Perceptual Influences on Place Assimilation: A Case Study

INTRODUCTION

Several cross-linguistic surveys of place assimilation (Cho 1990, Mohanan 1993, Jun 1995, Shin 2000) have revealed that this process targets nasals more often than plosives in the languages of the world. Moreover, Mohanan (1993), Jun (1995) and Shin (2000) have pointed out that this cross-linguistic tendency exists because there is an implicational relationship between nasals and plosives as targets of place assimilation: there are no languages in which plosives assimilate without nasals doing so as well, even though there are languages in which nasals but not plosives assimilate (e.g., Chiyao, Ponapean, Yoruba), and languages in which both groups of sounds assimilate (e.g., Malayalam, Catalan, Korean).

While Cho (1990) attributed this implicational asymmetry between plosives and nasals to an “accidental gap,” others have attempted to account for this cross-linguistic tendency by appealing to the comparative perceptual strength of cues for place of articulation in plosives and nasals. This argument—as first elaborated by Jun (1995)—maintains that nasals are more susceptible to the process of place assimilation because their place cues are perceptually weaker than the corresponding cues for plosives. Jun (1995) goes on to claim that speakers are aware of the perceptual differences between nasals and plosives, and are, therefore, more likely to permit the less-perceptible place information in nasals to undergo assimilation, since listeners would be less likely to notice this alternation. The interpretation of place assimilation as perceptually motivated in this way has been echoed in subsequent years by Steriade (2000), Boersma (1998) and Shin (2000).

There is little experimental evidence, however, showing that the perceptual weakness of nasals motivates their cross-linguistic likelihood to undergo place assimilation. On the contrary, experiments which have been specifically designed to test the perceptual underpinnings of place assimilation have shown that listeners seem to be able to perceive place information in nasals as well as or better than they perceive place information in plosives.

Ohala (1990), for instance, investigated the perceptual motivation for place assimilation by testing listener responses to place cues in an appropriate context for this process to occur. Ohala had theorized that the context-varying nature of place cues might account for the fact that place assimilation is usually regressive and not progressive. Specifically, Ohala reasoned that the relative perceptual strength of place cues leading into and out of stop consonants determined the most likely direction of place assimilation. For instance, the first stop consonant in a VC_1C_2V sequence often lacks a release burst. In this case, the stop will only have a post-vocalic transition cue to its place of articulation. The C_2 consonant, however, will always have both burst and transition cues to its place of articulation, thereby giving it a perceptual advantage over C_1. Ohala also cited evidence from Fujimura et al. (1978) , which showed that pre-vocalic stop transitions are perceptually stronger than post-vocalic stop transitions. On the basis of this evidence, Ohala hypothesized that the combined strength of place cues for C_2 might be able to overwhelm the corresponding place cues for C_1, perceptually, and thereby result in the listener perceiving the entire consonant sequence as one, homorganic stop cluster with the place of articulation of C_2. Ohala claimed that such an unintentional and "innocent" misperception on the part of the listener might be the perceptual motivation underlying the cross-linguistic phonological asymmetry between progressive and regressive assimilations.

In order to demonstrate the ability of the burst and transition cues for C_2 to override the corresponding cues for C_1, Ohala constructed a series of VCV stimuli which had conflicting place cues leading into and out of the intervocalic consonant. To give one example, Ohala spliced together a VCV sequence which had a post-vocalic /p/ transition leading into the intervocalic consonant while a pre-vocalic /k/ transition (preceded by a /k/ burst) led out. Ohala also constructed a similar set of VNCV stimuli, with post-vocalic nasal stop transition cues leading into the intervocalic consonant. In a small case study, Ohala then played these stimuli to listeners and asked them to identify the intervocalic consonant; in the case of the aforementioned example, the listeners would have had the option of replying that the sound they heard was either "apa," "aka" or "other." Ohala found that listeners consistently identified the place of the intervocalic consonant according to the place of the pre-vocalic, C_2 stop. On the basis of this evidence, then, Ohala concluded that "innocent misapprehensions" in perception could account for the general cross-linguistic tendency towards regressive place assimilations. Ohala did not, however, find any significant differences in identification rates between the VCV and VNCV groups of stimuli. Ohala's results do not, therefore, support the hypothesis that perception motivates the nasal/plosive asymmetry in place assimilation, since they showed no significant perceptual differences between these two groups of sounds.

Hura et al. (1992) also hypothesized that perception might play a role in shaping cross-linguistic patterns in place assimilation, but their understanding of perception's influence on phonology differed significantly from Ohala's. Hura et al.'s interpretation of the relationship between perception and phonology followed that of Kohler (1990), who claimed that perception influenced phonology inasmuch as "speakers are likely to sacrifice articulations that would be difficult for listeners to hear." With respect to place assimilation, this idea meant that place assimilation would only occur if listeners would not be likely to notice the loss of place information in the potential targets of assimilation. Kohler (1990) himself claimed that, "...fricatives are not assimilated under any conditions, because they are acoustically and auditorily far more distinct than nasals and unreleased stops with regard to place cues so that their articulatory reduction would be too salient, and is, therefore, not tolerated."

Hura et al. (1992) ran an experimental study in an attempt to test the truth of this claim. For this study, Hura et al. constructed a set of stimuli with VC_1C_2V sequences embedded in pairings of first and last names, with the break between names coming between the two intervocalic consonants (e.g., "Shan**ick Te**rry"). The C_1 consonants were either stops, fricatives or nasals of three different places of articulation (labial, alveolar and post-alveolar), while the C_2 consonants were all stops, of the same three places of articulation. Listeners in this study were asked to listen to the first name-last name sequence and identify the first name only.

Hura et al. found that listener error rates were lowest when the C_1 consonants were fricatives. This supported Kohler's hypothesis that fricatives did not undergo place assimilation because their place cues were too salient, perceptually. Hura et al.'s results did not, however, confirm that a corresponding perceptual difference existed between nasals and stops: "Planned comparisons revealed a significant difference for nasal and stop consonants vs. fricatives...no significant difference was found between nasals and stops." Hura et al. also noted that most of the errors their listeners made were non-assimilatory and involved, instead, a bias in the mistaken responses towards what they called a "default" segment. Hura et al. suggested that this result mitigated against the likelihood that "innocent misapprehensions" could account for patterns in place assimilation--as Ohala had hypothesized--and instead argued for the Kohlerian view, in which the place distinctions which were most difficult for listeners to perceive (as in nasals and unreleased stops) were the most likely to become the targets of place assimilation.

The fact that listeners responded more often with "default" segments instead of assimilatory mistakes, however, may have been an artifact of Hura et al.'s experimental design, rather than an indication of a more general perceptual tendency. The "default" segments for nasals and stops, for instance, were the alveolars /n/ and /t/, respectively. This bias towards alveolars is unsurprising, since /n/ and /t/ are more frequently found at the ends of English words than their labial and post-alveolar counterparts, and it has been shown that lexical frequency biases listener responses in identification experiments (Newman, Sawusch and Luce, 2000). The “default” segment for fricatives, however, was /ʃ/, even though the most common word-final fricative in English is the alveolar /s/. This bias towards /ʃ/ responses may have arisen because listeners had heard an initial /ʃ/ in the “**Sh**ani-“ name context. The bias towards /n/-responses may have been influenced by the /n/ in this context, as well. The productive English suffix, /-ɪʃ/, may have influenced listener responses towards /ʃ/, as well. The dubious origins of Hura et al.’s “default” segments therefore leaves it unclear as to whether or not the results of their study convincingly eliminate the possibility that “innocent misapprehensions” might still influence processes of place assimilation, as Ohala (1990) suggested.

