full fledged sound patterns (e.g., Hyman 1976; Ohala 1993; Lindblom et al. 1995; Hume and Johnson 2001; Blevins 2004). This view of phonological typology is motivated by the commonly held assumption that sound patterns and sound changes that recur across unrelated languages originate in properties of human articulatory, perceptual and/or auditory mechanisms (Ohala 1983, 1993; Beddor et al. 2007). Context-induced phonetic variation in speech production and perception is taken to be the phonetic precursors to listener misperception-based sound changes. Phonologization refers to this transition of gradient phonetic variation (i.e. intrinsic allophones) becoming entrenched and developing into categorical phonological

---

*I have benefitted tremendously from discussion with Sharon Peperkamp and Morgan Sonderegger and comments from the three anonymous reviewers, the associate editors, and the audiences at the Manchester Phonology Meeting in 2009, the Workshop on Computational Modeling of Sound Pattern Acquisition at the University of Alberta, University of Pittsburgh, Michigan State University, and Stony Brook University. Naturally, all errors are my own. This material is based upon work partially supported by the National Science Foundation under Grant No. 0949754. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

patterns (i.e. the emergence of extrinsic allophones). Rare or uncommon sound patterns are due to the low probabilities of the corresponding phonetic effects being phonologized through sound change.

Much debate in recent years has focused on the relative contribution of analytic and channel biases in shaping the typology of sound patterns (Blevins 2004, 2006; Seidl and Buckley 2005; Hale and Reiss 2000; Hyman 2001; Kiparsky 2006, 2008; Moreton 2008, 2010; Wilson 2006; Zuraw 2007). The main issues are whether analytic and channel biases are in fact distinct and, if so, how they should be teased apart. Moreton (2008), in a recent thought-provoking paper, argues forcefully for the strength of analytic bias, such as Universal Grammar and other non-modality-specific cognitive biases that facilitate the learning of some phonological patterns and inhibit that of others, in creating typological asymmetries on its own unassisted by the robustness of phonetic precursors. His approach is three-pronged. Using as his case study the asymmetric rate of attestation between patterns involving vowel-to-vowel height dependencies (HH-patterns) and patterns concerning vowel height to consonantal voicing dependency (HV-patterns), he first conducted a brute-force library survey of grammars for the attestation of vowel height harmony and consonantal voicing-vowel height interaction. He then surveyed previous acoustic studies of vowel-to-vowel coarticulation and consonantal-voicing effect on vowel formants to assess the strengths of these two coarticulatory effects. He concluded that the skewed attestation of HH patterns over HV patterns cannot be explained by the relative robustness of their phonetic precursors since the relative magnitudes of coarticulatory effects are comparable, if not in the reverse direction of expectation (p.93). He then set out to test for potential analytic biases against HV patterns using an artificial grammar learning paradigm. He showed that English-speaking subjects are less able to learn HV patterns than HH patterns and concluded that there must be a preference for patterns involving the concurrence of a single feature over patterns involving the concurrence of multiple features.

This article focuses on the assessment of phonetic precursor robustness. The main goal of this article is two-fold: (i) to establish the inadequacy of Moreton's method of evaluating relative phonetic precursor robustness and to offer an alternative to his approach; (ii) to report the results of a cross-linguistic study comparing the nature of HH and HV interactions with the same languages – no previous studies have ever directly compare these two phonetic precursors before. The article is organized as follows: I begin by first establishing that Moreton's method of evaluating relative phonetic precursor robustness is inadequate as it measures inter-contextual variation, rather than intra-contextual one (Section 2). Using a rational optimal listener model to speech perception, I propose a method that more directly and objectively measures phonetic precursor robustness (Section 3). The effectiveness of

		Raising	Lowering	L/R ratio
a.	Manual 1990:	_Ci	_Cɑ	_Cɑ/_Ci
	Sotho	400.67	443.17	1.11
	Shona	410.67	471	1.14
	Ndebele	396.67	442.5	1.12
b.	van Summers 1987:	_VoicedC	_VoicelessC	L/R ratio
	English	578.67	691.67	1.2

Table 1: Vocalic F1 values in Hertz in different contexts and languages.

this measure is demonstrated using results of a cross-linguistic acoustic investigation of HH and HV effects in English and Turkish (Section 4). The results of this investigation suggest that Moreton's claim about channel bias having no place in explaining the skewed typology of HH and HV patterns is premature. This paper concludes with a discussion about the significance of the present study and the complexity of isolating analytic bias from channel bias (Section 5).

2 Variation and channel bias

Moreton (2008) tests the relevance of channel bias in explaining the asymmetric rate of attestation between HH patterns and HV patterns by carrying out a survey of existing phonetic studies "to assess the effect on target vowel F1 of the phonological height of a neighboring vowel, and compare it with the effect of phonological voicing, aspiration or fortis/lenis status of an immediately following consonant." (p. 93). He proposes to quantify the relative magnitude of the phonetic precursors (i.e. the magnitude of the coarticulatory effects) by comparing the target-vowel F1 in contexts where it is likeliest to raise or lower. In Moreton's own words, "[f]or HH studies, the Raising context consisted of high vowels, and the Lowering context consisted of low vowels. For HV studies, the Raising context was voiced, unaspirated or lenis obstruents, and the Lowering context was voiceless, aspirated or fortis obstruents." (p.93) The effect of context, which I refer to as the L/R ratio, is therefore defined as the target-vowel F1 in the Lowering context divided by the target-vowel F1 in the Raising context. This L/R ratio thus estimates the degree of inter-contextual variation; the further the ratio deviates from 1, the wider the range of F1 variation is going to be. To illustrate this concretely, consider the data in Table 1. Table 1a summarizes the average F1 values and the corresponding L/R ratios at the offset of /e/ and /ɑ/ preceding /i/ or /ɑ/ reported in Manual (1990)'s study of vowel-to-vowel coarticulation in three Bantu languages.

Table 1b shows the average F1 values at the vowel offset of /ɑ/ or /ae/ preceding voiced and voiceless obstruents *b, p, v, f* reported in van Summers 1987's study of final-consonant voicing on vowel production in English.

Drawing from measurements reported in a set of studies similar to Manual's and Summer's, Moreton calculated L/R ratios for the HH- and HV-precursors in different languages and found the HH-precursor to be less robust than the HV-precursor. This conclusion is surprising if the likelihood of phonologization is the primary driver behind the frequency typology of sound patterns since Moreton's own typological survey establishes that HH-interactions are more widely attested than HV-interactions. He thus concludes that the over-attestation of HH-patterns relative to HV-patterns must stem from some sort of analytic bias that favors the learning of intra-tier feature interaction over inter-tier featural co-occurrence restrictions.

While Moreton's L/R ratio uses existing phonetic studies to establish a phonetic typology, its effectiveness as a measure of phonetic precursor robustness is dubious. To begin with, it is limited to the comparison of precursors that influence segments along the same phonetic dimension. For example, the L/R ratio would have difficulty comparing the robustness of the precursor to velar palatalization compare to the precursor to HV-interaction since the main phonetic cues to velar palatalization are peak spectral frequency and formant transitions (Guion 1996), while the primary cue to HH- and HV-interaction is the trajectory of the first formant. More problematic is how the L/R ratio relates to the robustness of phonetic precursors. The L/R ratio is, at its core, a measure of intra-category variability induced by a phonetic precursor. That is, it provides a rough measure of the spread of the category distribution along certain phonetic dimensions; the numerator and denominator are the means of a category's variants at the extremes of a category's distribution. To illustrate this point more concretely, consider Figure 1, which shows the distribution of category A within the context of a phonetic precursor k along some phonetic dimension X. The grey lines delineate the distributions of A in the L and R contexts of precursor k. From the perspective of regressive HH-interaction, for example, where A is a vowel, the phonetic cue X is F1, and the precursor k is the presence of a following vowel, the L context (dashed line) would be A in the context of a low vowel while the R context (dotted line) is A in the context of a high vowel. The black solid line shows the overall distribution of A in the presence of precursor k. What this illustration shows is that the further apart the category distribution in the L and R contexts (i.e. the further the L/R ratio deviates from 1), the wider range of variation the precursor would induce.

