A CROSS-LINGUISTIC STUDY OF STOP PLACE PERCEPTION

Elizabeth Hume, Keith Johnson, Misun Seo, and Georgios Tserdanelis
Ohio State University

ABSTRACT

Perceptual salience may motivate phonological processes such as consonant place assimilation. This paper reports on an experiment designed to test the perceptual salience of consonant place of articulation by listeners who are native speakers of Korean and by listeners who are native speakers of American English. The results, when averaged across vowels and language, reveal dorsal and labial consonants to be highest on the salience ranking with coronals lowest. However, it was also found that phonetic salience of stop place depends on vowel environment. Speakers of Korean and American English also showed somewhat different patterns of confusions and overall sensitivity to place information in bursts and transitions.

1. INTRODUCTION

The role of perceptual phenomena as a source of explanation for cross-linguistic phonological patterns has, in recent years, become an increasingly active area of research in phonological theory [1, 2, 3, 5, 7, 8, 14, 17]. An underlying assumption in these works is that greater perceptual salience of a given segment is a result of more robust cues and a better cue package. Thus, Jun 1995 proposes the universal salience ranking in (1) for unreleased stop consonants.

(1) Perceptual salience:
dorsal > labial > coronal

Though, to our knowledge, the ranking in (1) has never been directly tested, the acoustic cues for the unreleased coronal are claimed to be weak as compared to those for the dorsal and labial, [7] inter alia. This is attributed primarily to the observation that in certain vowel contexts the tongue gesture of the coronal stop is rapid, resulting in short transition cues. Under the assumption that longer transitions provide more information, labials and dorsals can be considered to have more robust perceptual cues since their gestures are more sluggish, giving longer transitions. The higher degree of salience accorded the dorsal is based on the observation that dorsal consonants have an additional acoustic cue for place of articulation resulting from the convergence of F2 and F3 of a neighboring vowel [7, 15].

Under the assumption that acoustically less salient segments are more likely to be targets of place assimilation than acoustically more salient ones [7], the observed pattern of place assimilation in Korean can be drawn on as support for the ranking in (1). As illustrated in (2a), coronals assimilate in place to a following labial or dorsal consonant. Labials, on the other hand, assimilate in place to dorsal but not coronal consonants, as shown in (2b). Dorsal consonants do not assimilate to either a following coronal or labial, as in (2c).[1]

An alternative account of these facts is provided by Iverson & Lee (1994), who also assume the hierarchy: dorsal > labial > coronal. Their ranking, however, is defined in terms of phonological markedness, an abstract notion which they characterize representationally within feature geometry: the more marked the segment, the greater the amount of representational structure (for related discussion, see [9, 12, 13]).

(2)

a.	/mit+ko/	[mikk'o]	'believe and'
	/mit^h+pota/	[mipp'ota]	'more than the bottom'
b.	/ip+ko/	[ikk'o]	'wear and'
	/nop+ta/	[nopt'a] *[nott'a]	'high'
c.	/nok+ta/	[nokt'a] *[nott'a]	'melt'
	/kuk+pota/	[kukp'ota] *[kupp'ota]	'more than soup'

The identity between the proposed rankings of perceptual salience and phonological markedness raises the question as to whether markedness can be grounded solely in terms of perceptual salience. One of the goals of this paper is to seek an answer to this question. A further objective is to provide a direct test of the salience ranking in (1), given the hypothesized relation between perceptual salience and phonological processes such as Korean place assimilation.

The experiment reported in this paper was designed to investigate these issues by comparing the perceptual salience of place of articulation by listeners who are native speakers of Korean and by listeners who are native speakers of American English. This experiment allowed us to construct a ranking of perceptual salience from a direct test of salience rather than from presumptions based on acoustic properties. The stimuli were composed of release bursts in addition to aspiration noises (which contain formant transitions) which were digitally spliced from naturally produced utterances of /p^h/, /t^h/ and /k^h/. We chose to investigate the perception of place in aspirated stops because the aspirated stops in Korean and American English are practically identical acoustically. In testing release bursts and transition portions independently we were interested in possible differences in the salience of individual cues. Further, by testing both Korean and American English speakers we were able to explore possible language-specific differences between the groups of listeners. We also investigated the salience of place of articulation in three different vowel environments (/i/, /a/, /u/) to test whether the ranking of perceptual salience is dependent on the following vowel (for one or both groups of listeners).