In spite of these potential problems with Hura et al.’s study, however, it has given credence in ensuing years to the Kohlerian interpretation of perception’s influence on place assimilation. Despite the fact that Hura et al.’s results could only account for the perceptual differences between fricatives and (collectively) nasals and stops, it has become commonplace for phonologists to assume that perception motivates other asymmetries in place assimilation. A number of such asymmetries were unearthed, for instance, by Mohanan (1993) in a small, four-language survey. Mohanan’s research uncovered such asymmetries as: if noncoronals undergo assimilation, so do coronals; if labials trigger assimilation, so do velars; if nonsonorants undergo assimilation, so do sonorants, etc.

By the time Jun (1995) expanded upon this survey, he did so under the assumption that the cross-linguistic asymmetries he found would all have perceptual origins. Jun therefore modeled these asymmetries in an OT framework, along Kohlerian lines. In his model, Jun represented a segment's (or a class of segments') susceptibility to place assimilation with a perceptually-based "Preserve" constraint. "Preserve" constraints were satisfied by the preservation of a particular underlying segment on the surface, while articulatorily-motivated "Weakening" constraints were satisfied by the elimination of particular articulatory gestures on the surface. If a preservation constraint for a particular class of segments was ranked lower than the "Weakening" constraint, it could not, therefore, prevent the elimination of certain gestures in that class of segments—as happens in place assimilation. Furthermore, preservation constraints were ranked according to a universal hierarchy of perceptual salience; the least salient (or distinctive) segments were, therefore, the least likely to be "preserved"—or, as Kohler would have it, the most likely to be eliminated through the processes of articulatory weakening.

1) Place Assimilation Example from Jun (1995)

Korean /ip + ko/ → [ikko] 'wear and…'

	/ip + ko/	Preserve (place (onset))	WEAKENING	Preserve (place (coda))
	ipko		**!	
☺	ikko		*	*
	ippo	*!	*	

Jun applied this framework to the results of his expanded cross-linguistic survey of place assimilation processes. In this framework, then, the fact that plosives only underwent place assimilation if nasals did fell out from the fact that the preservation constraint for plosives was universally ranked higher than the corresponding constraint for nasals. This ranking, in turn, depended on the place cues for plosives being stronger than the same cues for nasals.

Without explicit evidence demonstrating the greater likelihood of nasals to undergo assimilation in an experimental setting (as in, e.g., Ohala 1990 and Hura et al. 1992), Jun's argument for the comparative perceptual weakness of nasals rested on the results of a few studies on the perception of place in post-vocalic position. For instance, Jun cited Malecot (1956) as showing that, in post-vocalic position, nasal place cues were less salient than were plosive place cues. Malecot's (1956) study had included a condition in which post-vocalic stop and nasal transitions were cross-spliced with nasal murmurs of conflicting place (e.g., /œ{d}/ transitions + /m/ murmurs, or /œ{n}/ transitions + /m/ murmurs) and then played to listeners. Malecot's results showed that, for the plosive transition + nasal murmur combinations, the place information in the plosive transition consistently determined the place of articulation given in the listener responses, regardless of the place information in the nasal murmur. However, for the nasal transition + nasal murmur combinations, the place information in the nasal murmur occasionally overrode the conflicting place information in the nasal transitions. From these results, Jun concluded that post-vocalic plosive transitions must provide stronger place cues than do post-vocalic nasal stop transitions. Jun also argued for the greater confusability of nasals by appealing to the results of Mohr & Wang (1968). Mohr & Wang (1968) had run a similarity estimation study in which they played sequences of two sounds to listeners and then asked them to estimate how similar the two sounds were, on a scale from one to five. Listeners in this study consistently ranked contrasting nasals as more similar to each other than contrasting plosives, or sounds of any other type.

Other phonologists have concurred with Jun’s idea that the perceptual weakness of nasals motivates their susceptibility to place assimilation. The argument that nasals are perceptually weaker than plosives has appeared in a handful of other studies since Jun (1995). Ohala & Ohala (1993), for instance, mention that, “Although nasal consonants as a class are highly distinct from other consonants, their place cues are less salient than those for comparable obstruents.” Boersma (1998) also claims--in describing the limitation of place assimilation to /n/ (and not just /t/) in Dutch--that, “The restriction to nasals can be explained by the fact that, e.g., the nasals /m/ and /n/ are perceptually much more alike than the plosives /p/ and /t/, so that the listener will rely less on place information for nasals than for plosives, so that the speaker has more freedom to mispronounce a nasal than a plosive.” Boersma (1998) cites the results of Pols’ (1983) consonant confusion study as justifying the claim that /m/ and /n/ are more confusable than /p/ and /t/ for Dutch speakers. Pols (1983) played post-vocalic stops and nasals to Dutch listeners in various conditions of reverberation and noise; the results of this study seemed to show that nasals were less likely to be identified correctly than stops under these conditions. Similarly, Shin (2000)—in another analysis of cross-linguistic patterns in place assimilation—claims that, “All the peculiar assimilatory patterns of nasals are due to the fact that the place features of nasals are less salient compared to oral stops and fricatives.” Finally, Steriade (2001) recounts that, “Kohler (1990) notes that nasals are more likely to assimilate than stops and stops are in turn more likely than fricatives, observations confirmed by Jun's (1995) survey...The correspondence between place assimilability and rates of place confusion was later established by Hura, Lindblom, Diehl (1992)...The resulting misperception rates display the hierarchy nasals > stops > fricatives, with nasals being the most confusable class."

Steriade's (2001) comments reveal an interesting trend to assume—on the basis of phonological facts—a perceptual explanation that does not, in fact, exist. Hura et al.'s (1992) results did not reveal a significant difference between the error rates for these nasals and plosives, despite Steriade's (2001) claims to the contrary. Ohala (1990) also failed to show significant differences between these two groups of sounds. Interestingly enough, Ohala & Ohala (1993) also cite studies such as Singh & Black (1966) and Wang & Fillmore (1961) as demonstrating the perceptual weakness of nasals. Upon closer examination, however, the discrimination of nasals seems to be almost perfect in the results of Singh & Black (1966), who themselves claimed, "the two nasal sounds /m/ and /n/ were highly intelligible among all groups of speakers and listeners." Wang & Fillmore's (1961) results also show no clear trend towards the perceptual weakness of nasals as opposed to stops; Wang & Fillmore's listeners identified nasal place correctly about as often as they did place in plosives. These results are, however, of dubious relevance to claims about potential targets of place assimilation, since Wang & Fillmore only reported results for nasals and stops in pre-vocalic position, and claimed that responses for consonants in post-vocalic position were "nearly random," anyway.

