Effects of Vowel Context On Consonant Place Identification: Implications for a Theory of Phonologization

Jennifer Cole **Khalil Iskarous**

j-cole5@uiuc.edu iskarous@uiuc.edu

University of Illinois at Urbana-Champaign

1 Phonetic factors in phonology[1]

A perennial question in linguistics concerns the nature of the relationship between phonology and phonetics. On casual observation the synchronic evidence is conflicting: while phonetically "natural" phonological processes abound, e.g., the palatalization of coronals preceding palatal vocoids, there are also plenty of examples of phonological processes for which phonetic motivation is not readily apparent (Anderson 1981). The fact that phonetic conditioning factors are not always evident can be taken as support for the dual claims of phonological theory that phonological systems are defined at a level distinct from phonetics, and that phonological processes need not be held to a criterion of phonetic plausibility or naturalness. In principle, any pattern that can be defined over the units of phonological representation, in accordance with formal constraints that hold over those representations, can form the basis for a phonological rule or constraint.

Despite this separation of phonetics and phonology, phonetic considerations often play a major role in phonological analysis. Given a choice between two competing analyses for the alternation X~Y, /X/ → [Y] or /Y/ → [X], preference is often given to the analysis that has a plausible phonetic basis, all else being equal. It is not hard to accept that phonetic conditioning underlies certain phonological processes. For instance, few would dispute that categorical place assimilation (viewed as a phonological process) may have its roots in patterns of gradient coarticulation (viewed as a phonetic process). What has posed more of a challenge to phonological theory has been to explain why languages vary in the extent to which patterns of phonetic variation are "phonologized", i.e., promoted to the status of a phonological process or constraint. In other words, why is it that in some languages the sequence /ti/ fails to undergo phonological palatalization to [ʧ i], despite the fact that in the realization of [ti] the [t] may be subject to palatal 'coloring' due to coarticulation?

Constraint-based approaches to phonology appear to offer a solution to this problem: phonetically motivated constraints on phonological representation may form part of a universal phonological grammar, but need not be uniformly enforced in all languages. Thus, in Optimality Theory (Prince & Smolensky 1993, Archangeli & Langendoen 1997) a particular constraint may be upheld in one language and violated in another, as a function of the interaction between it and other constraints in the phonological grammar. The possibility of violable constraints has re-opened the door to phonetic explanation in phonology, and recent years have witnessed an increasing number of phonological analyses that draw on phonetic constraints, such as those that govern speech articulation and perception. For the sake of illustration, we cite several examples here.

Constraints arising from speech articulation are invoked to explain:

- the failure of tongue root advancement harmony to affect low vowels (Archangeli & Pulleyblank 1994);

[1] This research was supported by NSF grant SBR 98-73450. The authors are grateful for the assistance of Joo-Kyeong Lee in running subjects for the experiment reported here, and for helpful advice from Gary Dell. Errors or omissions in this work are the full responsibility of the authors.

- the occurrence of vowel reduction and consonant lenition due to articulatory undershoot (Lindblom 1983);
- place assimilation due to undershoot (Lindblom 1983), reflecting patterns of coarticulation (Lee 1999);
- the voicing of stops conditioned by place of articulation, closure duration, position in a consonant cluster, and phrasal position (Ohala 1983; Westbury & Keating 1986; Hayes 1996).
- the failure of geminates to undergo lenition, viewed as a phenomenon of physical effort reduction (Kirchner 1998).

Constraints arising from speech perception are invoked to explain:

- place assimilation arising from listener misperception (Hura, Lindblom & Diehl 1992; Ohala 1990);
- the neutralization of laryngeal contrasts in consonants conditioned by the salience of acoustic cues that signal the contrasts (Steriade 1995).

2 Evaluating phonetic naturalness

While phonetic constraints may offer insight into the origin and function of a phonological pattern, their role in phonological analysis raises one fundamental question: which patterns in the phonetics may give rise to constraints that operate in phonology? Without a clear answer to this question, phonetic accounts of phonological phenomena will never make testable predictions, and thus will never offer more than convenient post-hoc analysis. This problem with functional, phonetic explanation in phonology is cited by Kentowicz (1981) as the primary reason that functionalism has not been well-received in American (generative or structuralist) linguistics, "chiefly because they [functional theories that draw on phonetic explanation (JC/KI)] require independently motivated theories of the external domain appealed to (i.e., theories of the vocal apparatus, speech perception, … etc.), which are not developed to the point where appeals to these domains lead to testable empirical consequences" (p. 431). The number of phonetic patterns that can be defined along continuous acoustic and articulatory parameters and that could in principle serve as the basis for phonological generalization is much greater than the number of phonological patterns for which a phonetic basis has been claimed. We must come to know more precisely *which* patterns in phonetics may be incorporated in phonology before evaluating claims about phonetic naturalness or the consequences of phonetic variation in phonology.

A related question concerns the form in which a phonetic constraint gets expressed when it is incorporated into phonology. One view, expressed by Ohala (1991), is that phonetic constraints are imported directly into phonological grammar, without mediation or reformulation in terms of phonological features. An alternative view is that phonetic patterns are the seeds of development for phonological constraints, but that the form and even the function of the constraint may change in the mapping from the phonetic domain to the phonological domain. This is the view put forth by Hayes (1996), who argues that "there is a considerable gap between the raw patterns of phonetics and phonological constraints,"(p. 5). Hayes claims that "the influence of phonetics in phonology is not direct, but is mediated by structural constraints that are under some pressure toward formal symmetry" (p. 11).