While variation is a necessary condition to change, it is not a sufficient one. A category that has a wide variance may nonetheless show great stability if such a category overlaps

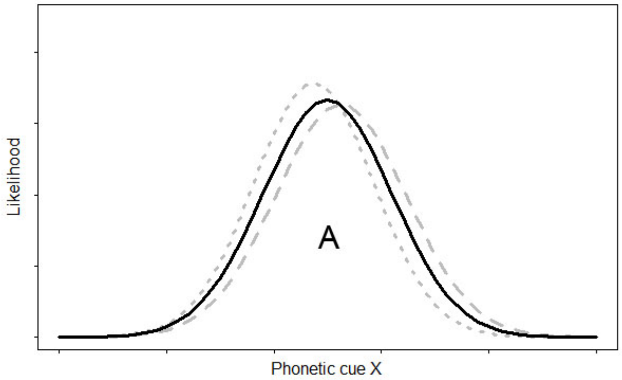

Figure 1: The distribution of Category A along some phonetic cue dimension X. The grey lines delineate distributions of Category A in the L (dashed line) and R (dotted line) contexts.

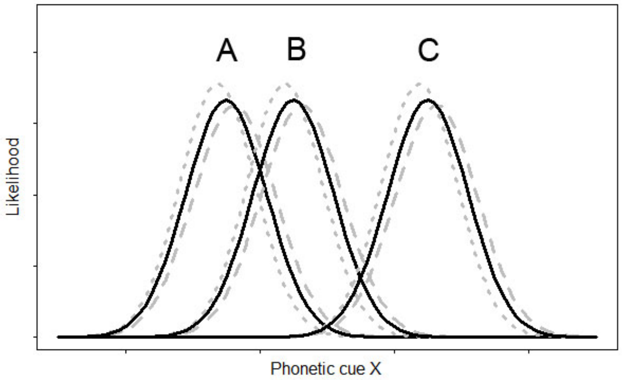

Figure 2: Distributions of Categories A, B, and C along some phonetic cue dimension X. All share the same L/R ratio and mean variance.

minimally with other contrastive categories in the perceptual space. On the other hand, a narrowly distributed category might be highly prone to sound change if it has a high degree of overlap with other contrastive categories in the system. Consider, for example, Figure 2, which shows the distributions of three categories, A, B, and C, in the presence of precursor k. While these categories exhibit identical distributions with respect to cue X (i.e. they have the same L/R ratio and variance; they only differ in their means), they nonetheless have very different phonologization profiles: categories A and B are prone to undergo sound change under precursor k, while Category C is less likely to do so. The logic behind these conclusions follows from the way these distributions overlap with each other. From the perspective of a listener-misperception view of sound change, the likelihood of sound change is determined by the likelihood of misperception prompted by the presence of some phonetic precursor. The more overlap there is between two category distributions, the more uncertain and less accurate the recognition of the two categories are going to be. From this point of view, then, the likelihood of sound change amounts to the degree of overlap between a set of category distributions along some perceptually relevant phonetic dimensions. The robustness of a phonetic precursor is therefore a measure of the degree of confusion induced by the presence of that precursor. That is, the mean and variance of a category might shift as a result of the presence of a phonetic precursor. Such a shift may result in an increase or decrease of overlap with other sound categories along the same phonetic dimension. The larger the precursor-induced shift toward the direction of greater overlap, the more robust this precursor is going to be. Thus given that categories A and B show much overlap under precursor k, we expect these categories would more likely undergo sound change in the presence of k than category C, which has only a small overlap with categories A and B.

In sum, an objective measure of phonetic precursor robustness ideally should be feature- or cue-independent and should offer a way to capture the degree of confusion between sounds in the presence of a phonetic precursor. One way to evaluate the relative robustness of a phonetic precursor is to ascertain the likelihood one sound being confused with another in a given context. That is, a means to evaluate intra-contextual variation is needed. For example, to assess the magnitude of anticipatory coarticulation /i/ has on /ɑ/, it is imperative to determine the likelihood of /ɑ/ being confused with a higher vowel, say /e/, in the same context. Likewise, to evaluate the robustness of /i/-lowering due to the presence of a following /ɑ/, it is useful to know how likely /i/ is confused with /e/ in that context. The most direct way of arriving at this estimate is to conduct a perceptual study. For example, Beddor et al. (2007) asked English and Shona speakers to classify an /ɑ/∼/e/ continuum embedded in contexts of __C/ɑ/ and __C/i/. The magnitude of the effect of vocalic context on the classification of /e/ or /ɑ/ can be evaluated by examining the resulting identification function, the canonical S curve. In particular, the relative strengths of different contexts have on sound categorization can be assessed by comparing the slope at the crossover point across contexts (e.g., the point where the probability of a signal is maximally ambiguous between categories). Clayards et al. (2008), for example, show that listeners are sensitive to the entire probability distribution of acoustic-phonetic cues for a sound category and the precision or amount of certainty about a category that a particular cue provides is inversely proportional to the variance of that cue for that category. That is, the more overlap in the probability distributions of two categories, the more uncertainty the listeners will have about which category they are hearing.

To be sure, cross-linguistic perception studies are few. Perceptual studies of vowel identification in the context of different following obstruents with varying voicing specification are even fewer, if any exist. Thus, a meaningful cross-linguistic phonetic survey based on existing perceptual studies is not yet feasible. However, recent rational models of human cognition have shown that human perceptual behaviors can be insightfully modeled as an "ideal observer". Building on this line of reasoning, I propose in the next section a method for evaluating phonetic precursor robustness even in the absence of perceptual data; human perceptual responses may be estimated given certain assumptions about human perceptual behavior and the probability distributions of the cues of the target categories.

3 Phonetic precursor robustness as degree of uncertainty

Optimal rational listener models of speech perception (Feldman and Griffiths 2007; Clayards 2008; Clayards et al. 2008; Flemming 2010) assume that decisions about perceptual information are guided by certain basic principles that optimize signal identification. To illustrate what a rational analysis might look like, let us consider a scenario where the listener has to decide whether some signal S in context k belongs to some category, say c_1 or c_2, as in a canonical two-alternative forced choice task. Adopting Sonderegger and Yu's (2010) model of perceptual compensation for coarticulation, itself an elaboration of Feldman and Griffiths 2007 and Feldman et al. 2009, listeners are assumed to perceive a noisy stimulus S that is normally distributed around a target pronunciation T as listeners cannot recover the target pronunciation T directly. The target pronunciation T is itself normally distributed around a category mean. Formally,

$$T|c_i, k \sim N(\mu_{c_i,k}, \sigma_c), \quad S|T, c_i \sim N(T, \sigma_S)$$

where $\mu_{c_i,k}$ is the mean of category i in context k, σ_C^2 is the variance in T around the category mean, and σ_S^2 is the variance in S around T. For the sake of simplicity, the variance of the category σ_C^2 and the variance of the signal σ_S^2 are assumed to be the same for categories 1 and 2.

The ideal optimal listener categorizes based on the likelihood of S being an instance of the speaker producing an example from c_i in context k, with target T. The probability of S coming from category c_1 can be calculated with Bayes' rule:

$$P(c_1|S, k) = \frac{P(c_1|k)P(S|c_1, k)}{P(c_2|k)P(S|c_2, k) + P(c_1|k)P(S|c_1, k)} \tag{1}$$

As shown in Equation 1, $P(c_i|k)$ is the probability of category i occurring in context k, i.e. in the lexicon as a whole. The $P(S|c_i, k)$ are calculated by integrating over all possible target T, giving the Equation in 2.[1]

---

[1] Perceptual compensation for coarticulation is treated in Sonderegger and Yu 2010 as a consequence of the variability in the differences between the means, variances, and/or frequencies of two categories in a given context. For example, if the mean F1 shifts downwards for both categories in context k_i compared to context k_j, then the perceptual boundary between the two categories in context k_i would also shift downwards relative to the perceptual boundary in context k_j. One of the reviewers noted that such a model of perceptual compensation for coarticulation might not be compatible with the approach of sound change Moreton implicitly assumes, that is, new sound variant arises when unsophisticated listeners, such as children and second language learners, fail to properly perceptual compensate for coarticulation (Ohala 1981, 1983, 1989, 1992, 1993, 1995). To fully address this matter would take the present discussion too far afield. However, briefly, there are reasons to think that this is not an insurmountable problem. To

(2) $$P(c_1|S,k) = \left(1 + \frac{f_2}{f_1}\, e^{b-Sg}\right)^{-1}$$

where $f_i = P(c_i|k)$ is the frequency of category i in context k and the variables b and g are defined as follows:

$$b = \frac{1}{2}\frac{\mu_{c_1,k}^2 - \mu_{c_2,k}^2}{\sigma_S^2 + \sigma_c^2}, \quad g = \frac{\mu_{c_1,k} - \mu_{c_2,k}}{\sigma_S^2 + \sigma_c^2}$$

Given that the confusability between two sounds is characterized by the nature of the identification function of these sounds, an objective measure of the nature of the identification function between two sounds is also a measure of the confusability of these sounds. With the abovementioned rational model in mind, I propose to quantify context-induced confusability in terms of the normalized slope of the identification function (3).[2]

(3) $$\text{Slope of Precursor Robustness (SPROB)} = \frac{(\mu_{c_1,k} - \mu_{c_2,k})^2}{4(\sigma_S^2 + \sigma_c^2)}$$

Thus, the magnitude of perceptual confusion in the identification of category A relative to category B under the influence of Precursor 1 and Precursor 2 (i.e. the relative robustness of Precursor 1 and Precursor 2) is determined by the respective SPROB scores of the $A \sim B$ identification functions in Context 1 and in Context 2. The lower the SPROB score a precursor has, the more uncertainty the precursor introduces to the discrimination between A and B. The more uncertainty there is between A and B in the presence of the precursor, the more robust the precursor is.