A pair of more recent studies have also failed to demonstrate a clear difference between the perceptual strength of nasals and plosives. Winters (2000), in an attempt to provide empirical underpinnings for the hierarchies of perceptual salience in Jun (1995), had listeners identify plosives and nasals of three different places of articulation in VC syllables at both a comfortable listening level and at speech reception threshold. Winters' (2000) results did not show a significant difference between the perception of these two groups of sounds, although listeners did show a slight tendency to identify nasals accurately more often than plosives. Winters (2001) discovered the same trend—though also not significant—when he expanded the previous study by playing post-vocalic stops and nasals to listeners in VCCV sequences, an appropriate context for place assimilation to occur. This last study confirmed the trend that those studies which have investigated the perception of place of articulation in an assimilatory context (Winters 2001, Ohala 1990, Hura et al. 1992) all show no significant differences between the perception of nasals and plosives. The claim that nasals' susceptibility to place assimilation therefore rests solely on the results of such studies as Malecot (1956), Mohr & Wang (1968) and Pols (1983), all of which tested the perception of place in a VC context.

These studies may have yielded conflicting results not only because of the different phonological contexts used in each, but also because of their varying methodologies and listening conditions. Pols (1983), for instance, degraded the signals that he played to listeners by masking them in various amounts of noise and reverberation. In contrast, Winters (2000, 2001) degraded the signal by playing it to listeners at their speech reception threshold. The discrepant results of these studies suggest that any perceptual differences which exist between nasals and plosives might only emerge under noisy conditions, without appearing at speech reception threshold.

The following study therefore expanded on the Winters (2001) paradigm in order to test the hypothesis that the perceptual differences between nasals and plosives might emerge in a VCCV context, as well, under similar conditions to those listeners experienced in the Pols (1983) study. Listeners in this study therefore had to identify stimuli of the form VC_1C_2V--where C_1 was either a plosive or a nasal, in the appropriate context for place assimilation—in four different listening conditions: at a comfortable listening level, in noise at both +6 dB and –6 dB signal-to-noise ratios, and at speech reception threshold. Given the results of previous experiments, it was expected that the perceptual differences between plosives and nasals would only emerge in the noisy listening conditions, while any differences between the two groups of sounds would remain indeterminate in both the comfortable listening level and speech reception threshold conditions.

METHOD

Stimulus Production, part I

VCCV stimuli were constructed from items that were recorded as read by two native speakers (one male and one female) of American English. The speakers were recorded while reading a script of sentences in a sound-attenuated booth. The speakers spoke through a Shure SM10A head-mounted microphone, connected through a Symetrix SX202 Dual mic pre-amplifier (gain ≈ 50 dB) into a Denon DRM-700 stereo cassette tape deck, where recordings were made onto Maxell Professional/Industrial C60 tape cassettes.

The script that the speakers read contained the elicitation items—written in pseudo-phonetic English—in three separate carrier contexts. The three contexts for the elicitation items "EEB" and “BEE” were, for example:

2a) Say EEB. (context one) 2b) Say BEE.
3a) Don't say EEB, say EEG. (context two) 3b) Don’t say BEE, say GEE.
4a) Don't say EED, say EEB. (context three) 4b) Don’t say DEE, say BEE.

Each production item appeared in contexts 2 and 3 twice—once with each of the other items containing stops of contrasting place of articulation, in an identical vowel context.

For example, "AHN" appeared in four different sentences of the form in contexts two and three:

5) Don't say <u>AHN</u>, say AHM.
6) Don't say <u>AHN</u>, say AHNG.
7) Don't say AHM, say <u>AHN</u>.
8) Don't say AHNG, say <u>AHN</u>.

Elicitation items included post-vocalic voiced stops of three different places of articulation (labial, coronal, dorsal) as well as post-vocalic nasals at the same three places of articulation. The elicitation script also included pre-vocalic voiced stops of the same three places of articulation. The vowels for all three elicitation types—pre-vocalic stops, post-vocalic stops, and nasals—were the three corner vowels, /i/, /u/ and /a/. The speakers read each elicitation item in each of the various contexts five times. The order of items differed in each of the five readings, so that no single item was consistently read at the end of the sentence list.

All original elicitation items were digitized at a sample rate of 22050 Hz using MacQuirer (version 4.7) on a Macintosh Power G3, while playing back on a JVC TD-W354 double cassette deck. The original items were then edited out of the carrier contexts using MacQuirer, and saved to disk as .aiff files for later presentation in a stimulus quality pre-test.

Stimulus Presentation, part I

The stimulus quality pre-test was run in order to determine whether any of the original elicitation items had been produced unclearly or incorrectly. Each edited elicitation item was therefore played to five separate listeners in order to determine whether or not it would be perceived as it had been intended to be uttered. Listeners in this pre-test heard each item while sitting in a sound-attenuated booth; the items were played by a custom SuperCard program (version 2.0) on a Power Macintosh 7100/80 over Sony MDR-7502 Dynamic Stereo Headphones. Listeners worked through the stimuli in this pre-test in a self-paced task. The stimuli were split into nine separate blocks by vowel type, syllabic position of consonant, and consonant manner. Blocking the stimuli in this way simplified the listeners' task to one of only having to determine the place of articulation for the consonant in each stimulus they heard. Participants completed this task by first listening to the SuperCard program play an individual stimulus token (e.g., [bu]) and then answering the on-screen question "What did the speaker say?" Listeners registered their responses by simply clicking on the appropriate response button on the screen. These response options were written in pseudo-phonetic English (e.g., "boo," "doo," "goo") and only differed from each other in terms of the place of articulation of their final or initial consonant. Listeners had the option of changing their responses before moving on to the next sound, but could only hear each individual stimulus once.

Though each stimulus was presented to five different listeners, each listener worked through only six of the nine blocks in the pre-test. Each block contained 90 different stimuli (3 places of articulation x 5 carrier contexts x 3 repetitions x 2 speakers), so each listener heard a total of 540 different stimuli. In all, nine listeners participated in this pre-test; all listeners were introductory linguistics students at the Ohio State University, participating for course credit.

Stimulus Production, part II

Only those stimuli which were correctly identified by all five listeners in the stimulus quality pre-test were considered for inclusion as stimuli in the subsequent perceptual study. This study tested listener perception of place in an appropriate context for place assimilation (a sequence of VC-CV syllables) in four different listening conditions: at a comfortable listening level (condition CLL), in noise at a +6 dB signal-to-noise ratio (condition PLUS), in noise at a –6 dB signal-to-noise ratio (condition MINUS), and at speech reception threshold (condition SRT). The basic set of stimuli was identical in each of these four conditions, with the particular listening condition superimposed on the basic listening task in each case.

The creation of the basic set of stimuli involved splicing together VC and CV tokens (using Macquirer v. 4.7) from the set of original items to create a sequence of VC-CV syllables. VC and CV syllables were matched according to vowel, carrier context and speaker—i.e., VC syllables from carrier context one were only matched with CV syllables from carrier context one; VC syllables containing the vowel [i] were only spliced together with CV syllables containing the vowel [i], etc. Release bursts—which had been present in the experimental pre-test—were removed from all stops and nasals in the VC utterances before splicing them together with CV syllables. This was done according to the received wisdom (Abercrombie, 1967; Ladefoged, 1975; Jun, 1995) that stops which precede other stops in consonant-consonant clusters are generally unreleased. There is some evidence, however, which suggests that this claim may be an oversimplification of the phonetic facts; for more details, see Henderson & Repp (1982) and Kim & Jongman (1996).

VC and CV syllables were spliced together so that 150 milliseconds separated the first vowel offset and the onset of the vowel in syllable two. The duration of 150 milliseconds was chosen because it was reported by Stetson (1951) to be the average duration of intervocalic CC clusters in English. In the case of vowel-nasal-stop-vowel sequences, this 150 millisecond intervocalic interval included the nasal murmur. Some original VN utterances included nasal murmurs that were longer than 150 milliseconds in duration; in these cases, the nasal murmur was cut to be just under 150 milliseconds in length.