2. METHODS

The stimuli were taken from a recording of one phonetically untrained male native speaker of Seoul Korean. Recordings were made in a sound-attenuated booth at the Ohio State University. Each syllable, which consisted of an aspirated stop (/p^h, t^h, k^h/) and following vowel (/a, i, u/), was presented to the speaker in Korean orthography in random order. Each syllable was recorded five times.

Syllables were digitized at 22050 Hz. Stimuli for the perception experiment included the burst and transition which were digitally spliced from the signal, with the transition comprised of the portion from 10 ms after the burst to the onset of vocalic voicing. The whole syllables and syllables with the burst removed were also included as fillers. The average peak amplitude was calculated for the five repetitions of each syllable and the amplitude of each stimulus was subsequently scaled in accordance with the average.

The listeners were 21 native speakers of Seoul Korean (11 males, 10 females) and 20 native speakers of American English (7 males, 13 females). The American listeners were all students at the Ohio State University and ranged in age from 18 to 37 years. The age range for the Korean listeners was 21 to 41 years, with all having lived in the U.S. between 1 and 15 years (average 4.8 years).

Listeners were instructed by the experimenter to listen carefully to the stimuli and then select the syllable that they thought they heard by clicking on a response button on a computer screen. The screen contained three different printed syllables, written in Roman alphabet, for each stimulus, i.e. *pa*, *ta*, *ka*; *pi*, *ti*, *ki*; *pu*, *tu*, *ku*. Listeners heard each syllable once with a two second response-to-stimulus interval. There were sixty trials in each block and three blocks of trials. Trials were blocked by vowel.

3. PERCEPTUAL SALIENCE

3.1. Results

We measured the perceptual salience of stop place in the burst and transition stimuli for each listener using a non-parametric measure of sensitivity, I [4]. This measure of sensitivity takes into account a listener's bias to choose a particular response alternative by calibrating 'hit rate' (the proportion of stimuli correctly identified) with 'false alarm rate' (the proportion of incorrect uses of the 'hit' response alternative). The formula for I is given in (3):

(3) $I = (1-P(fa) + P(hit))/2$,
where $P(hit) = P(\text{'r'}|R)$ and $P(fa) = P(\text{'r'}|X)$, and 'r' is a response, R is a stimulus having place 'r', and X is a stimulus not having place 'r'.

I varies from 0 to 1. A value of 1 means that labelling performance was 100% correct (no false alarms and 100% hits), a value of 0.5 means that the proportion of false alarms was equal to the proportion of hits, and a value of 0 means that there were no hits and 100% false alarms.

We analyzed these sensitivity data in a repeated-measures analysis of variance (ANOVA) having one between-listeners factor (**language**: Korean or English), and three within-listeners factors (**place**: labial, coronal or dorsal - **vowel**: [i], [a], or [u] - and **cue**: burst or transition). The results of the analysis of variance are shown in Table 1.

Table 1. Repeated measures ANOVA of sensitivity (I).