To further clarify the relationship between variance and slope and their effects on the

---

begin with, the idea that only unsophisticated listeners might participate in sound change is empirically dubious; recent experimental evidence have suggested that the sound system might change throughout an individual's life time (Sancier and Fowler 1997; Harrington et al. 2000; Sankoff 2004; Harrington 2006; Evans and Iverson 2007). Even if unsophisticated listeners were contributors to change, there are reasons to believe that differences in perceptual compensation responses by unsophisticated listeners such as children may be experientially-driven. Children have been shown to adjust the range of cues they attend to as they develop (Nittrouer 2002; Nittrouer and Lowenstein 2009); children initially weight dynamic, spectral cues (i.e., formant transitions) more than adults, and weight stable spectral cues (e.g., noise spectra) less (Mayo et al. 2003; Nittrouer 1992; Nittrouer and Miller 1997a,b; Nittrouer and Studdert-Kennedy 1987; Watson 1997). Finally, recent studies have suggested that the role of perceptual compensation for coarticulation in sound change might be better conceptualized in terms of across-individual differences than in terms of within-individual variation. That is, the introduction of stable new variants into a language might be better conceptualized as resulting from systematic differences in perceptual and production norms of individuals with different perceptual compensation strategies, rather than from occasional misperception by individuals (Yu 2010, pear).

[2] Normalization is accomplished by multiplying the actual slope of the identification function with the differences between the means of the two categories ($\mu_{c_1,k} - \mu_{c_2,k}$). This normalization feature allows the SPROB score to be dimensionless. Thanks to Morgan Sonderegger for suggesting this method of normalization.

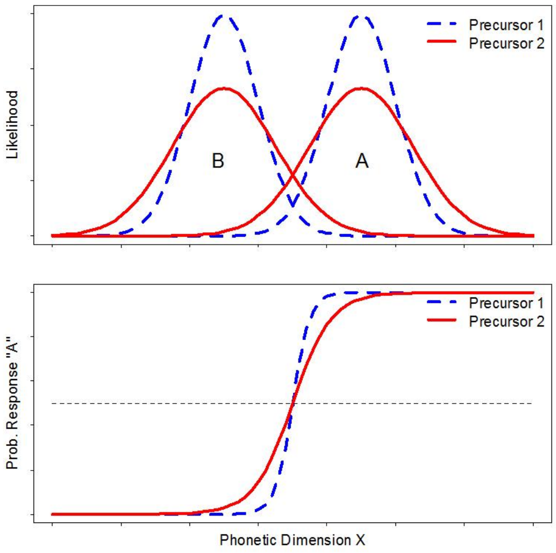

Figure 3: The top panel shows the probability distributions of categories A and B along some phonetic dimension X in the contexts of two phonetic precursors: the variance of the cue distributions under Precursor 1 is narrow (dashed lines); the variance under Precursor 2 is wide (solid lines). The bottom panel shows the optimal response curves calculated from the probability distribution using Eq. (2) for the narrow condition (dashed line) and the wide condition (solid line). The horizontal line in the bottom panel indicates the crossover point; SPROB measures the normalized slope at that crossover point.

interpretation of phonetic precursor robustness, consider Figure 3. The top panel of Figure 3 shows hypothetical probability distributions for two vowel categories along some phonetic dimension X under the influence of two phonetic precursors, 1 and 2. The means of the distributions are the same distance apart but the variances differ. The dashed lines correspond to some sound categories, such as vowels, which are produced consistently and thus have narrow distributions under the influence of Precursor 1; the solid lines correspond to the same sound categories under the influence of Precursor 2 where the sounds are produced less consistently with wide distributions. The optimal solution for each pairs of distributions in the top panel of Figure 3 is illustrated in the bottom panel of the same figure. Even though for both solutions the category boundary (the point where the categorization of A and B is at chance) is in the same place along the x-axis, distributions with greater degree of overlaps (solid lines) have higher degree of uncertainty in category identification, as reflected in the shallowness of slope of the categorization function.

Here, I propose to measure the strength of a phonetic precursor (i.e. the likelihood of a phonetic precursor resulting in sound change) in terms of SPROB introduced in Eqn. 3, which measures the degree of uncertainty (quantified in terms of the normalized slope (3)) engendered by the intra-context variation inherent in the signal. The lower the SPROB, the more robustness the phonetic precursors since a lower SPROB score corresponds to a higher degree of uncertainty. The main idea motivating this approach is the hypothesis that an increase in uncertainty in perception leads to an increase in the likelihood of listener misperception-driven sound change (i.e. the strength of the channel bias).

To better illustrate how this estimate is assessed and its application to understanding relative phonetic precursor robustness, I conducted a cross-linguistic production studies contrasting the effects of anticipatory vowel-to-vowel height coarticulation and anticipatory interaction between vowel height and consonantal voicing.[3] The relative robustness of these phonetic precursors is evaluated in terms of the uncertainties of /i/ and /e/, on the one hand, and /ɑ/ and /e/ on the other, in the context of a following obstruent that is either voiced or voiceless (the HV precursors) and a following vowel that is either /ɑ/ or /i/ (see (4) for a summary).

(4) Summary of the contexts and vowel targets investigated in the production experiment.

Vowel pairs	HH context	HV context
/i/ ∼ /e/	__Ca	__[+voi] or __[-voi]
/ɑ/ ∼ /e/	__Ci	__[+voi] or __[-voi]

The results of the production study serve as the basis for estimating the SPROB scores for the HH and HV phonetic precursors in English and Turkish. Given that previous studies have shown that obstruent voicing has the effect of lowering the F1 value of a preceding vowel, /e/ is expected to encroach on the vowel space of /i/ and /ɑ/ to encroach on the vowel space of /e/. Given the F1-raising effect of /ɑ/ on the preceding vowel, /i/ is expected to encroach on the vowel space of /e/ and /e/ to encroach on the vowel space of /ɑ/. Likewise, the F1 values of /ɑ/ and /e/ should fall in the context of /i/. Thus, if the robustness of HH is stronger than HV, the SPROB scores for both pairs of vowels in the HH contexts should be lower than the same pairs of vowels in the HV contexts. Simply put, vowels in HH contexts should behave more like the wide variance distributions (solid lines) in Figure 3, while the same vowels in HV context should resemble the narrow variance distributions (dashed lines).

---

[3] While it would have been instructive to apply the present model to the phonetic studies reviewed in Moreton's phonetic survey, this is unfortunately not tenable given that, first, only English has been studied for both vowel-to-vowel coarticulation as well as height-voice interaction, and second, the relevant studies did not report all the parameters (e.g., standard deviations) needed by the present model to calculate the SPROB scores.