Stimuli which had been spliced together in this way were then used as is in the comfortable listening level (CLL) portion of the perceptual study. For the PLUS and MINUS listening conditions, noise was added to the stimuli by altering them with a custom perl script (courtesy of Keith Johnson). This script (see Appendix 1) operates by first calculating the RMS amplitude of an .sd file and then adding to it a pre-specified level of noise. In order to add noise to the basic set of VC-CV stimuli, the test stimuli first had to be converted from .aiff to .sd format. This was accomplished by first converting all basic stimuli from .aiff to .wav format using SoundApp (ver. 1.5.1., by Norman Franke) on a Power Macintosh G3. The resultant .wav files were then transferred to a UNIX system, where the custom perl script converted them from .wav to .sd format. The script then took the converted .sd files, added half a second of silence both before and after the file, calculated the RMS amplitude of the sound file, and added white noise to the file in a pre-specified signal-to-noise ratio (SNR). Noise was added at both a SNR of +6 dB and a SNR of –6 dB to create two separate sets of sound files for the PLUS and MINUS listening conditions. The resultant files were then converted back to .wav format (using the custom perl script), and then transferred back to a Power Macintosh G3, where SoundApp was once again used to convert them to .aiff format.

Stimulus Presentation, part II

Edited VC_1C_2V stimuli were presented to listeners in nine separate blocks, split according to vowel and C_1 manner (nasal or plosive). Blocking the stimuli in this way ensured that the listeners' only task was to determine the place of articulation of the two consonants in the VC_1C_2V sequence. Each block contained 54 (3 C_1 x 3 C_2 x 3 carrier contexts x 2 speakers) stimuli, so each listener heard a total of 324 stimuli in six separate blocks.

Listeners heard the stimuli as played by a custom Supercard program (version 2.0) on a Power Macintosh 7100 in a sound-attenuated booth, over Sony MDR-7502 Dynamic Stereo Headphones. Listeners worked through the six blocks of stimuli in a self-paced task; after they heard each stimulus, the Supercard program would present them with the question “What did the speaker say?” on a screen that included nine different response options, written in pseudo-phonetic English. The response options only differed from each other in the place of articulation of the two intervocalic consonants. The response options were also written as sequences of two words. Presenting the choices in this way helped clarify the distinction between coronal and dorsal nasals before /g/ (e.g., “een gee” as opposed to “eeng gee”). Listeners registered their responses by simply clicking on the appropriate on-screen option. Listeners had the option of changing their responses before moving on to the next stimulus, but could hear each individual sound only once.

All stimuli were played to listeners through an Atus AM100 stereo mixer, which enabled relatively fine-tuned adjustments of the loudness level of the stimuli through the headphones. Before testing began, the loudness level of stimuli being played through the mixer was calibrated for each of the four listening conditions using a Quest 155 sound level meter (ANSI standard type 1). For the comfortable listening level (CLL) and the two noise conditions (PLUS and MINUS), the output level was adjusted so that the sound level of the stimuli through the headphones was approximately 70 dB SPL. Previous calibration—using the same instruments—had determined that the appropriate loudness level for "speech reception threshold" (as determined by the method described in Winters

2000, 2001) was approximately 40 dB SPL. The output level of the mixer was therefore adjusted before each listening condition so that the listeners heard the stimuli at the appropriate loudness level. In the Speech Reception Threshold (SRT) condition, the extra step of removing the Macintosh CPU from the sound booth was taken so that the noise from its fan would not interfere with the listening task.

There were twenty listeners in each of the four separate listening conditions, for a total of 80 listeners altogether. All listeners were introductory linguistics students at the Ohio State University, participating for course credit.

RESULTS AND ANALYSIS

Table 1 shows the raw response totals for all listeners, broken down by the manner and place of the C_1 consonant, for all four listening conditions. The percentages of correct responses reveal some of the general trends in the data; there are, for instance, more correct responses in the CLL and PLUS listening conditions than there are in the other two listening conditions. Analyzing response data only in terms of correct identification percentages may be problematic, however, in that these percentages may be influenced by listener response bias. Such bias seems to exist, for example, in the PLUS listening condition, where listeners strongly favored the coronal response option over its labial and dorsal alternatives. Thus, the high number of correct “coronal” responses in this listening condition does not reflect listener sensitivity to these sounds so much as it reflects their biased tendency to simply give coronal responses in this condition.

In an attempt to filter out the effects of listener response bias, raw response totals were converted into values of A’, a non-parametric measure of listener sensitivity to particular contrasts (Grier, 1971). A’ accounts for listener bias in part by incorporating the percentage of “false alarms” for a particular response category into a measure of listener sensitivity. Since the task in this study included three response options, false alarms for a particular response category were counted any time that response option (e.g., labial) was chosen in response to stimuli of either alternative category (e.g., coronal or dorsal). Table 1 lists the percentage of hits and false alarms for each response category, along with the values of A’ for each place and manner of articulation, in all four listening conditions. The values of A’ were calculated with the following formula, as given in Grier (1971):

$$2)\ A' = 1/2 + \frac{(H - FA) * (1 + H - FA)}{4 * H * (1 - FA)}$$

where H and FA represent the percentages of hits and false alarms, respectively.

A repeated measures analysis of variance (ANOVA) was run in order to determine whether any trends in the resultant A’ values were statistically significant. This ANOVA included Place (labial, coronal, dorsal) and Manner (plosive, nasal) as within-subjects factors, and ListCond (CLL, PLUS, MINUS, SRT) as between-subjects factors. The ANOVA did not include several other potential factors—vowel context differences, speaker differences, carrier context differences, and the place of articulation of the C_2 consonant—in an attempt to limit the complexity of the ANOVA and also to focus the analysis on the crucial issues: whether nasals’ lack of salience (in comparison to plosives) accounts for their susceptibility to place assimilation, and whether or not this difference in salience is affected by listening condition. The variability inherent in the other, unanalyzed factors was included in the experiment to shield the results of interest from potential confounds and artifacts arising from any single experimental condition--e.g., the possibility that either nasals or plosives might be more salient in an [i] vowel context, or when produced by the female speaker, etc. Table 2 lists those factors which were found to be significant in the ANOVA. Post-hoc, two-tailed t-tests were also run on individual A’ populations in order to determine the direction of significance for the individual factors and their interactions in the analysis of variance. Table 2 provides a summary of the results of these t-tests, which are spelled out in more detail in Table 3.

These statistical tests confirm neither of the predictions made prior to the study. The results of the ANOVA and the post-hoc t-tests indicate that listeners are more sensitive to the distinctions between nasals in the appropriate context for place assimilation than they are to the distinctions between plosives. The Manner factor (in Table 2) is significant and the post-hoc t-tests reveal that the mean A' value for nasals across all conditions is reliably greater than the same value for plosives. Such results indicate that perceptual weakness does not motivate the susceptibility of nasals to place assimilation in the manner proposed by Kohler (1990) and Jun (1995). If the Kohlerian interpretation of place assimilation applied to these results, speakers would be more likely to allow plosives to undergo place assimilation than nasals.