Source of Variance	DF	F	P
Between listeners			
Language	1,39	13.5	<0.01
Within listeners			
Place	2,78	22.6	<0.01
Vowel	2,78	3.8	0.03
Cue	1,39	703.2	<0.01
Place * Language	2,78	0.7	0.5
Vowel * Language	2,78	0.4	0.6
Cue * Language	1,39	7.7	<0.01
Place * Vowel	4,156	24.4	<0.01
Place * Cue	2,78	12.9	<0.01
Vowel * Cue	2,78	10.8	<0.01
Place * Vowel * Language	4,156	1.1	0.37
Place * Cue * Language	2,78	1.4	0.26
Vowel * Cue * Language	2,78	1.7	0.18
Place * Vowel * Cue	4,156	2.8	0.03
Place * Vow * Cue * Lg.	4,156	0.5	0.77

The largest effect in this analysis was the effect of **cue**. On average, the value of I for the transition stimuli was 0.88 while the value of I for the burst stimuli was 0.55. There was also a main effect for the native **language** of the listener and in the **cue * language** interaction (figure 1) we see that the difference between Korean-speaking listeners and American English-speaking listeners occurred primarily with the transition stimuli.

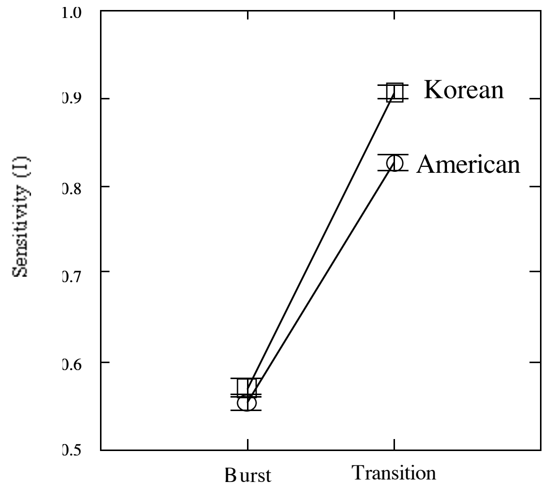

Figure 1. Cue * Language interaction.

The effect of **place** of articulation was also reliable. The average sensitivity showed dorsal to be highest in salience, followed by labial and coronal (the average values of I were: dorsal 0.73, labial 0.71, coronal 0.68). The **place * cue** interaction (figure 2) showed that the differences for place of articulation were due primarily to the transitions. While the trend in both effects is consistent with the ranking in (1), the distinction between dorsal and labial was not significant in a Tukey post-hoc comparison of means. Salience of the burst stimuli did not differ much by place of articulation.

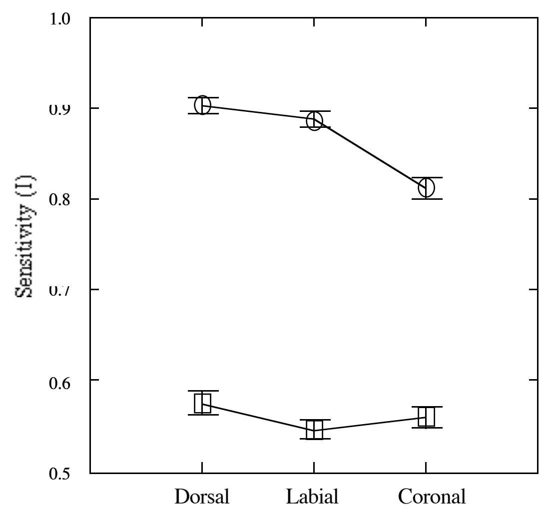

Figure 2. Place * cue interaction.

Two other interactions indicate that, for both groups of listeners, salience of stop place of articulation is mediated by the identity of the neighboring vowel. The **place * vowel** interaction is shown in figure 3, which shows that the salience of the burst and transition cues for the dorsal stop /k/ is dependent on the identity of the following vowel. In back vowel environments /k/ is more salient than /p/ or /t/, while in the front vowel environment /k/ is less salient than the other stops. The **vowel * cue** interaction was caused by the relatively higher salience of place in the burst preceding /u/ as compared with the burst in the other two vowel environments.