4 A production study

The following study examined the effect of anticipatory vowel-to-vowel height coarticulation (HH) against the effect of consonantal voicing and vowel height interaction (HV). Previous studies on vowel-to-vowel coarticulation in English reported greater carry-over effects, particularly for unstressed vowels. Using CVCV reiterant speech, Majors (1998), for example, reported that unstressed vowels are subject to greater V-V influence in English than stressed ones. This stress asymmetry is subject to vowel-specific variation; the stress asymmetry is significant for /i/ but not for /o/. Focusing on the production of bVˈbVbV strings where V = /i e a o u/, Beddor et al. (2002) also found more extensive carry-over V-to-V coarticulation than anticipatory effects. While coarticulation is strongest at vowel edges, statistically significant anticipatory coarticulation is also found in English when the vowel is unstressed. With respect to V-to-V coarticulation in Turkish, Beddor and Yavuz (1995), which focused on coarticulation of vowel frontness, found that anticipatory coarticulation is more consistent than carryover coarticulation in Turkish. In particular, they found anticipatory and carryover fronting of /a/ with neighboring /i/ and anticipatory backing of /i/ with following /a/ and slight dissimilation with respect to preceding /a/. Beddor and Yavuz (1995) speculate that the reason why anticipatory coarticulation is the more consistent pattern than carryover effect might have to do with the word-final stress pattern of Turkish in light of the fact that stress vowels tend to exert a greater coarticulatory influence on neighboring vowels than unstressed ones do. Inkelas et al. (2001), which expanded on Beddor and Yavuz (1995)'s study by including stress as a factor of vowel-to-vowel coarticulation in Turkish, replicated Beddor and Yavuz (1995)'s results but found no stress effects. To the best of the knowledge of this author, there is no existing study that investigated HV effects in Turkish. Previous studies on HV effects in English have focused on the effect of final obstruent voicing in CVC strings. For example, van Summers (1987), who focused on bVC syllables where V = /a/ or /ae/ and the final consonant = /b p v f/, reported higher F1 frequencies throughout the vowel (i.e. during the onset, steady-state, and offset) when the final consonant was voiceless relative to utterances with a final voiced consonant. Hillenbrand et al. (1984) found that CVC stimuli identified at least 75% of the time as ending in a voiceless consonant tend to have higher F1 offset frequency than stimuli identified at least 75% as ending in a voiced consonant. This effect is much larger for the vowels with higher first formant frequency; F1 offset frequency difference ranged from 76 to 258 Hz when the vocalic context is /ɛ/ or /a/ but only 17 Hz for /i/ and 23 Hz for /u/. The nature and extent of coarticulation is language-specific. The range of phonemic contrasts within a language, for example, restricts the magnitude of variability induced by coarticulation (Manual 1990; Beddor et al. 2007);

the greater the number of phonemic contrast along an acoustic-perceptual dimension, the less coarticulatory variability is observed. Thus given the small range of height contrast in Turkish than in English, HH and HV coarticulatory effects in Turkish is expected to be more pronounced than in English.

4.1 Methods

The stimuli consist of nonsense trisyllabic strings of the form /ˈɑdV_2CV_3/ where C is either /p/ or /b/ while V_2 and V_3 are /i, e, ɑ/.[4] The three vowels in the stimulus materials were selected as the closest counterparts in the two languages: phonetically, [i, e, a] in Turkish and [i, e, ɑ] in English. A total of eighteen stimuli were constructed. In order to avoid the influence of stress on coarticulation (Beddor et al. 2002), subjects were instructed to accentuate the first syllable of each target word. The medial vowel, V_2, was the target of this analysis.

Six native speakers of North American English (2 female and 4 male) and four native speakers of Turkish (1 female and 3 males) were recruited to take part in the experiment. The English speakers were linguistics graduate students at the University of Chicago with some phonetic training and the Turkish speakers are graduate students at the University of Chicago with no linguistic training. Following Beddor et al. (2002), phonetically trained native English speakers were chosen to avoid difficulties with English orthography and excessive vowel reduction and diphthongization of the mid vowel /e/. Subjects were paid a nominal fee for participation. Speakers were recorded reading a randomized list of test stimuli ten times. The target trisyllables (given in IPA transcription in angled brackets [ ]) were embedded in a sentence context, English "Give me an ___ tomorrow" and Turkish "açıkça ___ söyle" ("clearly ___ say"). The test stimuli were displayed on a computer screen one at a time using a Java script. Subjects were instructed to read aloud the displayed sentence in a normal speed. English subjects were additionally instructed to read the target trisyllables as closely to their IPA values as possible. Turkish subjects were instructed to read the target trisyllables the way they would sound in Turkish. The utterances were recorded in a sound-proofed room and were digitized directly at 48 kHz using an Ediral digital recorder. F1 values were measured using an LPC algorithm implemented in PRAAT.[5] Measurements were taken

---

[4] One reviewer pointed out that the use of potentially phonotactically illegal strings here might present difficulties with the native English subjects especially since vowels are often reduced in unstressed positions. This concern is mitigated by the fact that this is a cross-linguistic study. Given that the main focus of this research is in uncovering the nature of phonetic precursors across languages, to the extent that coarticulatory effects are observed even in phonotactically illegal strings in the native language of the speaker, it only strengthens the universal basis of such effects.

[5] Perceived vowel height is influenced by multiple acoustic-phonetic factors (e.g., F0 and duration). However, to ensure maximal comparability with Moreton's original study, the present study following Moreton's

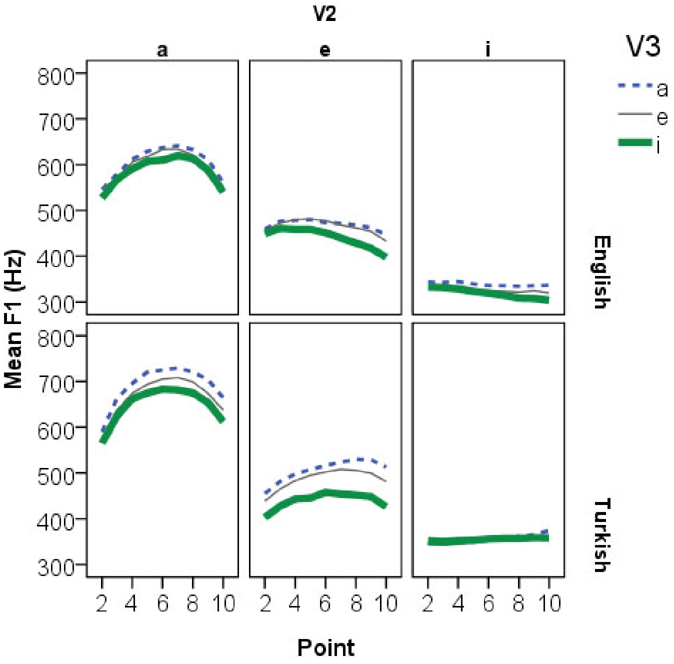

Figure 4: F1 values of /ɑ/, /e/, /i/ preceding /ɑ/, /e/, or /i/ from speakers of English and Turkish.

from eleven equidistant points of the target vowel, starting from the onset of the vowel (i.e. immediately after the commencement of voicing) measuring at ten percentage increment of the total duration of the vowel. Measurements at the eleventh measure point were taken at the beginning of stop closure. Only the values from Point 2 to Point 10 were considered in the statistical analysis to avoid errors introduced by closure boundaries. Outliers, defined as F1 values more than two standard deviations from the means of their condition, were discarded from further analysis. The linear formant frequency scale was converted to a Bark scale (Traunmller 1990). Statistical analyses were performed on the Bark-transformed formant frequencies.

4.2 Results

Figures 4 and 5 summarize the results of the production study (See Appendix for a detailed summary). The effect of vowel-to-vowel coarticulation on F1 is given in Figure 4 while the

---

lead and focuses on just F1 values. However, it is worth noting that the same SPROB calculation can be conducted for all relevant acoustic-phonetic cues.

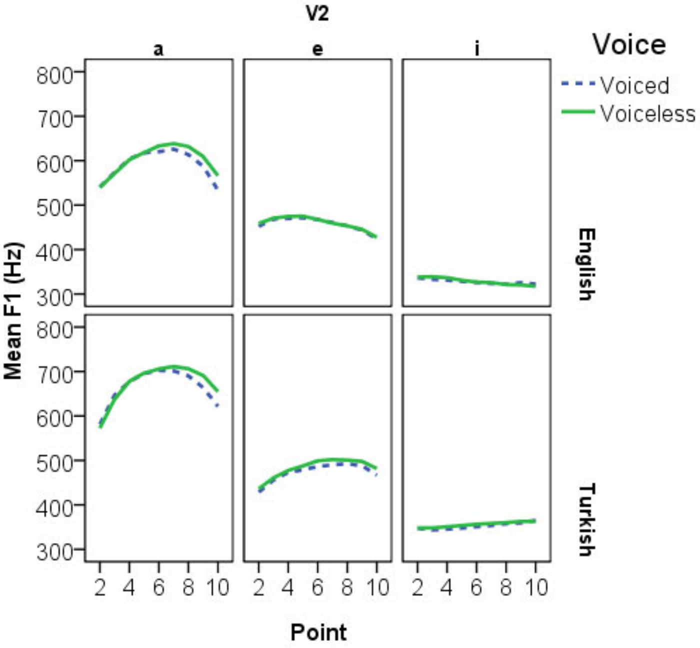

Figure 5: F1 values of /ɑ/, /e/, /i/ preceding [±voice] obstruents from speakers of English and Turkish.

nature of the plosive voicing-F1 interaction is presented in Figure 5. For ease of reference, the F1 values are given in Hz rather than in the Bark-transformed scale. Visual inspection confirms that the subjects indeed articulated three vowels with distinct vowel height, suggesting that the subjects were able to accomplish the task as instructed.