The statistical analysis of the response data also reveals that the perceptual strength of nasals varies in comparison to plosives across listening conditions, although not in the manner originally expected. Nasals were perceptually stronger than plosives in the CLL, PLUS and MINUS conditions, but not in the SRT listening condition. Here, listeners were more sensitive to the distinctions between plosives than they were to the distinctions between nasals. This result unexpectedly conflicts with the earlier findings of Pols (1983) and Winters (2000, 2001). These previous studies had indicated that the perceptual salience of nasals diminishes in noisy listening conditions (as in Pols 1983), while remaining on a par with plosives in clear listening conditions or at speech reception threshold (as in Winters 2000, 2001). The opposite pattern emerged here: listeners were more sensitive to nasal distinctions in the CLL and noisy listening conditions, but more sensitive to plosives at speech reception threshold.

Considering the acoustic nature of the cues for place of articulation in plosives and nasals may provide one means of resolving this conflicting pattern of results. Malecot (1956), for instance, showed that post-vocalic plosive transitions provide stronger cues to place of articulation than do post-vocalic nasal transitions. The articulatory mechanics of producing post-vocalic nasals may account for this difference; the opening of the velo-pharyngeal port in anticipation of a subsequent nasal stop closure may obscure the transition cues at the end of a vowel by nasalizing that portion of the vowel (Mary Beckman, p.c.). However, post-vocalic nasals in a VNCV context also have place cues in the nasal murmur (as was shown by Malecot 1956) and in the transition between the vowel and the nasal murmur (Kurowski & Blumstein 1984). The results of this study suggest that the combination of post-vocalic nasal transition and nasal murmur cues is perceptually stronger than post-vocalic plosive transitions alone. The added salience provided by nasal murmur cues could account for the perceptual strength of nasals in the CLL condition. The combination of transition and murmur cues also appears to be resistant to the masking effects of noise in the PLUS and MINUS conditions. However, listeners may not have been able to pick up on the nasal murmur cues—which have a relatively low amplitude to begin with—in the SRT condition. The inaccessibility of these cues for listeners would result in post-vocalic transitions providing the only cues to place for nasals and plosives at speech reception threshold. In this case, the perceptual dominance of plosive transitions over nasal transitions apparently re-emerged.

This account is not inconsistent with the findings of Winters (2000, 2001), but still does not account for the comparatively poor perception of nasals in Pols' (1983) study. One possible explanation for this discrepancy is that the stimuli in Pols (1983) may have also included release bursts on the post-vocalic plosives; this extra cue to place may have made plosives more distinctive, perceptually, than nasals. Malecot (1958), Winitz et al. (1972), and Blumstein & Stevens (1980) have all shown that release bursts provide very salient place information for plosives. Pols' study was also carried out in Dutch, with Dutch listeners, and may therefore reflect fundamental differences between the ways Dutch and English listeners perceive place of articulation.

ANOTHER ANALYTICAL VIEW

While analyzing response data in terms of listener sensitivity (i.e., A') addresses the issue of whether or not perceptual salience may influence place assimilation as Kohler (1990) envisioned, it still leaves open the question of whether or not "innocent misapprehensions" yield particular patterns in place assimilation, as Ohala (1981, 1990) suggested. In Ohala's theory, perception's influence on phonology arises solely from the listener, and the types of perceptual mistakes that they are prone to make. In order to address the question of whether or not the perceptual mistakes listeners made in this study can account for nasals' susceptibility to place assimilation, then, it is necessary to look for any tendency in the listeners to mistakenly perceive heterorganic consonant-consonant sequences as homorganic sequences—i.e., the sort of "innocent" perceptual mistake that could lead to a phonological process of place assimilation.

Tables 4 and 5 summarize the listener responses to stimuli with heterorganic clusters, grouped together by C_1 place of articulation and listening condition. These tables list the responses to these stimuli according to their relationship, in terms of place of articulation, with the original stimulus. For example, if the original stimulus was a labial-coronal cluster, "X" would denote the labial place of articulation while "Y" would denote the coronal place of articulation (and a "Z" in this case would designate dorsal). If a listener mistakenly identified this stimulus as a labial-dorsal sequence, then, the response would count as an XZ mistake, since the labial place in the response is identical to the original X place of articulation in the stimulus, while the Z (dorsal) place in the response matches neither the X (labial) nor the Y (coronal) place in the original stimulus cluster. The mistake type that is of particular interest in these Tables is the "YY" response to heterorganic "XY" stimuli—i.e., those "innocent" listener misapprehensions which result in perceptually-derived regressive place assimilations. Table 6 lists the frequency of these mistakes for each C1 manner and listening condition, while grouping all other response types (including correct responses) together under the heading of "non-assimilations." Given this small set of contingency tables, it is possible to determine, using a χ^2 test, whether or not listeners are more likely to "innocently misperceive" place assimilations when the original stimuli are nasal-plosive sequences than when they are plosive-plosive sequences. The χ^2 tests reveal that this is, in fact, the case in the MINUS and SRT listening conditions ($p = .001$ and $p = .004$, respectively). On the basis of this evidence, then, one may draw the conclusion that, under such extreme listening conditions, the comparative likelihood of nasal-plosive clusters to be misperceived as regressively assimilated, homorganic clusters may account for their cross-linguistic susceptibility to place assimilation.

Whether or not this story is particularly realistic remains unclear, however. The proportion of regressively-assimilated, YY misperceptions is still much smaller than the number of correct responses, and is generally about the same as the proportion of ZY responses. The number of ZY responses, in fact, presents an interesting conundrum in its own right. Table 6 also lists contingency tables broken down into ZY responses and non-ZY responses for each manner type and listening condition; χ^2 tests on these tables reveal that listeners are more likely to misperceive plosive-plosive sequences as having the corresponding ZY places of articulation in the CLL and PLUS conditions ($p < .001$ and $p = .003$, respectively) than they are to misperceive the nasal-plosive sequences in the same way. On the basis of Ohala's theory, then, one ought to expect that the "innocent misapprehensions" of listeners should be just as likely to spawn XY → ZY processes as they are to produce XY → YY assimilations, since phonological processes such as these are entirely listener-based in origin. Moreover XY → ZY processes should more commonly involve plosive-plosive sequences than they do nasal-plosive sequences. This, of course, is not the case; XY → ZY processes are not as common as regular assimilatory processes of the form XY → YY, if they occur at all.

XY → ZY processes do not exist in phonology for an apparent reason: there is no connection between the target of the process and the phonological environment in which the process occurs. The "Z" component—while apparently a viable response option in perception—does not exist in either the target or the trigger of the XY → ZY process and therefore appears in the output from out of the blue, as it were. While this process would be predicted never to occur by the mechanics of a phonological paradigm like feature geometry, it cannot be ruled out by Ohala's theory that sound changes originate from listeners alone. If listeners alone cannot account for the cross-linguistic propensity of nasals to undergo place assimilation, then what can? It follows from the failure of Ohala's theory that more than just listeners must play a role in shaping patterns of place assimilation; constraints on speakers must play a part, as well. Viewed from an articulatory perspective, it becomes clear that processes of XY → YY assimilation occur because they simplify the articulation of heterorganic clusters. XY → ZY processes do not occur, however, because they lack any similar articulatory motivation. They neither simplify the articulation of the cluster, nor preserve all of the place information in the input string.