2.1 Articulatory factors

Hayes' argument is based on the phonetic factor of ease of articulation, in particular, on the ease of producing voicing or voicelessness on a stop consonant based on aerodynamic factors that arise from the local environment. Constraints on stop consonant voicing that are observed to function in phonological systems reflect the relative ease of producing voicing (or voicelessness) in a particular environment. Voicing is ruled out in contexts where it is harder to produce than voicelessness, and vice-versa. Under a principle of formal symmetry, the voicing constraint is extended from its phonetically-determined environment to other environments that fall into the same phonological class. For example, based on the predictions of the aerodynamic model, Hayes claims that voicing is easier to produce than voicelessness for bilabial and coronal stops following a liquid, but in the same context voicelessness is easier than voicing for the velar stop (Figure 1). Despite the difference between the velar stop and bilabial or coronal stops, Hayes' analysis predicts that languages that suspend a voicing contrast in post-liquid position will tend to affect all three places of articulation in a uniform manner, e.g., allowing [lb, ld, lg] while disallowing [lp, lt, lk].

	p	t	k	b	d	g
l/r ___ :	45	28	15	10	20	30

Figure 1. Phonetic difficulty indices for oral stops in post-liquid position. High values indicate more difficulty in the production of voicelessness for [p,t,k] or voicing for [b,d,g]. Adapted from Hayes (1996).

2.2 Perceptual factors

Crucial to Hayes' account of phonetic constraints in phonology is a concrete model of the relative difficulty of producing speech sounds in various environments. The aerodynamic model Hayes uses provides an independent measure of phonetic difficulty that can be applied to evaluate the phonetic naturalness of phonological patterns of stop voicing. While Hayes' focus is on phonetic difficulty in the domain of speech production, of equal interest is phonetic difficulty arising from speech perception, and its role as a conditioning factor in phonology. As noted above, perceptual difficulty (or conversely, perceptual salience) is claimed by Ohala (1991) and Steriade (1995) as the basis for certain phonological patterns such as assimilation.

In order to fully understand how perceptual ease or difficulty may give rise to constraints that shape phonological systems, we must first have, in Hayes' terminology, a landscape of perceptual difficulty for each type of phonological distinction that languages make. Hayes evaluates production difficulty on the basis of a formal model of speech aerodynamics. In this paper, we propose that a map of perceptual diffucuity can be obtained by considering results from speech perception research. In the remainder of this paper, we report on a speech perception experiment that investigates contextual effects on the perception of consonantal place of articulation (C-Place). The goal of the experiment is to determine if there are perceptual biases for the identification of C-Place that stem from the context of the adjacent vowels. If such effects are found, then a further goal is to determine how the perceptual biases relate to known patterns in phonology governing the occurrence of contrastive C-Place features. The results of this experimental work are offered here as a first step in the construction of a complete landscape of perceptual difficulty, towards the goal of evaluating the role of perception in phonology.

3 Phonetic and phonological evidence for contextual effects on place identification in oral stops

It is well-known that the acoustic cues that identify the place of articulation for an oral stop consonant vary according to the adjacent vowel context. The second formant transition provides an important cue to C-Place (Delattre, Liberman & Cooper 1955), but the slope of the transition varies depending on the place of articulation of the adjacent vowel (V-Place). Another important cue to C-Place is the spectrum of the release burst (Blumstein & Stevens 1979, 1980; Stevens & Blumstein 1978), which also varies as a function of the V-Place of the adjacent vowel. Thus, the possibility exists that the C-Place cues obtained in a specific vowel context are more salient than in other vowel contexts, for a given C-Place feature.

Consistent with the observed effects of adjacent vowel context on C-Place cues, phonological research provides evidence for the interaction between C-Place features for stop consonants and the V-Place features of adjacent vowels. Clements (1991) and Clements & Hume (1995) note that there are phonological systems in which labial consonants co-occur with round vowels, and systems in which palatal consonants co-occur with front vowels. These interactions include cases where the consonant conditions the place features of an adjacent vowel, as well as cases where the vowel feature conditions the place features of the adjacent consonant. Other evidence for interactions between C-place and adjacent V-Place comes from research on the lexical frequency of C-Place features in five languages (Janson 1986, Kawasaki-Fukumori 1992), as reported in Figure 2. Phonological theory has focused on formal models of these C-Place/V-Place dependencies, but makes no direct claims about the phonetic origins of these patterns of place dependency. Plausible explanations can be constructed both in terms of speech production, drawing on coarticulation, and in terms of speech perception, drawing on the salience of C-Place and V-Place cues. It is even conceivable that both kinds of effects conspire to render the observed distributional restrictions on place of articulation.

More frequent combinations:

- alveolar + front V
- labial + back, round V
- velar + back, round V

Less frequent combinations:

- velar + front V
- velar + round V

Figure 2. Evidence for C-Place and V-Place interaction from lexical frequency. Adapted from Janson (1986).