Figure 4 shows that the F1 values of the target vowel is highest when the following vowel is /ɑ/ (dashed line) and the F1 values are low when the following vowel is /i/ (thick solid line). As illustrated in Figure 5, the effect of consonantal voicing on the F1 values of a preceding vowel is less pronounced. The F1-lowering effect of a voiced stop is most apparent when the target vowel is low. Visual inspection also reveals a difference in the temporal dynamics of HH and HV coarticulations. While HH coarticulation is evidenced across the entire target vowel, the HV effect, if any, is evident only in the latter half of the target vowel (i.e. the portion closest to the consonantal trigger). To examine the statistical significance of these differences, a series of four-way repeated measures ANOVA with V2 (/i e ɑ/), V3 (/i e ɑ/), VOICING (Voiced vs. Voiceless) and MEASUREMENT POINT (Point 2 to Point 10) as fixed factors and SUBJECT as the error term was performed for each language group. The statistical results are summarized in Table 2.

		English		Turkish	
	df	F value	p	F value	p
V2	2	26240.028	<0.001	9737.023	<0.001
V3	2	208.851	<0.001	133.910	<0.001
Voice	1	10.301	0.001	14.457	<0.001
Point	8	91.830	<0.001	56.112	<0.001
V2:V3	4	10.429	<0.001	43.038	<0.001
V2:Voice	2	8.364	<0.001	0.622	0.537
V3:Voice	2	3.863	0.021	3.852	0.021
V2:Point	16	55.420	<0.001	15.002	<0.001
V3: Point	16	2.887	0.000	1.080	0.368
Voice:Point	8	0.699	0.693	0.731	0.665
V2:V3:Voice	4	4.989	0.001	0.939	0.440
V2:V3:Point	32	1.166	0.238	0.088	1.000
V2:Voice:Point	16	2.846	<0.001	0.695	0.802
V3:Voice:Point	16	0.423	0.978	0.092	1.000
V2:V3:Voice:Point	32	0.504	0.991	0.158	1.000

Table 2: Results of ANOVA for F1 values (based on 9518 responses from six English speakers and 6050 measurements from four Turkish speakers). P-values below 0.05 are bolded.

As expected, main effects of V3 and Voicing, as well as V2 and Measuring Point, are observed in both the English and Turkish data. Significant interaction between V2 and V3 is observed in both English and Turkish, suggesting that the degree of anticipatory coarticulation varies depending on the quality of the target vowel. Post-hoc analyses (alpha level adjusted to 0.017 for 3 comparisons) shows that, with the sole exception of V2 = /i/ in Turkish where the V3 effect does not reach the adjusted alpha level (F(2, 2032) = 3.14, $p = 0.04$)[6], the effect of V3 on all other target vowels in both languages are extremely robust ($p < 0.001$). Turkish shows greater degree of coarticulation than English, at least with respect to the non-high V2. This finding is consistent with the idea that a language with a smaller number of phonemic contrast along an acoustic-perceptual dimension should be more susceptible to coarticulation (Manual 1990; Beddor et al. 2007). The lack of a strong coarticulatory effect when V2 is /i/ in Turkish might be attributed to the effect of the preceding vowel; previous studies show that /i/ dissimilates (i.e. /i/ is higher and fronter rather than lower and backer) when the preceding vowel is /ɑ/ (Beddor and Yavuz 1995; Inkelas et al. 2001). The cause of this dissimilatory effect is not known, although Beddor and Yavuz (1995) speculates that "dissimilation" may actually be lack of centralization when V1

---

[6] The V3 effect on /i/ is trending in the right direction.

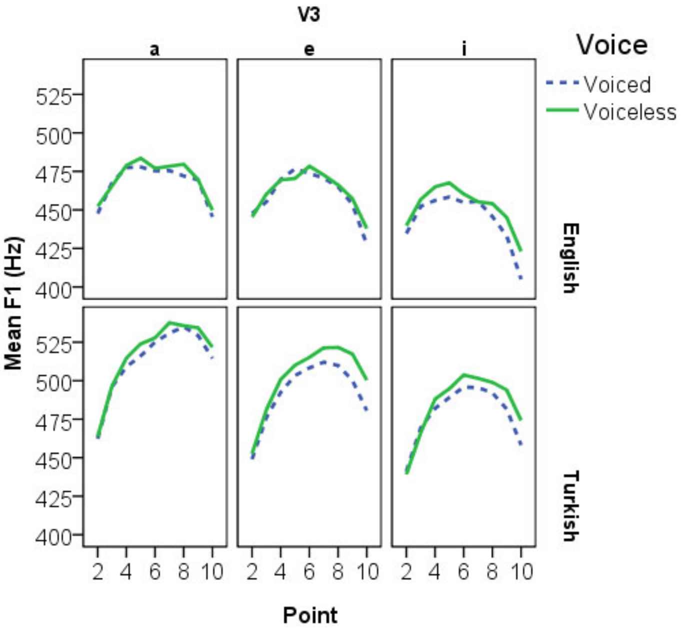

Figure 6: F1 values of V2 under the influence of V3 and the voicing of the following obstruent from speakers of English and Turkish.

≠ V2. In the case of English, significant interactions between V2 and VOICING, and between V2, VOICING and MEASURING POINT suggest that the degree of F1-raising from obstruent voicing depends not only on the target vowel but also varies in time. Post-hoc analyses (alpha level adjusted to 0.017 for 3 comparisons) show the V2-dependent voicing effect to be significant only when V2 = /ɑ/ (F(1, 3113) = 12.23, $p < 0.001$). It is noteworthy that previous studies of HV-effects have mainly focused on low or non-high vowels in English. The fact that the present study shows that no significant HV-effect is observed with /e/ and /i/ further highlights the specificity of the HV effect, suggesting that this might be an additional obstacle to HV-effects being phonologized. One of the more intriguing findings is a significant, albeit weak, interaction between V3 and VOICING in both English and Turkish. Figure 6 shows that F1 values generally trend lower preceding a voiced plosive when a non-low vowel follows the target vowel. However, this V3-dependent voicing effects did not reach statistical significance in post-hoc analyses.

	i∼e			ɑ∼e		
	_Cɑ	_[+voi]	_[-voi]	_Ci	_[+voi]	_[-voi]
English	1.13	1.33	1.25	0.58	0.60	0.66
Turkish	1.32	0.91	0.93	0.50	0.64	0.69

Table 3: SPROB scores for /i∼e/ and /ɑ∼e/ in four vowel raising and lowering contexts.

4.3 Evaluating relative phonetic precursor robustness

4.3.1 Predictions of the SPROB approach

The main goal of the cross-linguistic production study is to assess the relative robustness of HH and HV precursors. To quantify the relative phonetic precursor robustness, the SPROB scores for /i∼e/ and /ɑ∼e/ were calculated using Eqn. 3. Specifically, four contextual effects were considered: /i∼e/ in a lowering context (V3 = /ɑ/), /ɑ∼e/ in a raising context (V3 = /i/) and, /i∼e/ and /ɑ∼e/ followed by a plosive that is voiced (a vowel-raising context) or voiceless (a vowel-lowering context). With respect to V_1 and V_2 in Eqn. 3, in the /i∼e/ comparison, V_1 = /e/ and V_2 = /i/, while in the /ɑ∼e/ comparison V_1 = /ɑ/ and V_2 = /e/. Formant values were based on the midpoint of the target vowel in the specific context (see the Appendix for actual values). Following Sonderegger and Yu 2010, variances for the target vowels (the $\sigma_S^2 + \sigma_c^2$ term in Eqn. 3) were taken to be the mean of the variances for the target vowels in a given context. For example, for /i∼e/ in a lowering context (V3 = /ɑ/), the variances would be the mean of the variances for the /ˈɑdibɑ/ and /ˈɑdebɑ/ stimuli. Table 2 summarizes the SPROB results.

Recall that a larger SPROB score corresponds to a steeper category boundary. Thus the shallower the slope, the more categorization uncertainty exists. Table 2 shows that, all else being equal, listener' classification of vowels is more uncertain when the target signal is under HH influence than when under HV influence. This result suggests that the HH precursor is more robust than the HV precursor; the HH precursor is more likely to result in sound change than the HV precursor. The only exception to this generalization is /i∼e/ in Turkish which shows HV environments yielding lower SPROB scores than the vowel-lowering _Cɑ context. This exceptional behavior is to be expected in light of the results of the acoustic study and the general dissimilatory behavior of /i/ when the preceding vowel is /ɑ/ as reported in earlier studies (Beddor and Yavuz 1995; Inkelas et al. 2001). That is, recall that the effect of V3 on /i/ is weak in Turkish. Thus while /e/ is being drastically lowered by the following /ɑ/, little lowering is evidenced in /i/. This leads to an increase separation between /i/ and /e/ before /ɑ/. Thus, all else being equal, a wider separation between distributions of two vowel categories translates into a steeper category boundary slope (i.e. a higher SPROB score). The behavior of /i∼e/ in Turkish is thus the exception that proves the rule. Exceptionalities of this sort highlight the importance of not treating channel bias as a monolithic factor that applies to all languages indiscriminately; the nature and magnitude of coarticulatory effects, and intra-context variation in general, is largely language-specific.