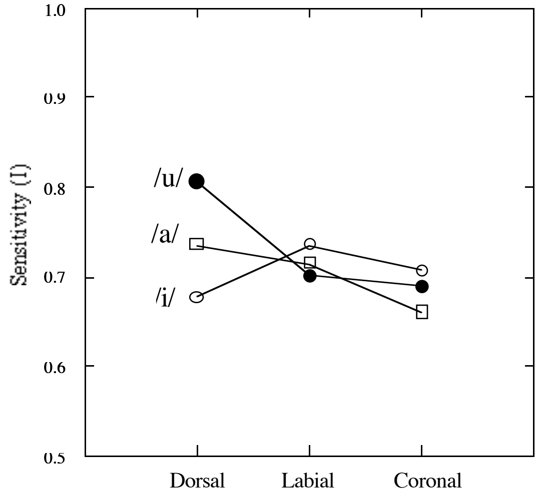

Figure 3. Place * vowel interaction.

3.2. Discussion

It is not too surprising that the transition stimuli have a greater amount of place information given that the burst stimuli were much shorter than the transition stimuli. However, it is interesting that for Korean listeners this difference between bursts and transitions was greater than it was for American English listeners. One explanation for this finding relates to differences in the system of phonological contrasts in each language. Unlike English, Korean includes the set of phonological contrasts among tense, lax, and aspirated stops, which is cued in part by the amplitude of aspiration [8]. We suggest that the presence of these phonological contrasts lead Korean listeners to focus greater attention on the interval of time following stop release burst; that is, on the transitions. This explanation is consistent with earlier studies in showing an influence of the phonological system on perception [11].

The results also reveal a main effect for place of articulation with dorsal and labial transitions being most salient and coronal transitions least salient. The trend revealed in these results with dorsal having higher and coronal lower salience lend some support to the ranking in (1) and hence, to Jun's [7] speculation about the role of formant transitions in the salience of place in syllable coda stops. However, the results also show that the relative salience of place cues, particularly those for dorsal, are highly dependent on the neighboring vowel. In fact, in the context of /i/ dorsal is less salient than both the labial and coronal. This suggests caution in appealing to auditory-perceptual data as the sole motivation for processes such as place assimilation and markedness, in general. This is underscored by the observation that patterns of place assimilation can differ cross-linguistically. For example, place assimilation in English and Sri Lankan Portuguese Creole differ from the pattern noted in (2) above for Korean: in English, only coronals typically undergo assimilation, while in Sri Lankan Portuguese Creole, of the three places of articulation, only labials and dorsals appear to be targets of assimilation [16].

4. PERCEPTUAL MAPS

To further analyze the contrastive salience of places of articulation, we calculated dissimilarity scores for pairs of the stimuli. The dissimilarity measure that we used is ϕ^2, a normalized χ^2 calculated from confusion matrices. For example, we counted the number of times that the release burst of [t^ha] was identified by Korean listeners as 't', 'p', and 'k' and compared this pattern of responses with the pattern of responses for [k^ha], using ϕ^2 to measure the dissimilarity of these two stimuli. We then entered these dissimilarity measurements into multidimensional scaling (MDS) analyses. There were a total of 12 one-dimensional MDS solutions: 2 languages (Korean & English), 2 cues (burst versus transition), and 3 vowels (/i/, /a/, and /u/). These solutions are shown in figure 4.