4 Experiment

To investigate how the adjacent vowel context might affect the perception of C-Place in stop consonants, we conducted a speech perception experiment in which subjects were asked to identify the C-Place feature of an intervocalic stop consonant in noise. Subjects listened to noisy stimuli consisting of nonsense words containing VCV sequences, and were asked to identify the intervocalic consonant as either [b], [d], or [g]. The goal of the experiment was to determine if and how the identification accuracy for [b], [d], and [g] may vary as a function of adjacent vowel context, and how such variation relates to observed patterns of phonological C-Place and V-Place dependencies. Furthermore, since a major cue to C-Place lies in the F_2 locus of a stop consonant, our study compares C-Place identification of stops in the context of back vowels, (with a relatively low F_2) and front vowels (with a relatively high F_2).

4.1 Stimuli

The stimuli for the experiment were constructed from a set of 75 nonsense words of the form: ómV$_1$CV$_2$. The vowels preceding (V_1) and following (V_2) the target consonant were chosen from the set of five tense vowel monophthongs, [i,e,u,o,a]. The intervocalic target consonant was chosen from the set [b,d,g]. The total number of distinct V_1CV_2 sequences was 75 (5x3x5), yielding a set of 75 nonsense words such as *ómadi, ómebu, ómigo,* etc. The nonsense words contain sounds which are readily identifiable by English listeners, but the words themselves do not sound like individual words of English, in large part because of the sequence of three syllables, each with a full vowel. Stress is located on the first syllable (expressed by a H*L pitch accent), but the second and third syllables also contain full vowels. In a typical English word, only the syllables with primary or secondary stress can contain a full vowel, and a three syllable word would typically have only one or two stressed, full-voweled syllables, e.g., *Omaha* [ˈowməˌha], *capital*

[ˈkæpərəl]. The 75 nonsense words were produced in a clear and slow speaking style by one of the experimenters, a native speaker of Egyptian Arabic.[2]

The speech signals were digitally recorded (48Hz) and checked for uniform amplitude in order to rule out the possibility that subjects would respond more accurately to an individual word due to greater loudness of the speech signal. The speech signals were transferred from tape to a Sun Ultra 5 workstation (sampled at 16kHz, 16-bit encoding) for creation of the noisy stimuli using Entropics ESPS/Waves+. A uniform noise signal was synthesized at five different amplitude levels, ranging from .5 to .75 times the average amplitude level of the clear speech signals (the nonsense words). The waveform for each nonsense word was composed with each level of noise signal, for 375 (75x5) different noisy nonsense word signals.

4.2 Method

Nine subjects, all young adult native speakers of American English, listened to the stimuli through headphones in a quiet laboratory room. The 375 noisy stimuli were presented to subjects in three randomized blocks, with a two second response-to-stimulus interval, and a short break between successive blocks. The total session duration was between 1-1.5 hours. Subjects were asked to respond immediately after hearing each noisy word, identifying the target consonant as [b], [d] or [g] by selecting one of three response buttons appearing on the computer screen. Subjects were told before beginning that all nonsense words contained one of these three consonants, and were given a brief familiarization with the form of the nonsense words.

4.3 Results

Data for vowel contexts involving the front vowels [i,e] and the back vowels [o,u] are reported here. The results for C-Place identification are presented below in terms of the frontness or backness of the adjacent vowels, rather than in terms of individual vowel qualities. The complete set of stimulus-response data for the lowest (N0) and highest (N4) noise levels are presented in the form of confusion matrices in the Appendix. We report here on the result of statistical analysis using repeated-measures ANOVA, with separate ANOVA's run for each of the three consonants. The independent variables, all of which were within-subject variables, were: Repetition (3 levels), Vowel1 (2 levels: front, back), Vowel2 (2 levels: front, back), and Noise (5 levels: noise-to-signal ratio values ranging from .5 to .75). The dependent variable was a non-parametric measure of sensitivity ($\underline{I}$), which normalizes the number of correct responses (P(hit)) for each consonant by the number of false alarms (P(fa)) for the same consonant (Hume et.al. 1999; Grier 1972). The formula for $\underline{I}$ is given below; higher $\underline{I}$ values indicate greater sensitivity, which in the present case means more accurate identification of the C-Place feature.

$$\underline{I} = (1 - P(fa) + P(hit))/2$$

Table 1 presents the ANOVA results for the main effects and interactions. Significant results are marked with an asterisk, and discussed individually below.

[2] The Egyptian Arabic speaker was chosen over an American English speaker in order to get clear, monophthongal vowel qualities. Had the stimuli included the diphthongal vowels (of the sort produced by speakers of American English) in positions adjacent to the target consonant, we would not be able to make comparisons of individual vowels in pre-consonantal vs. post-consonantal position.