The robustness of the HH precursor over that of the HV precursor extends beyond the SPROB comparison. While the production study shows that anticipatory vowel-to-vowel coarticulation is evidenced in both English and Turkish, the extent of HV interaction is mediated temporally as well as by target vowel quality, at least in the case of English. That is, while anticipatory HV coarticulation dissipates the further away the vocalic gesture is from the following obstruent, anticipatory HH effect shows no such temporal dependency. The locality of the HV effect may further weaken its effectiveness as a phonetic precursor to sound change. The temporal restrictiveness of the HV precursor points to two potential venues for further investigation. To begin with, the phonological effect of HV-interaction might also be localized to the part of the vowel closest to the triggering consonant. For example, voicing in a preceding obstruent may only affect the first portion of a following vowel's height. Such a pattern is observed in the history of the Austronesian language Haroi (Lee 1977; Denning 1989): following a voiced obstruent, *a > ɨa, *au > ɨaʊ, *əi > ɨi; voicelessness in the preceding obstruents, on the other hand, conditioned lowering of part or all of a high vowel (after a voiceless obstruent, *u > ô, *i > êy, e) .[7] Another potential venue for exploring HV effects might be in its interaction with other precursors. Recall, for example, that there is an interaction effect between voicing and V3. Such an interaction may manifest itself in the form of passive resistance in natural language. For example, in a language with height harmony, voiced obstruents may serve as blockers. In the Buchan dialect of Scots, for example, there exists a partial height harmony where an unaccented high vowel in a suffix would agree in height with the preceding stressed root vowel unless the intervening consonant is either a voiced consonant or [l m n ŋ] followed by voiceless obstruents (Dieth 1932; Fitzgerald 2002; Paster 2004). Thus, *messy* is pronounced [mɛse] and *mealie* as [mili], but *doggie* as [dogi] and *bendy* as [bɛndi] (Paster 2004: 365-366). Assuming that partial height harmony in Buchan Scots originated from vowel-to-vowel coarticulation, the

---

[7] The opposing effect of obstruent voicing on vowel height is consistent with the fact that obstruent voicing lowers F1, hence raises vowel height, while obstruent voicelessness raises F1, thus lowering vowel height. That post-voiced obstruent raising only affects low vowels might be due to the fact that high vowels already have low F1 to start with. Likewise, the fact that lowering after a voiceless obstruent only targets high vowels presumably is because low vowels already have high F1.

		Raising	Lowering	L/R ratio
a.	V-to-V:	_Ci	_Cɑ	_Cɑ/_Ci
	English	540.1	560.2	1.04
	Turkish	612.3	664.5	1.09
b.	Voicing:	_VoicedC	_VoicelessC	L/R ratio
	English	533.5	566.2	1.06
	Turkish	621.6	654.8	1.05

Table 4: Vocalic F1 values in Hertz measured at the target vowel offset (Point 10) in different contexts and languages.

local F1-lowering effect of obstruent voicing might have been sufficient to counteract the stressed non-high vowel's effect of F1 raising on the unstressed high vowel of the suffix. This type of interaction between the effect of voicing on F1 and vowel-to-vowel coarticulation is readily evident in the production data presented above. Further investigation is needed to substantiate this type of synergistic interaction between consonantal voicing and vowel height.

4.3.2 Comparison with the L/R ratio approach

It is instructive at this point to return to Moreton's original proposal and to consider what predictions the L/R ratio analysis might make given the present dataset. Following Moreton, the L/R ratios are calculated here based on measurements taken at the target vowel offset. Table 4 shows the L/R ratios for vowels before /ɑ/ or /i/ and before a voiced or voiceless stop in English and Turkish.

The results show that, at least in English, the L/R ratio is higher in the HV context than in the HH context. The opposite pattern, however, is observed in the Turkish case. Even though the L/R ratios based on the English data appear to show that the HH precursor is less robust than the HV precursor while the Turkish results point to a more robust HH precursor than an HV precursor, the magnitude of difference in L/R ratios across precursor contexts are very small (the average L/R ratios for HH and HV precursors are 1.061 and 1.057 respectively), especially relative to the differences observed in Moreton's survey (the average L/R ratios for HH and HV precursors in Moreton's study are 1.065 and 1.207 respectively). The results of the L/R ratio analysis would have led to the conclusion that the precursors of HH and HV interactions are comparable in robustness and thus do not support the hypothesis that relative precursor robustness predicts typological frequency of phonologization. This conclusion is contradicted by the overall prediction of the SPROB analysis.

This difference in conclusion is to be expected since Moreton's approach only measures a category's variation across contexts. If a category varies approximately the same amount in both HH and HV contexts, the L/R ratio analysis would have suggested no differences in phonetic precursor robustness. The SPROB approach, on the other hand, emphasizes the effects context-induced variation have on between-category confusion. Thus, even though the HH and HV precursors might induce the same amount of variation with respect to one category, their effects might not be the same with respect to another category. The amount of potential confusion generated by a phonetic precursor might therefore not be identical across sound categories.

5 Conclusion

This paper begins with a discussion of the inadequacy of Moreton (2008)'s L/R ratio as a measure of phonetic precursor robustness. A method of evaluating the robustness of phonetic precursors, called SPROB, was proposed. This parameter, derived from a rational model of speech perception and production, offers a means to estimate listeners' perceptual response to intra-contextual variation, even when actual perceptual data from human listeners is absent. Based on the results of a cross-linguistic production experiment and the corresponding SPROB comparison, I show that the HH precursor is more robust than the HV precursor. As such, a channel bias explanation for the underphonologization of HV precursor remains very much plausible.

To be sure, the viability of a channel bias account of the observed underphonologization between the HH and HV precursors does not obviate a possible contribution of analytic bias in shaping this typology. In fact, I would submit that the type of rational model proposed in this and other work can be viewed as a type of analytic bias since rational models assume humans perform optimally and rationally in cognitive tasks. Such models are also not modular-specific in that they have been invoked to account for problems in other domains of human cognition beyond language. In a very real sense, the model advocated in this paper provides a natural framework for bridging the divide between a purely experience-based approach to the emergence of sound pattern and a formal/cognitive approach to sound pattern since the parameters the rational model depends on are experience-driven. The outcome of categorization is just as much affected by the model of optimization selected as it is by the distribution of cues assumed by the speaker-listener who performs the categorization task.

Naturally, there remains the issue of whether analytic bias can be diagnosed apart from channel bias. As Moreton painstakingly explains, the contribution of analytic bias is best revealed when the confound of channel bias is ruled out (this is the basic logic behind all underphonologization research). The best evidence for the independence of analytic bias from channel bias would come from a case that allows the double-dissociation between these two biases. The present study argues that the HH/HV asymmetry does not rise up to the challenge of a double-dissociation test. Some might point to the results of the artificial grammar learning experiment as evidence for hints of an analytic bias at work. There are, however, reasons to be cautious, at least with respect to the significance of Moreton's experimental results. Moreton constructed an artificial language where the V1 and C2 of C1V1C2V2 words are such that V1 and C2 were high and voiced or non-high and voiceless. Subjects were exposed to samples of this artificial language and were tested subsequently for awareness of the co-occurrence restriction. Participants were found to have a hard time learning the HV patterns. Moreton attributes this difficulty to the learning of cross-featural tier dependencies (i.e. it might be more difficult to learn the dependency between voicing features and height features than the dependencies among height features). The relative difficulty of acquiring feature dependencies notwithstanding, there is a potential confound that is obscuring the interpretation of the experimental findings. As is well known, certain varieties of English possess a phonological rule, generally known as Canadian-raising, whereby a low vowel is raised when preceding a voiceless stop. Despite the name of this pattern, this raising phenomenon is not limited to Canada. Moreton and Thomas (2007), for example, noted that the/ai/∼/ʌi/ alternation has been reported in other parts of North America, including Virginia and adjacent parts of Maryland and North Carolina (Shewmake 1925, 1943, 1945; Greet 1931; Lowman 1936; Tresidder 1941, 1943; Dorrill 1986:86), and coastal South Carolina and Georgia (McDavid 1955; Kurath and McDavid 1961). As such, it cannot be discounted that students at the University of North Carolina at Chapel Hill, the subject population of Moreton's artificial grammar learning experiment, might either possess Canadian Raising in their phonologies or are at least aware of such an alternation in other varieties of English. If correct, the fact that subjects in Moreton's experiment exhibit difficulties in learning HV-dependencies might be attributed to the fact that they were asked to learn a pattern that conflicts with either the phonology of their own language or the phonology of some neighboring English varieties they are aware of. This contradiction might be sufficient to hinder the learning of the HV patterns to such a degree that the subjects yielded a lower success rate in learning HV patterns than in learning the HH patterns.