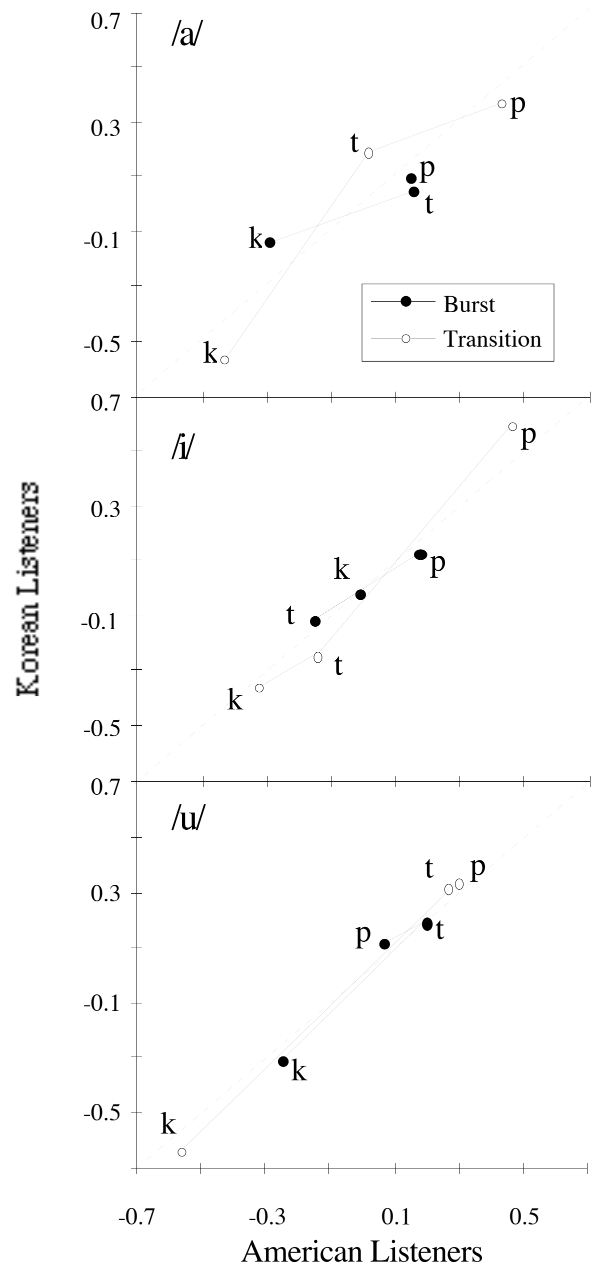

Figure 4. Results of the multidimensional scaling analyses.

In figure 4, the MDS solutions have been scaled by the original ϕ^2 values, making it possible to compare the solutions for different sets of stimuli and listeners. In these graphs, a steeply sloped line means that the Korean listeners distinguished place of articulation more accurately than did the

American listeners, and a shallowly sloped line means the opposite. Any deviation from a straight line indicates a difference between the two groups of listeners in the relative discriminability of the stimuli.

A number of patterns in the MDS solutions shown in figure 4 are remarkably similar for Korean and American English listeners. For example, the line connecting burst stimuli is consistently shorter than the line connecting transition stimuli - this is parallel to the results of the perceptual salience analysis described in the preceding section. Both analyses thus indicate that transitions provided more information concerning place of articulation than did bursts.

It is also interesting that for both groups of listeners the dorsal transition stimuli were quite distinct from the labial and coronal place stimuli in the context of /a/ and /u/, while in the context of /i/ it is the labial transition stimuli that were more distinct. This also parallels the result of the earlier analysis (vowel * place interaction), showing the influence of vowel context on the salience of different consonants. Note further that in no context does the coronal emerge as particularly distinct from the other two places, supporting its low ranking in the salience hierarchy.

A further similarity shared by the groups of listeners is the relative ordering of the burst stimuli as a function of the following vowel. When /a/ follows, both groups of listeners have the burst for /p/ and /t/ merged. When /i/ follows both groups have a larger distance between /t/ and /p/ than between /k/ and /p/. And when /u/ follows both groups have /p/ located between /t/ and /k/. It could be argued that these patterns shared by both groups of listeners reflect patterns of raw acoustic similarity which will have the same effect regardless of the native language of the listener. A similar outcome can be observed with respect to transitions.

However, figure 4 also shows some interesting differences between the groups of listeners. The most general effect, which we also found in section 3, is that Korean listeners showed greater separation of the places of articulation for the transition stimuli than did the American English listeners. The slopes of the line relating Korean and American English spaces for transition stimuli for each vowel is greater than one. It is interesting also that for /a/ and /i/ the slope of the line for burst stimuli is less than one, suggesting that American English listeners had somewhat greater separation of the place of articulation of the bursts. The particular pattern of confusions among transition stimuli were also slightly different for the Korean and American English listeners. Preceding the vowel /a/, /t/ and /p/ were not as well separated for the Korean listeners, while preceding the vowel /i/ Korean listeners had greater separation of /p/ and /t/ than did the American listeners.