Subject Df 8	Df	F [b]	Pr(F) [b]	F [d]	Pr(F) [d]	F [g]	Pr(F) [g]
Repetition	2,16	4.0295	**0.01839***	1.7476	0.17532	3.66506	**0.02634***
Vowel1	1,8	37.3732	**0.00000***	2.9592	0.08605	36.94403	**0.00000***
Vowel2	1,8	419.5836	**0.00000***	106.5299	**0.00000***	3.47781	0.06282
Noise	4,24	2.6459	**0.03295***	5.1583	**0.00048***	8.31169	**0.00001***
Repetition: Vowel1	2,16	2.3654	0.095026	1.5098	0.222024	2.01918	0.133909
Repetition: Vowel2	2,16	1.0353	0.355909	0.3171	0.728446	0.06798	0.934286
Vowel1: Vowel2	1,8	28.2962	**0.00000***	3.8650	**0.04988***	0.01908	0.890188
Repetition: Noise	8,64	1.1614	0.320894	1.4208	0.185106	1.11276	0.352948
Vowel1: Noise	4,24	1.7508	0.137674	0.5176	0.722862	1.17001	0.323222
Vowel2: NoiseL	4,24	0.2616	0.902521	1.1498	0.332404	0.75257	0.556641
Repetition: Vowel1: Vowel2	2,16	2.7247	0.066599	0.7964	0.451550	1.51826	0.220159
Repetition: Vowel1:Noise	8,64	1.0416	0.403526	1.1382	0.335903	0.83546	0.571641
Repetition: Vowel2: Noise	8,64	1.1220	0.346686	0.6296	0.753070	0.93863	0.484064
Vowel1: Vowel2: Noise	4,24	0.0567	0.994008	0.0929	0.984671	0.49019	0.742964
Residuals	472						

Table 1. Results from repeated-measures ANOVA for identification of three consonants.

4.3.1 Effect of noise

As can be seen in Table 1, the Noise effect was significant for all three consonants. To find out which pairs of noise levels were significantly different and in which direction the difference occurred, a multicomparison procedure using Tukey's method was performed.

	Estimate	Std.Error	Lower Bound	Upper Bound
n0-n1	0.0100	0.0208	-0.046800	0.0669
n0-n2	**0.0571**	**0.0208**	**0.000273**	**0.1140 ****
n0-n3	**0.0679**	**0.0208**	**0.011100**	**0.1250 ****
n0-n4	**0.0733**	**0.0208**	**0.016500**	**0.1300 ****
n1-n2	0.0471	0.0208	-0.009760	0.1040
n1-n3	**0.0579**	**0.0208**	**0.001040**	**0.1150 ****
n1-n4	**0.0633**	**0.0208**	**0.006450**	**0.1200 ****
n2-n3	0.0108	0.0208	-0.046000	0.0676
n2-n4	0.0162	0.0208	-0.040600	0.0730
n3-n4	0.0054	0.0208	-0.051400	0.0622

Table 2. Multicomparison of Noise Levels

Table 2 presents the results of the multicomparison procedure, which are shown graphically in Figure 3. The significant comparisons were for the pairs n0-n2, n0-n3, n0-n4, n1-n3, and n1-n4.

Figure 3 shows the confidence interval for the paired comparisons. Confidence intervals in the positive range indicate that identification was better at the lower noise level. In general, the closest differences (e.g., n0-n1) did not have significant effects on the sensitivity, but the larger differences (e.g., n0-n2) did, with confidence intervals well in the positive range. The largest effect of noise is seen in paired comparisons between n0 (the lowest noise level) and n3-n4 (the two highest noise levels) and between n1 and n4. These results show that the noise factor had the expected degrading effect on identification: identification at the highest noise levels was significantly worse than at the lowest noise levels.

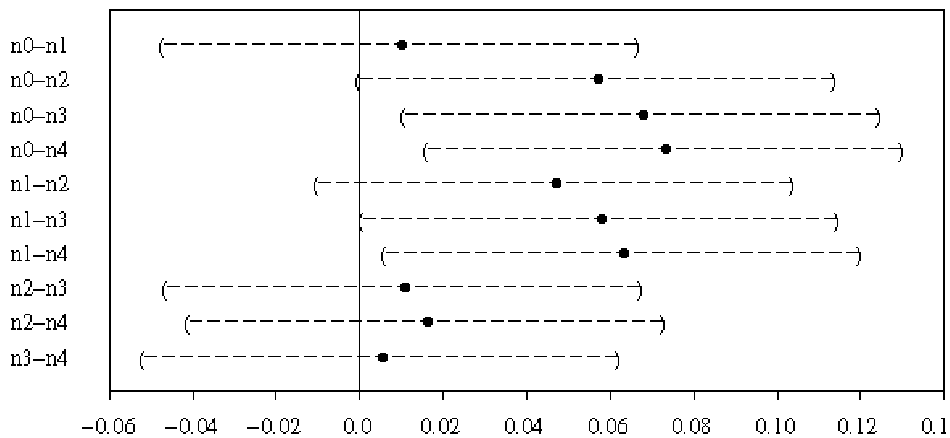

Figure 3. Multicomparison of Noise level. Simultaneous 95% confidence limits determined by Tukey's Method. Significance is detected by lack of intersection between confidence interval and 0 level.

4.3.2 Effect of adjacent vowel quality

We look next at the effect of the adjacent vowel quality on C-Place identification. The vowel contexts were defined in terms of two vowel groups: the front vowels [i,e] and the back vowels [u,o] (Table 3). Responses for the low vowel [a] were not included in this phase of the analysis in order to keep the comparison sets of equal size and diversity. The four contexts are defined in terms of the quality of the preceding and following vowels, and are identified below using the example of the target consonant [g]. Similar labels identify vowel contexts for the target consonants [b,d].