A reviewer of this article suggests that the strength of the HH precursor itself might have originated from the effect of analytic bias. That is, coarticulatory effects may be seen as the application of phonetic rules in the language rather than as mechanical side effects that fall out as a consequence of the coproduction of various articulators (Whalen 1990;

Kingston and Diehl 1994). Such an interpretation is particularly apt under models of the phonetics-phonology interface that eschew any differences between phonetics and phonology and propose instead to account for both the categorical and gradient patterns within the same grammatical system (Flemming 1995, 1996, 2001a,b; Ní Chiosáin and Padgett 2001; Padgett 2003). The idea that vowel-to-vowel coarticulation as well as the effects of voicing on F0 and F1 are part of the phonetic knowledge of the speaker and are, if only partially, under the control of the speaker is certainly not new (Whalen 1990; Kingston and Diehl 1994; Scarborough 2004). However, if phonetic knowledge is to be classified under the purview of analytic bias, then it is not clear to what extent it is meaningful to differentiate channel bias from analytic bias in the first place. To what extent, for example, should the type of rational model advocated in this and other work be conceptualized in terms of channel or analytic bias, particularly in light of the applicability of Bayesian rational models to many domains of human cognition beyond language, as alluded to earlier? Given that this paper aims at addressing the more narrow issue of the proper measurement of the robustness of phonetic precursor as conceptualized within Moreton's framework, I shall take refuge from the fact that Moreton himself took vowel-to-vowel coarticulation and the interaction between voicing and F1 as part of the channel bias and leave the ontological question of where to draw the channel bias and analytic bias divide to another occasion.

Finally, it is worth mentioning that the SPROB approach, as it is described here, admittedly has its limitation since it only provides an expression of phonetic precursor robustness in the context of structure-preserving changes (i.e. the reassignment of ambiguous phonetic signals to *existing* phonological categories) as it requires the comparisons of cues between two already-established sound categories. In the case of structure-building changes (i.e. the emergence of new allophones or phonological categories based on some ambiguous phonetic signals), the question of phonetic precursor robustness is no longer about confusion and reassignment between categories, but about inducing a distinct allophonic or phonemic category out of a set of context-specific cues. Without taking the discussion too far afield, it is worth noting that there has been a host of recent work applying computational techniques of clustering and pattern recognition to the problem of phonological category induction (de Boer and Kuhl 2003; Lin 2005; Vallabha et al. 2007; Feldman et al. 2009; McMurray et al. 2009; Kirby 2010, pear). Kirby (2010, pear), for example, illustrates a promising method of phonetic category induction using a model-based clustering approach. He models phonetic categories as Gaussian mixtures of cue distributions and shows that, given a set of phonetic cues, a two-category solution (i.e. the positing of two allophones or two phonological categories) may be justified under certain combination of cues; other cue combinations would lead to complete merger. To be sure, further research is needed to ascertain whether such a model-based clustering approach is adequate for inducing novel allophones in the context of a structure-building phonetic precursor. The study of the robustness of phonetic precursors is only in its early years. As formal models of speech perception, such as the rational model promoted here, are developed and refined, better and more sophisticated methods of measuring precursor robustness will no doubt emerged.

In sum, channel bias, i.e. the relative robustness of phonetic precursors, remains a strong contender in explaining the underphonologization of HV patterns over HH patterns. To be sure, the ideal of isolating analytic bias by controlling for potential channel bias is a tall order. Moreton is to be saluted for laying out an interesting approach to this thorny issue and for attempting to offer concrete evidence for analytic bias' role in shaping the typology of sound patterns. Nevertheless, at least in the case of the underphonologization of HV over HH patterns, irrefutable evidence for analytic bias remains elusive.

6 Appendix

Summary of F1 values measured at the mid point of V2 in /ˈɑdV2CV3/ sequences.

			English		Turkish	
V2	V3	Voice	Mean (Hz)	SD	Mean (Hz)	SD
a	a	Voiced	629	104	733	85
		Voiceless	645	94	718	103
		Total	637	99	725	94
	e	Voiced	630	105	693	116
		Voiceless	637	89	718	80
		Total	634	97	705	100
	i	Voiced	603	106	684	139
		Voiceless	618	124	683	123
		Total	610	115	684	130
	Total	Voiced	620	105	702	117
		Voiceless	633	103	706	105
		Total	627	104	704	111
e	a	Voiced	479	73	514	81
		Voiceless	468	71	519	79
		Total	473	72	516	80
	e	Voiced	474	68	495	64
		Voiceless	481	78	509	60
		Total	478	73	502	62
	i	Voiced	449	58	450	104
		Voiceless	454	63	467	117
		Total	451	60	458	110
	Total	Voiced	467	68	486	88
		Voiceless	468	71	499	89
		Total	468	69	493	89
i	a	Voiced	336	43	351	37
		Voiceless	336	42	356	37
		Total	336	42	353	37
	e	Voiced	328	37	348	31
		Voiceless	322	38	355	39
		Total	325	38	351	35
	i	Voiced	315	33	355	41
		Voiceless	323	36	358	44
		Total	319	35	356	42
	Total	Voiced	326	39	351	37
		Voiceless	327	39	356	39
		Total	327	39	354	38

References

Beddor, P., Brasher, A., and Narayan, C. (2007). Applying perceptual methods to the study of phonetic variation and sound change. In Solé, M.-J., Beddor, P., and Ohala, M., editors, *Experimental Approaches to Phonology*, pages 127–143. Oxford University Press, Oxford.

Beddor, P. and Yavuz, H. (1995). The relation between vowel-to-vowel coarticulation and vowel harmony in turkish. In *Proceedings of the International Conference of the Phonetic Sciences 1995 Vol. 2*. 44-51.

Beddor, P. S., Harnsberger, J., and 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: the emergence of sound patterns.* Cambridge University Press, Cambridge.

Blevins, J. (2006). A theoretical synposis of Evolutionary Phonology. *Theoretical Linguistics*, 32:117–165.

Clayards, M. (2008). *The ideal listener: making optimal use of acoustic-phonetic cues for word recognition.* PhD thesis, University of Rochester.

Clayards, M., Michael K. Tanenhaus, R. N. A., and Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition*, 108:804–809.

de Boer, B. and Kuhl, P. (2003). Investigating the role of infant-directed speech with a computer model. *Acoustic Research Letters On-line*, 4(4):129–134.

Denning, K. (1989). *The diachronic development of phonological voice quality, with special reference to Dinka and the other Nilotic languages.* PhD thesis, Stanford University, Palo Alto.

Dieth, E. (1932). *A grammar of the Buchan dialect (Aberdeenshire), vol. I, phonology-accidence.* W. Heffer & Sons, Cambridge.

Evans, B. and Iverson, P. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *Journal of the Acoustic Society of America*, 121:38143826.

Feldman, N. and Griffiths, T. (2007). A rational account of the perceptual magnet effect. In McNamara, D. and Trafton, J., editors, *Proceedings of the 29th Annual Cognitive Science Society*, pages 257–262. Cognitive Science Society, Austin, TX.

Feldman, N. H., Griffiths, T. L., and Morgan, J. L. (2009). The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological review*, 116(4):752–782.

Fitzgerald, C. M. (2002). Vowel harmony in buchan scots english. *English Language and Linguistics*, 6:6179.

Flemming, E. (1995). *Auditory representations in phonology.* PhD thesis, University of California, Los Angeles, Los Angeles, CA.

Flemming, E. (1996). Evidence for constraints on contrast: The dispersion theory of contrast. *UCLA Working Papers in Phonology*, 1:86–106.

Flemming, E. (2001a). *Auditory representations in phonology.* Garland Press, New York.

Flemming, E. (2001b). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology*, 18(1):7–44.

Flemming, E. (2010). Modeling listeners: Comments on pluymaekers et al. and scarborough. In Fougeron, C., Kühnert, B., D'Imperio, M., and Vallée, N., editors, *Papers in Laboratory Phonology 10*, pages 587–606. Mouton de Gruyter, Berlin, Germany.

Guion, S. (1996). *Velar palatalization: Coarticulation, perception, and sound change.* PhD thesis, University of Texas, Austin, Austin, TX.