5. CONCLUSION

Perceptual salience of stop place of articulation averaged across vowels and language in this experiment is similar to the phonetic salience ranking which Jun (1995) proposed. However, we also found that perceptual salience of stop place depends on vowel environment, while phonological markedness (as gauged by consonant place assimilation patterns) is not conditioned by vowel environment. This indicates that phonetic salience is only one of possibly several factors influencing phonological markedness.

We also found that speakers of Korean and American English showed somewhat different patterns of confusions (section 4) and overall sensitivity to place information in bursts and transitions (section 3). This suggests that although some universal patterns in phonological systems may be driven at least in part by perceptual phenomena, it may also be the case that listeners' perceptual abilities or strategies are shaped by their native language experience. This bidirectional interplay between phonology and speech perception is ripe for further exploration.

ACKNOWLEDGMENTS

We would like to thank John Ohala, Donca Steriade and the members of the phonetics/phonology discussion group at the Ohio State University , in particular, Mary Bradshaw, David Odden, Christian Uffmann and Pauline Welby, for valuable comments and suggestions.

NOTES

1. Phonemically lax stops are realized as tense, represented with an apostrophe in (2), following another obstruent. Place assimilation can involve fricatives and nasals as well. See Iverson & Lee (1994) for related discussion.

REFERENCES

[1] Boersma, Paul. 1998. Functional phonology. PhD dissertation. University of Amsterdam.

[2] Côte, Marie-Hélène. 1997. Phonetic salience and consonant cluster simplification. In B. Bruening, Y. Kang and M. McGinnis (eds.), MITWPL 30. Cambridge, MA, MIT.

[3] Flemming, Edward. 1995. Auditory features in phonology. PhD dissertation. UCLA.

[4] Grier, J. Brown. 1972. Nonparametric indexes for sensitivity and bias: Computing formulas. Psychological Bulletin 75, 424-9.

[5] Hume, Elizabeth. 1998. The role of perceptibility in consonant/consonant metathesis. Proceedings of WCCFL 17. Stanford, CSLI.

[6] Iverson, Gregory & Shinsook Lee. 1994. Variation as optimality in Korean cluster reduction. Proceedings of ESCOL 94. University of South Carolina.

[7] Jun, Jongho. 1995. Place assimilation as the result of conflicting perceptual and articulatory constraints. Proceedings of WCCFL 14, 221-237.

[8] Kim, Mi-Ran Cho. 1994. Acoustic characteristics of Korean stops and perception of English stop consonants. PhD dissertation. The University of Wisconsin, Madison.

[9] Lombardi, Linda. 1997. Coronal epenthesis and markedness. University of Maryland Working Papers in Linguistics 5, 156-175.

[10] Ohala, John. 1992. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. Papers from the Parasession on the Syllable: Chicago Linguistics Society, 319-338.

[11] Pitt, Mark. 1998. Phonological processes and the perception of phonotactically illegal consonant clusters. Perception and Psychophysics 60(6), 941-951.

[12] Prince, Alan and Paul Smolensky. 1993. Optimality Theory. ms. Rutgers University and The University of Colorado at Boulder.

[13] Rice, Keren. 1996. Default variability: The coronal-velar relationship. NLLT 14(3), 493-543.

[14] Steriade, Donca. 1997. Phonetics in phonology: The case of laryngeal neutralization. ms. UCLA.

[15] Stevens, Kenneth. 1989. On the quantal nature of speech. Journal of Phonetics 17, 3-45.

[16] Tserdanelis, Georgios. 1998. Nasal assimilation in Sri Lanka Portuguese Creole. ms. The Ohio State University.

[17] Wright, Richard. 1996. Consonant clusters and cue preservation in Tsou. PhD dissertation. UCLA.