F1	Preceding vowel is front:	*igV, egV*
F2	Following vowel is front :	*Vgi, Vge*
B1	Preceding vowel is back:	*ugV, ogV*
B2	Following vowel is back:	*Vgu, Vgo*

F1_G_F2	*igi, ige, egi, ege*
F1_G_B2	*igu, igo, egu, ego*
B1_G_F2	*ugi, uge, ogi, oge*
B1_G_B2	*ugu, ugo, ogu, ogo*

Table 3. Vowel context groups for [g] stimuli.

4.3.2.1 Response patterns for [g] stimuli

The ANOVA results for [g] from Table 1 indicate a significant effect of the preceding vowel. Figure 4 plots the interaction between Vowel1 and Vowel2 for [g] identification at the highest noise level (N4). The effect of the preceding vowel is seen in the higher sensitivity values for B1 compared to F1. This is the only significant vocalic effect on the identification of [g]; no comparable difference is found for B2 and F2. The interaction between the two factors is also not significant for this consonant.

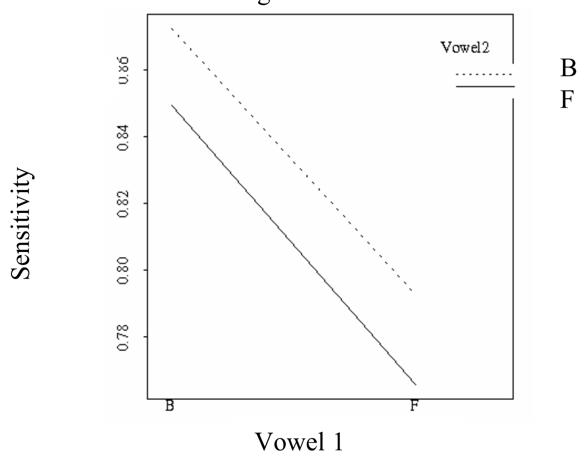

Figure 4. Interaction Plot for the factors Vowel1 and Vowel2 in the identification of [g] at noise level N4.

4.3.2.2 Response patterns for [d] stimuli

Table 1 shows that for [d] identification there is one significant main effect of vocalic quality and one interaction. These findings are shown for the noise level N4 in the interaction plot in Figure 5. When V2 is back, sensitivity is significantly greater, regardless of the quality of V1. A significant interaction between V1 and V2 is found only when V2 is front: under that condition, sensitivity is greater when V1 is back.

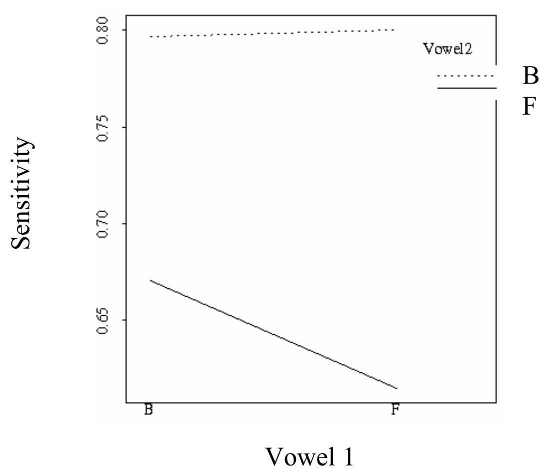

Figure 5. Interaction Plot for the factors Vowel1 and Vowel2 in the identification of [d]at noise level N4.

4.3.2.3 Response patterns for [b] stimuli

The results for [b] identification from Table 1 are similar to those for [d], with the addition of a significant main effect of V1. In all, there are two significant main effects and one significant interaction for [b], illustrated by the interaction plot in Figure 6 for the noise level N4. As with [d], sensitivity for [b] is higher when V2 is back than when V2 is front. And as with [g], sensitivity is higher overall when V1 is back, though the effect is much greater under the condition that V2 is front. This interaction between V1 and V2 is significant for [b], as seen above for [d].

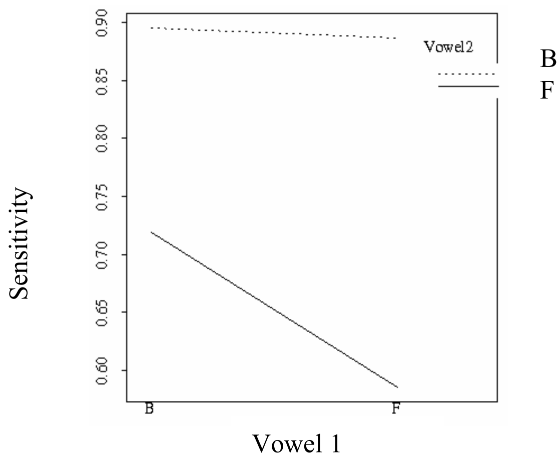

Figure 6. Interaction Plot for the factors Vowel1 and Vowel2 in the recognition of [b].

4.3.5 Summary of results

Noise is seen to have the expected effect of degrading C-Place identification. The greatest effect of noise is found in comparing results from stimuli that differed by more than one step on the noise gradient. Overall, the clearest stimuli were identified with highest accuracy and the noisiest stimuli were identified with the lowest accuracy.