The need to simplify articulatory tasks seems to provide a reasonable explanation for the susceptibility of nasals to place assimilation, as well. The production of a nasal stop requires two active, articulatory gestures: the opening of the velo-pharyngeal port and a stop closure with an articulator somewhere in the oral tract. The production of a voiced plosive, on the other hand, requires only one of these articulations—the stop closure at some point in the oral tract. In a heterorganic cluster, then, the transition from consonant to consonant is more complex, articulatorily, if the first consonant is a nasal instead of a plosive. In this case, the speaker must not only close the velo-pharyngeal port, but also change the place of articulation of the stop closure in the oral tract--as well as coordinate these two gestures so that they take place at the same time. Any effort on the speaker's part to simplify the articulatory task, therefore, would reasonably target heterorganic nasal-plosive clusters before plosive-plosive clusters, since the nasal-plosive clusters are more difficult to produce. Place assimilation provides one way of simplifying this articulatory task; it therefore seems sensible, from an articulatory standpoint alone, to expect that it would target nasals more often plosives in the languages of the world.

CONCLUSION

The results of this study do not support the hypothesis that nasals' cross-linguistic tendency to undergo place assimilation more often than plosives has perceptual origins. Listeners can distinguish place in nasals better than in plosives, in an appropriate context for place assimilation, in both clear and noisy listening conditions. Though the perceptual advantage of nasals over plosives disappears at speech reception threshold, the extra murmur cue for place in nasals seems to provide them with added perceptual salience over burst-less plosives in the most realistic listening conditions. An analysis of listeners' perceptual mistakes also indicates that "innocent misapprehensions" alone cannot account for the tendency of nasals to become targets in place assimilation.

Instead, articulatorily motivated constraints on consonant cluster combinations seem to provide a stronger argument for the susceptibility of nasals to place assimilation.

Along with this re-evaluation of the hypothesis that perception motivates this particular trend in place assimilation, this study has a broader and perhaps more important message for the study of the interaction of perception and phonology: it is important to get the perceptual facts right before assuming that they can account for any given phonological pattern. While testing hypotheses (such as the one above) about the presumed “difficulty” of certain articulations remains a challenging endeavor, recent advances in speech technology have made testing the perception of certain phonological patterns a relatively straightforward process. It is no longer necessary to base arguments about perceptual salience on the basis of cue analysis or tests on the perception of segments in irrelevant environments when it is possible to tackle the perceptual issue head on, and investigate how listeners perceive the sounds in question, in whichever context they are undergoing a phonological change. Without empirical verification of this sort, any argument attempting to link a phonological pattern with some aspect of perception remains incomplete.

Special thanks go to Keith Johnson, Terah Schamberg, Paul Davis and Andrea Sims for their help with this project. This work is based on material supported by a National Science Foundation Graduate Fellowship.

Appendix 1: Script for adding noise to .wav files

```
#!/usr/local/bin/perl

#set USE_ESPS_COMMON=off

$SNR = 6;

print "SNR = $SNR\n";

while(<*.wav>){
    $signal = $_;

    @args = ("wav2sd", "$signal");
    system(@args) == 0 or die "system @args failed: $?";

#make needed file names
    ($first, $ext) = split(/\./,$signal,2);
    $sd = sprintf("%s.sd",$first); #sd file
    $padded = sprintf("%s.pad",$first);  #silence added
    $noise = sprintf("%s.noise",$first); #scaled noise
    $splusn = sprintf("%s_noise.sd",$first); #output

#get basic data about sound file
    $ndrec = int ‘hditem -i ndrec $sd‘;
    $record_freq = int ‘hditem -i record_freq $sd‘;

#how loud is the signal?
    ($i1, $i2, $i3, $i4, $i5, $i6, $rms, $i8) =split(’ ’, ‘stats -n $sd‘);
    print $sd," ", $rms, " ",$ndrec/$record_freq,"\n";

#make the noise one second longer than the signal
    $noise_ndrec = $ndrec + $record_freq;

#add silence at the beginning of the signal
    $silence = $record_freq/2;  #0.5 sec of silence
    @args = ("zero_pad","-i", "-l $silence", "$sd", "$padded");
    system(@args) == 0 or die "system @args failed: $?";

#how loud should the noise be?
    $noise_rms = $rms/(10**($SNR/20));
    print "noise level (in RMS): ",$noise_rms,"\n";   #  <- new line

#make the noise sample
```

```
    @args = ("testsd","-Tgauss","-l$noise_rms","-r$record_freq","-p$noise_ndrec", "-tSHORT", "$noise");
    system(@args) == 0 or die "system @args failed: $?";

#add the files
    @args = ("addsd","-z","$noise","$padded","$splusn");
    system(@args) == 0 or die "system @args failed: $?";

#convert back to .wav
    @args = ("sd2wav", "$splusn");
    system(@args) == 0 or die "system @args failed: $?";

    system("rm $sd $padded $noise $splusn");

}
```

REFERENCES

Abercrombie, D. 1967. *Elements of general phonetics.* New York: Aldine.

Blumstein, S.E. and Stevens, K.N. 1980. Perceptual invariance and onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America* **67**, 648-662.

Boersma, P. 1998. *Functional phonology: formalizing the interactions between articulatory and perceptual drives*. PhD Dissertation, University of Amsterdam.

Cho, Y.-M. 1990. *Parameters of consonantal assimilation*. PhD Dissertation, Stanford University.

Fujimura, O., Macchi, M.J., and Streeter, L.A. 1978. Perception of stop consonants with conflicting transitional cues: a cross-linguistic study. *Language and Speech* **21**, 337-346.

Grier, J. B. 1971. Nonparametric indexes for sensitivity and bias: computing formulas. *Psychological Bulletin* **75**, 424-429.

Henderson, J. B., & Repp, B. H. 1982. Is a stop consonant released by another stop consonant? *Phonetica* **39**, 71-82.

Hura, S.L., Lindblom, B., and Diehl, R.L. 1992. On the role of perception in shaping phonological assimilation rules. *Language and Speech* **35**, 59-72.

Jun, J. 1995. *Perceptual and articulatory factors in place assimilation: an optimality theoretic approach*. PhD Dissertation, UCLA.

Kim, H., & Jongman, A. 1996. Acoustic and perceptual evidence for complete neutralization of manner of articulation in Korean. *Journal of Phonetics* **24**, 295-312.

Kohler, K. 1990. Segmental reduction in connected speech in German: phonological facts and phonetic explanations. In *Speech production and speech modelling* (W. J. Hardcastle and A. Marchal (eds.)), 69-92. Netherlands: Kluwer.

Kurowski, K.M. and Blumstein, S.E.. 1984. Perceptual integration of the murmur and formant transitions for place of articulation in nasal consonants. *JASA* **76**, 383-390.

Ladefoged, P. 1975. *A course in phonetics* (1st edition). Fort Worth: Harcourt Brace Jovanovich.

Malecot, A. 1956. Acoustic cues for nasal consonants:an experimental study involving a tape-splicing technique. *Language* **32**, 274-284.

Malecot, A. 1958. The role of releases in the identification of released final stops. *Language* **34**, 370-380.

Mohanan, K. P. 1993. Fields of attraction in phonology. In *The last phonological rule* (John Goldsmith (ed.)), 61-116. Chicago: The University of Chicago Press.

Mohr, B. and W.S.-Y. Wang. 1968. Perceptual distance and the specification of phonological features. *Phonetica* **18**, 31-45.

Newman, R., Sawusch, J., and P. Luce. 2000. Underspecification and Phoneme Frequency in Speech Perception. In *Papers in laboratory phonology V: acquisition and the lexicon* (M. Broe and J. Pierrehumbert, eds.), 299-312. Cambridge: Cambridge University Press.