Hale, M. and Reiss, C. (2000). Phonology as cognition. In Burton-Roberts, N., Carr, P., and Docherty, G., editors, *Phonological knowledge: Conceptual and empirical foundations*, pages 161–184. Oxford University Press, Oxford.

Harrington, J. (2006). An acoustic analysis of 'happy-tensing' in the queen's christmas broadcasts. *Journal of Phonetics*, 34:439–457.

Harrington, J., Palethorpe, S., and Watson, C. (2000). Monophthongal vowel changes in received pronunciation: an acoustic analysis of the queen's christmas broadcasts. *Journal of the International Phonetic Association*, 30:63–78.

Hillenbrand, J., Ingrisano, D. R., Smith, B. L., and Flege, J. E. (1984). Perception of the voiced-voiceless contrast in syllable-final stops. *Journal of the Acoustical Society of America*, 76:18–27.

Hume, E. and Johnson, K. (2001). A model of the interplay of speech perception and phonology. In Hume, E. and Johnson, K., editors, *The role of speech perception in phonology*, pages 3–26. Academic Press, San Diego.

Hyman, L. (1976). Phonologization. In Juilland, A., editor, *Linguistic studies presented to Joseph H. Greenberg*, pages 407–418. Anma Libri, Saratoga.

Hyman, L. (2001). On the limits of phonetic determinism in phonology: *nc revisited. In Hume, B. and Johnson, K., editors, *The Role of Speech Perception Phenomena in Phonology*, pages 141–185. Academic Press.

Inkelas, S., Barnes, J., Dolbey, A., Good, J., Hansson, G., Kavitskaya, D., Orgun, O., Sprouse, R., and Yu, A. (2001). The effects of stress on vowel-to-vowel coarticulation in turkish. In *International Conference on Turkish Linguistics*, Istanbul, Turkey.

Kingston, J. and Diehl, R. L. (1994). Phonetic knowledge. *Language*, 70(3):419–454.

Kiparsky, P. (2006). The amphichronic program vs. evolutionary phonology. *Theoretical Linguistics*, 32(2):217–236.

Kiparsky, P. (2008). Universals constrain change; change results in typological generalizations. In Good, J., editor, *Linguistic Universals and Language Change*, pages 23–54. Oxford University Press, Oxford.

Kirby, J. (2010). *Cue selection and category restructuring in sound change. University of Chicago dissertation.* PhD thesis, University of Chicago, Chicago.

Kirby, J. P. (ToAppear). Acquisition of covert contrasts: an unsupervised learning approach. In *Proceedings of the 46th Annual Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society.

Lee, E. W. (1977). Devoicing, aspiration, and vowel split in Haroi: evidence for register (contrastive tongue-root position. In Thomas, D., Lee, E. W., and Liem, N. D., editors, *Papers in South East Asian linguistics No. 4. Chamic studies*, number 4 in Pacific Linguistics Series A, pages 87–104. Australia National University, Canberra.

Lin, Y. (2005). *Learning Features and Segments from Waveforms: A Statistical Model of Early Phonological Acquisition*. PhD thesis, University of California Los Angeles.

Lindblom, B., Guion, S., Hura, S., Moon, S.-J., and Willerman, R. (1995). Is sound change adaptive? *Rivista di Linguistica*, 7:5–36.

Majors, T. (1998). *Stress-dependent harmony: phonetic origins and phonological analysis*. PhD thesis, University of Texas, Austin, Austin.

Manual, S. Y. (1990). The role of contrast in limit vowel-to-vowel coarticulation in different languages. *Journal of the Acoustic Society of America*, 88:1286–1298.

Mayo, C., Scobbie, J. M., Hewlett, N., and Waters, D. (2003). The influence of phonemic awareness development on acoustic cue weighting strategies in childrens speech perception. *Journal of Speech, Language, and Hearing Research*, 46(5):1184–1196.

McMurray, B., Aslin, R., and Toscano, J. C. (2009). Statistical learning of phonetic categories: insights from a computational approach. *Developmental Sciences*, 12(3):369–378.

Moreton, E. (2008). Analytic bias and phonological typology. *Phonology*, 25(1):83–127.

Moreton, E. (2010). Underphonologization and modularity bias. In Parker, S., editor, *Phonological argumentation*, pages 79–101. Equinox Publishing, London.

Ní Chiosáin, M. and Padgett, J. (2001). Markedness, segment realization, and locality in spreading. In Lombardi, L., editor, *Segmental phonology in Optimality Theory.* Cambridge University Press.

Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics*, 20:351–382.

Nittrouer, S. (2002). Learning to perceive speech: How fricative perception changes, and how it stays the same. *Journal of the Acoustical Society of America*, 112:711–719.

Nittrouer, S. and Lowenstein, J. H. (2009). Does harmonicity explain childrens cue weighting of fricative-vowel syllables? *Journal of the Acoustical Society of America*, 125:1679–1692.

Nittrouer, S. and Miller, M. E. (1997a). Developmental weighting shifts for noise components of fricative-vowel syllables. *Journal of the Acoustical Society of America*, 102:572–580.

Nittrouer, S. and Miller, M. E. (1997b). Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America*, 101:2253–2266.

Nittrouer, S. and Studdert-Kennedy, M. (1987). The role of coarticulatory effects on the perception of fricatives by children and adults. *Journal of Speech and Hearing Research*, 30:319–329.

Ohala, J. (1983). The origin of sound patterns in vocal tract constraints. In MacNeilage, P., editor, *The production of speech*, pages 189–216. Springer-Verlag, New York.

Ohala, J. (1989). Sound change is drawn from a pool of synchronic variation. In Breivik, L. E. and Jahr, E. H., editors, *Language Change: Contributions to the Study of its Causes*, pages 173–198. Mouton de Gruyter, Berlin.

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

Ohala, J. J. (1981). The listener as a source of sound change. In C. S. Masek, R. A. Hendrick, . M. F. M., editor, *Papers from the Parasession on Language and Behavior*, pages 178–203. Chicago Linguistics Society, Chicago, IL.

Ohala, J. J. (1992). What's cognitive, what's not, in sound change. In Kellerman, G. and Morrissey, M. D., editors, *Diachrony within synchrony: Language history and cognition*, pages 309–355. Peter Lang Verlag, Frankfurt. Reprinted in Lingua e Stile (1992) 27: 321–362.

Ohala, J. J. (1995). Experimental phonology. In Goldsmith, J. A., editor, *A Handbook of Phonological Theory*, pages 713–722. Blackwell, Oxford.

Padgett, J. (2003). Contrast and post-velar fronting in russian. *Natural language and linguistic theory*, 21(1):39–87.

Paster, M. (2004). Vowel height harmony and blocking in Buchan Scots. *Phonology*, 21(3):359–407.

Sancier, M. L. and Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25:421–436.

Sankoff, G. (2004). Adolescents, young adults and the critical period: two case studies from seven up. In Fought, C., editor, *Sociolinguistic Variation*, pages 121–139. Oxford University Press.

Scarborough, R. A. (2004). *Coarticulation and the structure of the lexicon.* PhD thesis, University of California, Los Angeles, Los Angeles.

Seidl, A. and Buckley, E. (2005). On the learning of arbitrary phonological rules. *Language Learning and Development*, 1:289–316.

Sonderegger, M. and Yu, A. C. L. (2010). A rational account of perceptual compensation for coarticulation. In Ohlsson, S. and Catrambone, R., editors, *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, pages 375–380. Cognitive Science Society, Austin, TX.

Traunmller, H. (1990). Analytic expressions for the tonotopic frequency scale. *Journal of the Acoustic Society of America*, 88:97100.

Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F., and Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the National Academy of Sciences*, 104(33):13273–13278.

van Summers, W. (1987). Effects of stress and final-consonant voicing on vowel production: articulatory and acoustic analysis. *Journal of the Acoustic Society of America*, 82:847–863.

Watson, J. M. M. (1997). *Sibilant-vowel coarticulation in the perception of speech by children with phonological disorder.* PhD thesis, Queen Margaret College, Edinburgh.

Whalen, D. H. (1990). Coarticulation is largely planned. *Journal of Phonetics*, 18:3–35.

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science*, 30:945–82.

Yu, A. C. L. (2010). Perceptual compensation is correlated with individuals' "autistic" traits: Implications for models of sound change. *PLoS One*, 5(8):e11950.

Yu, A. C. L. (ToAppear). Individual differences in socio-cognitive processing and the actuation of sound change. In Yu, A. C. L., editor, *Origins of Sound Change: Approaches to phonologization*. Oxford University Press, Oxford, UK.

Zuraw, K. (2007). The role of phonetic knowledge in phonological patterning: corpus and survey evidence from Tagalog infixation. *Language*, 83(2):277–316.