To summarize the effect of adjacent vowel quality, identification of [g], [d], and [b] is facilitated when the first vowel is back, although for [d] the effect of the first vowel backness is dependent on the second vowel being front. Also, [d] and [b] identification is strongly facilitated when the second vowel is back. A generalization that emerges from this experiment, therefore, is that facilitation of identification of all three consonants occurs in the context of a back vowel. A related fact is that identification for all three consonants is worst when both preceded and followed by a front vowel.

5. Discussion

Before discussing the main results from the perception study described above, we give consideration to two possible objections to the study. First, the nonsense words used as stimuli were spoken by a native speaker of Egyptian Arabic (the second author), whose consonant and vowel productions may differ from those of American English, the native language of all of the subjects. The possibility may exist that the subjects' response patterns might be influenced by this language difference. While we do not claim that the speaker produced speech sounds just as in American English, we think that the idiosyncratic features of his pronunciation did not reduce the subjects' ability to discriminate between his [b], [d] and [g] sounds. In support of our claim, we note that none of the subjects had any difficulty in understanding the speaker's productions in the clear, not even those subjects who had no prior experience hearing the speaker's voice. Even at the lowest noise level, when the noise signal was at half the amplitude of the clear signal, C-Place identification was very good. Moreover, the speaker produced the nonsense words with clear, hyperarticulated speech, in a style known to exhibit the least effects of coarticulation of the sort that might affect C-Place perception (Lindblom 1990, Lee 1999). We expect that the effects of vowel context on C-Place identification might be even greater under more casual, hypoarticulated speech styles.

The second objection that might be raised concerns the perception bias that is introduced by the native language of the listener. There is evidence that listeners with different native languages attend differently to the acoustic cues that signal phonological distinctions between sounds. For example, Hume et al. (1999) demonstrate that listeners whose native language is Korean show a different sensitivity to C-Place cues than listeners whose native language is American English. These findings suggest the possibility that the response patterns obtained in our study reflect a bias on the part of the listeners that stems from their native language experience with American English phonetics and phonology. Although we do not dispute the

Hume et al finding, we do not think that our findings suggest a primary effect of the subjects' native language. None of the response patterns (e.g., facilitation of [d] and [b] identification by a following back vowel, facilitation of [g] identification with a preceding front vowel, or inhibition of all C-Place identification with preceding and following front vowels) reflect any property of American English phonology or phonetics known to us. Furthermore, the nonsense words did not resemble existing English words, especially in the sequencing of syllables with full, long vowels, so the VCV contexts the subjects heard should not have been very familiar to them based on their knowledge of English word structure.[3]

The results show that identification for [b] and [d] was conditioned by the following vowel, with a following front vowel conditioning lower sensitivity to C-Place identification. The fact that sensitivity to C-Place is reduced in the context of a following front vowel is consistent with evidence in phonology that front vowels (especially /i,e/) frequently condition phonological place assimilation (palatalization). If certain place features are less distinct preceding a front (i.e., palatal) vowel, then it's possible that palatalization may serve to further enhance the perceptual distinction. Neutralization between [b] and [d], for example, may be avoided if one of them is replaced by a palatal or palatalized sound in the context of a following front vowel. This account offers a perceptual motivation for palatalization, but does not rule out an articulatory motivation as well, and does not indicate whether one source for the phonological process is primary. One possibility is that a sound change, such as palatalization, is conditioned in the first place by coarticulation, but then may offer additional advantages for perception.

One important difference between our findings and patterns of phonological palatalization concerns the behavior of the bilabial stop. Phonological palatalization processes only very rarely affect bilabial consonants (Lee 1999), while in our data the affect of a following front vowel is even slightly greater for the bilabial [b] than it is for the coronal [d]. If the perceptual pattern serves (even partially) as the basis for phonological patterns of palatalization, then based on our findings we would expect to find similarity in the behavior of [b] and [d], with more or less equal occurrences of bilabial and coronal palatalization. It may be that the rarity of bilabial palatalization compared with coronal palatalization reflects articulatory differences between bilabials and coronals--- the lingual gesture for a front vowel is more likely to affect the lingual gesture of a preceding [d] than it will affect the labial gesture of a preceding [b], and thus there may be a greater articulatory basis for palatalization with coronals than for bilabials. This line of reasoning suggests that the primary motivation for palatalization may lie in the articulatory domain, with a perceptual benefit for the discrimination of bilabials and coronals that does not by itself motivate the sound change for bilabials.

It is of interest to note that the vocalic effects on [b] indentification are not at all problematic when viewed from the perspective of the interaction between labial consonants and round vowels, as mentioned in Section 3. [b] identification is best when the following vowel is back and round, which is consistent with observed phonological patterns of co-occurrence between bilabials and round vowels. In a language with the five-vowel inventory used in the present experiment, a consequence of enforcing a co-occurrence restriction between bilabials and back, round vowels is that bilabials would then fail to appear in the palatalizing context of a following front (unrounded) vowel---precisely the environment where bilabials are perceptually less salient.