Ohala, 1981. The listener as a source of sound change. In Masek, C.S., R.A. Hendrik, M.F. Miller (eds.), *Papers from the parasession on language and behavior: Chicago Linguistics Society*. Chicago: CLS. 178-203.

Ohala, J.J. 1990. The phonetics and phonology of aspects of assimilation. In *Papers in laboratory phonology I: between the grammar and the physics of speech* (J. Kingston and M. Beckman, eds.), 258-275. Cambridge: Cambridge University Press.

Pols, L.C.W. 1983. Scaling of consonant confusion matrices. *Speech Communication* **2(4)**, 275-293.

Shin, E. 2000. English place assimilation: optimality theoretic analysis from a functional approach. ms., Yonsei University.

Singh, S and J.W. Black. 1966. Study of twenty-six intervocalic consonants as spoken and recognized by four language groups. *JASA* **39**, 372-387.

Steriade, D. 2001. Directional asymmetries in assimilation: a perceptual account. In *The role of speech perception in phonology* (E. Hume and K. Johnson (eds.)), 219-250. New York: Academic Press.

Wang, W.S.-Y. and C.J. Fillmore. 1961. Intrinsic cues and consonant perception. *Journal of Speech and Hearing Research* **4**, 130-136.

Winitz, H., Scheib, M., and Reeds, J. 1972. Identification of stops and vowels for the burst portion of /p,t,k/ isolated from conversational speech. *Journal of the Acoustical Society of America* **51**, 1309-1317.

Winters, S. 2000. Turning phonology inside out: testing the relative salience of audio and visual cues for place of articulation. *OSU WPL* **53**, 168-199. Columbus, OH: Ohio State.

Winters, S. 2001. VCCV perception: putting place in its place. *OSU WPL* **55**, 70-87. Columbus, OH: Ohio State.

Table 1: Raw Response Totals

	CLL		Respond =					
	VC	Labial	Coronal	Dorsal	Total	P(hits)	P(FA)	A-prime
	Labial	754	160	166	1080	69.81%	20.14%	0.833
Given {	Coronal	157	786	137	1080	72.78%	18.75%	0.852
	Dorsal	278	245	557	1080	51.57%	14.03%	0.791
	Total	1189	1191	860	3240			

	VN	Labial	Coronal	Dorsal	Total	P(hits)	P(FA)	A-prime
	Labial	938	68	74	1080	86.85%	8.66%	0.939
Given {	Coronal	75	837	168	1080	77.50%	13.10%	0.893
	Dorsal	112	215	753	1080	69.72%	11.20%	0.875
	Total	1125	1120	995	3240			

	PLUS		Respond =					
	VC	Labial	Coronal	Dorsal	Total	P(hits)	P(FA)	A-prime
	Labial	571	242	267	1080	52.87%	15.46%	0.788
Given {	Coronal	153	802	125	1080	74.26%	27.45%	0.819
	Dorsal	181	351	548	1080	50.74%	18.15%	0.760
	Total	905	1395	940	3240			

	VN	Labial	Coronal	Dorsal	Total	P(hits)	P(FA)	A-prime
	Labial	738	219	121	1078	68.46%	14.91%	0.853
Given {	Coronal	146	761	173	1080	70.46%	22.38%	0.825
	Dorsal	176	264	640	1080	59.26%	13.62%	0.825
	Total	1060	1244	934	3238			

	MINUS		Respond =					
	VC	Labial	Coronal	Dorsal	Total	P(hits)	P(FA)	A-prime
	Labial	440	304	336	1080	40.74%	28.47%	0.618
Given {	Coronal	322	412	346	1080	38.15%	30.65%	0.576
	Dorsal	293	358	429	1080	39.72%	31.57%	0.581
	Total	1055	1074	1111	3240			

	VN	Labial	Coronal	Dorsal	Total	P(hits)	P(FA)	A-prime
	Labial	486	313	281	1080	45.00%	27.55%	0.657
Given {	Coronal	291	446	343	1080	41.30%	28.61%	0.621
	Dorsal	304	305	471	1080	43.61%	28.89%	0.636
	Total	1081	1064	1095	3240			

	SRT		Respond =					
	VC	Labial	Coronal	Dorsal	Total	P(hits)	P(FA)	A-prime
	Labial	556	244	280	1080	51.48%	27.64%	0.698
Given {	Coronal	260	566	254	1080	52.41%	26.57%	0.711
	Dorsal	337	330	413	1080	38.24%	24.72%	0.633
	Total	1153	1140	947	3240			

	VN	Labial	Coronal	Dorsal	Total	P(hits)	P(FA)	A-prime
	Labial	523	287	270	1080	48.43%	29.12%	0.668
Given {	Coronal	269	476	335	1080	44.07%	26.30%	0.661
	Dorsal	360	281	439	1080	40.65%	28.01%	0.622
	Total	1152	1044	1044	3240			

Table 2: Significant effects in three-factor ANOVA

Factors	F	df	Sig.	Post-hoc T-tests:			
Manner	28.25	1, 76	<.001	Nasal >> Oral			
Listcond	48.877	3, 76	<.001	CLL, Plus >> SRT > Minus			
Place	44.347	2, 75	<.001	Labial > Coronal >> Dorsal			
Manner*Listcond	16.31	3, 76	<.001	CLL	Nasal >> Oral	Oral	CLL, Plus >> SRT >> Minus
				Plus	Nasal >> Oral	Nasal	CLL, Plus >> SRT, Minus
				Minus	Nasal > Oral		
				SRT	Oral >> Nasal		
Manner*Place	9.191	2, 75	<.001	Oral:	Coronal, Labial >> Dorsal	Labial:	Nasal >> Oral
				Nasal:	Labial >> Coronal, Dorsal	Coronal:	Nasal, Oral
						Dorsal:	Nasal >> Oral
Place*Listcond	5.221	6, 152	<.001	CLL:	Labial > Coronal >> Dorsal	Labial	CLL > Plus >> SRT, Minus
				Plus:	Labial, Coronal >> Dorsal	Coronal	CLL, Plus >> SRT >> Minus
				Minus:	Labial > Dorsal, Coronal	Dorsal	CLL, Plus >> SRT, Minus

Table 3: Results of post-hoc, two-tailed t-tests

Manner	Mean A’	T-Test
Oral	0.716	**<.001**
Nasal	0.748	-

		T-Tests	
Place	Mean A’	vs. Coronal	vs. Dorsal
Labial	0.752	**0.049**	**<.001**
Coronal	0.738	-	**<.001**
Dorsal	0.706	-	-

		T-Tests		
ListCond	Mean A’	vs. Plus	vs. Minus	vs. SRT
CLL	0.850	0.208	**<.001**	**<.001**
PLUS	0.813	-	**<.001**	**<.001**
MINUS	0.610	-	-	**0.039**
SRT	0.656	-	-	-

ListCond*Manner		Mean A’	T-Tests
CLL	Oral	0.811	**<.001**
	Nasal	0.888	-
Plus	Oral	0.791	**<.001**
	Nasal	0.835	-
Minus	Oral	0.587	**0.013**
	Nasal	0.632	-
SRT	Oral	0.675	**0.006**
	Nasal	0.638	-