Oddly, our findings do not show a significant effect of a following front vowel for [g] identification, despite the observation that phonological palatalization commonly affects velars in CV structures (Lee 1999). Our data indicate that it is the preceding vowel that influences [g] identification, with a preceding front vowel inhibiting identification. In fact, the inhibitory effect of a preceding front vowel is also found with [b] and [d] identification, but only when the following vowel is also front. This effect of a preceding front vowel is at odds with patterns of coarticulation and phonological palatalization. In her study of the acoustic evidence for coarticulation, Lee 1999 finds a greater coarticulatory influence of a vowel on an adjacent consonant in CV structures than in VC structures for bilabial, coronal, *and* velar place of

[3] One type of VCV vowel context with two full vowels that is found in English includes sequences where the second vowel is a word-final long /i/, as in *Edie, baby, Bobby, ruby, Toby.* Though listeners had greater experience with this context compared to others, we do not observe a consistent effect in either direction for contexts with a following front vowel.

articulation. She relates her finding to attested patterns of phonological palatalization, which also occur with much greater frequency in CV structures than in VC. In sum, the effects of a preceding front vowel on the perceptual patterns for C-Place identification are not consistent with the phonological patterns of palatalization, which points further to an articulatory rather than a perceptual basis for palatalization of velars.

A clear pattern that emerges for all three consonants in our study is that the context in which both the preceding and the following vowel are front is the worst overall for C-Place identification. This is a kind of "double-palatalizing" context, and suggests itself as a prime context for a phonological palatalization, or even more drastically, for the phonological suspension of a C-Place contrast altogether. Once more, we find a poor fit between the perceptual patterns and phonological patterns of place assimilation. In phonological systems, restrictions on the occurrence of C-Place features are typically one-sided. C-Place features may be restricted on the basis of a following vowel, as in palatalization, or the preceding vowel, as in the restriction of retroflex place of articulation to post-vocalic positions (Steriade 1995). But it is rarely, if ever, the case that a consonantal place feature is restricted on the basis of the combined preceding and following vowel contexts.[4] Of course, we recognize that there are phonological patterns that are restricted to VCV environments. For example, lenition processes typically affect stop consonants in intervocalic position, where the presence of both the preceding and following vowels are critical to the process, as in Spanish (Cole, Hualde & Iskarous 1999). But in that case the vowel quality is not a factor, and all vowels equally condition the lenition process. The two-sided environment with front vowels that inhibits C-Place identification in our study is more specific with respect to vowel quality than the contexts that are found to condition the distribution of C-Place in phonological systems.

To summarize, the perceptual patterns of C-Place identification from our study are partially consistent with phonological patterns of palatalization for coronal consonants: a following front vowel reduces the perceptual distinctiveness of the coronal (and bilabial) place feature. Our findings are also consistent with the phonological patterns of co-occurrence between bilabial consonants and back, round vowels. The parallel between perceptual and phonological patterns breaks down, however, when we consider the effects of a preceding front vowel on the bilabial and coronal, and even more clearly, on the velar consonant. Phonological palatalization processes are not typically conditioned by a preceding front vowel. Furthermore, we observe that the strongest influence of vocalic context occurs only when both the preceding and following vowels are front. The special status of this "two-sided" context is not observed in phonological processes that restrict C-Place features, such as palatalization. This finding presents us with indirect evidence about the kinds of structures that can serve as conditioning factors for phonological patterns. Although two-sided contexts may condition perception, to the extent that these perceptual patterns serve as a basis for phonological constraints, their incorporation into phonology may involve a process of generalization that locates the conditioning factor on either the preceding or following context, but not both. Or, perhaps, the conditioning factor may be interpreted solely in terms of the sequencing of C and V, disregarding the vowel quality (as in the example of lenition).

The findings from our study offer at best a loose fit between perceptual patterns and contexts for phonological constraints on the occurrence of C-Place. But we have considered thus far only those phonological processes which most fully respond to the perceptual effect: complete loss or assimilation of a C-Place feature. Another possibility is that languages respond to perceptual patterns not in an absolute way by *eliminating* less-salient structures, but in a gradient way, by exhibiting such structures with lower frequency than the more-salient structures. Thus, an analysis of VCV structures in the lexicon of a given language might reveal, consistent with our findings, that many fewer words exhibit the [b/d/g] place contrast in the context of flanking front vowels than in other contexts. Similarly, it's possible that the inhibiting effect of a preceding front vowel on [g] identification is reflected in a lower incidence of [igV] and [egV] sequences compared to [ugV] or [ogV]. It's also possible that the perceptual patterns found in our study are reflected in patterns of sporadic language change, with less-salient structures more frequently undergoing sound changes that increase their overall perceptual salience. We leave these matters for future exploration, noting here only that the effort to relate patterns of lexical frequency or sound change to

[4]We exclude here processes which are conditioned by position in prosodic structure, an effect of hierarchical rather than linear position in the phonological representation.

perception requires in the first place a thorough understanding of the perceptual landscape and the contexts that affect perceptual salience for a given sound feature.

In conclusion, we find that the precise nature of the mapping from perceptual constraint to phonological constraint is not yet clear, but the findings from our study support the view that the mapping is not direct. We observe some robust perceptual patterns that lack counterparts in phonological systems, and others which have phonological counterparts defined over contexts that are more general than those defined by the perceptual patterns alone. These findings suggest to us the use of caution and careful consideration in appealing to perceptual "explanations" of phonological phenomena, and the need for more research on the perceptual salience of speech sounds as a function of phonological context.

Bibliography

Archangeli, Diana & D. Terence Langendoen. 1997. *Optimality Theory.* Cambridge, Ma: Blackwell.

Archangeli & Pulleyblank. 1994. *Grounded Phonology.* Cambridge, MA: MIT Press.

Anderson, Stephen. 1981. Why phonology isn't natural. *Linguistic Inquiry* 12: 493-539.

Blumstein, Sheila & Kenneth Stevens. 1979. Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *J. Acoust. Soc. Am.* 66: 1001-17.

Blumstein, Sheila & Kenneth Stevens. 1980. Perceptual invariance and onset spectra for stop consonants in different vowel environments. *J.Acoust.Soc.Am.* 67: 648-662.

Clements, G.N. 1991. A unified set of features for consonants and vowels. Ms., Cornell University.

Clements, G.N. & Elizabeth Hume. 1995. The internal organization of speech sounds. In J. Goldsmith (ed.), *The Handbook of Phonological Theory*, pp. 245-306. Cambridge, MA: Blackwell.

Cole, J., Hualde, J. and K. Iskarous. 1999. Segmental and prosodic effects on /g/ lenition in Spanish. In O. Fujimura, B. Joseph, and B. Palek (eds.), *Proceedings of Linguistics and Phonetics 1998*. Prague: Charles University Press.

Delattre, P., Liberman, A. & F. Cooper. 1955. Acoustic loci and transitional cues for consonants. *J.Acoust.Soc.Am.* 27:769-773.

Grier, J. Brown. 1972. Nonparametric indexes for sensitivity and bias: computing formulas. *Psychological Bulletins*. 75:424-429.

Hays, William. 1994. *Statistics.* Orlando, Florida: Harcourt Brace College Publishers.

Hayes, Bruce. 1996. Phonetically driven phonology: the role of Optimality Theory and inductive grounding. Ms., UCLA. To appear in the proceedings volume of the 1996 Milwaukee Conference on Formalism and Functionalism in Linguistics.

Hume, Elizabeth, Keith Johnson, Misun Seo & Geortios Tserdanelis. 1999. A cross linguistic study of stop place perception. To appear in *Proceedings of the 14th International Congress of Phonetic Sciences 1999.*

Hura, Susan, Björn Lindblom and Randy Diehl. 1992. On the role of perception in shaping phonological assimilation rules. *Language and Speech* 35: 59-72.

Janson, T. 1986. Cross-linguistic trends in the frequency of CV sequences. *Phonology Yearbook,* 3: 179-195.

Kawasaki-Fukumori, Haruko. 1992. An acoustical basis for universal phonotactic constraints. *Language and Speech* 35: 73-86.

Kenstowicz, Michael. 1981. Functional explanations in generative phonology. In Goyvaerts, D.L. (ed.) *Phonology in the 1980's*. Ghent, Belgium: Scientific Publishers.

Kirchner, Robert. 1998. Geminate inalterability and lenition. Ms., UCLA.

Lee, Joo-Kyeong. 1999. *A Phonetic Examination of C-to-V Place Assimilation*. Doctoral dissertation, University of Illinois, Urbana.

Lindblom, Björn. 1990. Explaining phonetic variation: a sketch of the H&H Theory. In Hardcastle, W. and A. Marchal (eds.) *Speech Production and Speech Modelling*, pp. 403-439. Dordrecht: Kluwer Academic Publishers.

Lindblom, Björn. 1983. Economy of speech gestures. In MacNeilage , Peter F. (ed.), *The Production of Speech,* pp. 217-245. Heidelberg: Springer.

Ohala, John. 1983. The origin of sound patterns in vocal tract constraints. In MacNeilage, Peter F. (ed.) *The Production of Speech,* pp. 189-216. New York: Springer.

Ohala, John. 1990. The phonetics and phonology of aspects of assimilation. In Kingston, J. and M. Beckman (eds.) *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech,* pp. 258-75). Cambridge: Cambridge University Press.

Prince, Alan and Paul Smolensky. 1993. Optimality Theory. Ms. Rutgers University and the University of Colorado.

Steriade, Donca.1995. Positional neutralization. Ms., UCLA.

Stevens, Kenneth and Sheila Blumstein. 1978. Invariant cues for place of articulation in stop consonants. *J. Acoust. Soc. Am.* 64: 1358-68.

Westbury, John and Patricia Keating.1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22: 145-166.

Appendix

Front-Front Noise 0		Response		
		b	d	g
Stimulus	b	58	50	0
	d	36	66	5
	g	7	18	83

Front-Back Noise 0		Response		
		b	d	g
Stimulus	b	105	3	0
	d	23	78	7
	g	11	7	88

Back-Front Noise 0		Response		
		b	d	g
Stimulus	b	81	25	0
	d	0	106	0
	g	8	3	95

Back-Back Noise 0		Response		
		b	d	g
Stimulus	b	92	8	6
	d	11	93	3
	g	6	3	97

Front-Front Noise 4		Response		
		b	d	g
Stimulus	b	34	70	2
	d	18	72	18
	g	7	32	69

Front-Back Noise 4		Response		
		b	d	g
Stimulus	b	97	9	2
	d	19	71	17
	g	15	29	64

Back-Front Noise 4		Response		
		b	d	g
Stimulus	b	37	65	6
	d	6	92	10
	g	6	11	91

Back-Back Noise 4		Response		
		b	d	g
Stimulus	b	60	29	19
	d	9	89	9
	g	10	8	90