			T-Tests		
Manner*ListCond		Mean A’	vs. Plus	vs. Minus	vs. SRT
Oral	CLL	0.811	0.501	**<.001**	**<.001**
	Plus	0.791	-	**<.001**	**<.001**
	Minus	0.587	-	-	**0.001**
	SRT	0.675	-	-	-
Nasal	CLL	0.888	0.078	**<.001**	**<.001**
	Plus	0.835	-	**<.001**	**<.001**
	Minus	0.632	-	-	0.832
	SRT	0.638	-	-	-

Place*Manner		Mean A’	T-Tests
Labial	Oral	0.730	**<.001**
	Nasal	0.774	-
Coronal	Oral	0.735	0.355
	Nasal	0.742	-
Dorsal	Oral	0.684	**<.001**
	Nasal	0.728	-

Table 3 (continued)

			T-Tests	
Manner*Place		Mean A’	vs. Coronal	vs. Dorsal
Oral	Labial	0.730	0.553	**<.001**
	Coronal	0.735	-	**<.001**
	Dorsal	0.684	-	-
Nasal	Labial	0.774	**0.005**	**<.001**
	Coronal	0.742	-	0.145
	Dorsal	0.728	-	-

			T-Tests	
ListCond*Place		Mean A’	vs. Coronal	vs. Dorsal
CLL	Labial	0.877	**0.028**	**<.001**
	Coronal	0.857	-	**<.001**
	Dorsal	0.816	-	-
Plus	Labial	0.824	0.819	**<.001**
	Coronal	0.823	-	**<.001**
	Dorsal	0.792	-	-
Minus	Labial	0.634	**0.031**	**0.043**
	Coronal	0.594	-	0.478
	Dorsal	0.602	-	-
SRT	Labial	0.675	0.680	**<.001**
	Coronal	0.679	-	**<.001**
	Dorsal	0.614	-	-

Place*ListCond		Mean A’	vs. Plus	vs. Minus	vs. SRT
Labial	CLL	0.877	**0.041**	**<.001**	**<.001**
	Plus	0.824	-	**<.001**	**<.001**
	Minus	0.634	-	-	0.087
	SRT	0.675	-	-	-
Coronal	CLL	0.857	0.291	**<.001**	**<.001**
	Plus	0.823	-	**<.001**	**<.001**
	Minus	0.594	-	-	**0.001**
	SRT	0.679	-	-	-
Dorsal	CLL	0.816	0.453	**<.001**	**<.001**
	Plus	0.792	-	**<.001**	**<.001**
	Minus	0.602	-	-	0.615
	SRT	0.614	-	-	-

Table 4: Raw mistake totals for plosive-initial, heterorganic clusters by place and listening condition

CLL	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	Total
LC	268	2	26	3	5	3	0	52	1	360
LD	278	3	31	2	4	1	5	34	2	360
CL	253	4	36	5	4	4	2	50	2	360
CD	259	6	29	3	0	4	3	55	1	360
DL	203	1	67	6	3	5	6	66	3	360
DC	184	1	65	3	4	6	0	94	3	360
Total	1445	17	254	22	20	23	16	351	12	2160

PLUS	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	Total
LC	194	0	81	2	3	3	0	76	1	360
LD	181	0	86	2	14	2	0	74	1	360
CL	247	12	45	6	11	3	1	34	1	360
CD	252	9	42	5	3	3	2	44	0	360
DL	177	3	55	3	5	6	9	99	3	360
DC	180	0	120	3	3	1	1	52	0	360
Total	1231	24	429	21	39	18	13	379	6	2160

MINUS	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	Total
LC	145	8	70	5	12	17	7	83	13	360
LD	45	19	28	21	72	70	23	46	36	360
CL	78	37	44	46	27	22	35	60	11	360
CD	47	66	36	68	31	17	41	34	20	360
DL	79	16	45	18	40	26	27	69	40	360
DC	117	16	85	15	14	9	19	79	6	360
Total	511	162	308	173	196	161	152	371	126	2160

SRT	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	Total
LC	169	10	48	15	20	12	4	70	12	360
LD	131	20	58	17	28	26	21	38	21	360
CL	130	21	58	21	31	10	18	44	27	360
CD	154	30	47	19	19	17	18	34	22	360
DL	89	23	72	21	25	22	18	70	20	360
DC	118	12	78	20	12	4	16	93	7	360
Total	791	116	361	113	135	91	95	349	109	2160

Table 5: Raw mistake totals for nasal-initial, heterorganic clusters by place and listening condition

CLL	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	Total
LC	310	1	20	1	1	2	3	22	0	360
LD	317	0	16	2	2	2	2	12	7	360
CL	270	4	24	3	4	1	1	53	0	360
CD	271	3	61	7	1	2	1	13	1	360
DL	256	3	37	2	4	3	3	48	4	360
DC	248	2	68	0	2	4	2	33	1	360
Total	1672	13	226	15	14	14	12	181	13	2160

PLUS	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	Total
LC	257	0	72	0	0	0	1	30	0	360
LD	233	0	53	1	5	5	0	62	0	359
CL	250	1	49	5	3	1	1	50	0	360
CD	217	10	86	2	0	1	0	44	0	360
DL	214	5	60	3	4	2	3	63	6	360
DC	208	0	94	0	0	0	0	58	0	360
Total	1379	16	414	11	12	9	5	307	6	2159

MINUS	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	Total
LC	127	11	81	7	21	9	10	80	14	360
LD	64	17	45	17	69	36	20	55	37	360
CL	96	32	64	27	26	15	23	60	17	360
CD	70	43	55	58	19	22	50	32	11	360
DL	82	23	62	17	48	31	22	35	40	360
DC	123	18	83	18	7	11	20	76	4	360
Total	562	144	390	144	190	124	145	338	123	2160

SRT	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	Total
LC	142	7	80	14	16	12	8	70	11	360
LD	121	10	61	15	34	24	20	55	20	360
CL	126	18	54	26	16	19	12	66	23	360
CD	107	24	81	35	14	17	22	40	20	360
DL	99	32	83	12	20	20	13	70	11	360
DC	122	17	76	7	4	5	12	109	8	360
Total	717	108	435	109	104	97	87	410	93	2160

Table 6: Chi-squared tests of independence for assimilatory and non-assimilatory percepts

CLL	Non-YY	YY	Total	Chi-squared p
Plosives	1906	254	2160	0.175
Nasals	1934	226	2160	
Total	3840	480	4320	
PLUS	Non-YY	YY	Total	
Plosives	1731	429	2160	0.570
Nasals	1745	414	2159	
Total	3476	843	4319	
MINUS	Non-YY	YY	Total	
Plosives	1852	308	2160	0.001
Nasals	1770	390	2160	
Total	3622	698	4320	
SRT	Non-YY	YY	Total	
Plosives	1799	361	2160	0.004
Nasals	1725	435	2160	
Total	3524	796	4320	

CLL	Non-ZY	ZY	Total	Chi-squared p
Plosives	1809	351	2160	<.001
Nasals	1979	181	2160	
Total	3788	532	4320	
PLUS	Non-ZY	ZY	Total	
Plosives	1781	379	2160	0.003
Nasals	1852	307	2159	
Total	3633	686	4319	
MINUS	Non-ZY	ZY	Total	
Plosives	1789	371	2160	0.175
Nasals	1822	338	2160	
Total	3611	709	4320	
SRT	Non-ZY	ZY	Total	
Plosives	1811	349	2160	0.015
Nasals	1750	410	2160	
Total	3561	759	4320	