EMPIRICAL INVESTIGATIONS INTO THE PERCEPTUAL AND ARTICULATORY ORIGINS OF CROSS-LINGUISTIC ASYMMETRIES IN PLACE ASSIMILATION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for

the Degree Doctor of Philosophy in the Graduate

School of The Ohio State University

By

Stephen James Winters, M.A.

* * * * *

The Ohio State University
2003

Dissertation Committee:

Professor Keith Johnson, Adviser

Professor Mary Beckman

Professor Elizabeth Hume-O'Haire

Approved by

____

Adviser
Linguistics Graduate Program

ABSTRACT

Some researchers have claimed that nasals are cross-linguistically more likely than stops to undergo place assimilaton because they have weaker perceptual cues to their place of articulation.

This dissertation investigates this hypothesis by testing perceptual differences between speakers of English and Dutch, two languages which have different assimilatory patterns with respect to nasals and stops. The first perception experiment involves a magnitude estimation task, which requires Dutch and English listeners to make subjective estimates of the differences between nasals and stops of various places of articulation in VC syllables. The results of this study show that release burst cues significantly increase the estimated magnitudes of differences in stops over nasals, but that stops without these cues do not have a perceptual advantage over nasals. A second experiment, testing the perception of place in VCCV sequences in an AX discrimination task, yielded similar results to the first.

A subsequent production experiment tested the possible relevance of stop release burst cues to perceptual influences on the stop/nasal asymmetry in place assimilation. The results of this study showed that stops did have release bursts more often in this context than nasals, for speakers of both languages. Re-interpreting the data from the previous perception experiments indicates the proportion of bursts in the two languages is not great enough to give stops a consistent perceptual advantage over nasals.

A final experiment tests the hypothesis that articulatory constraints might motivate nasals to undergo place assimilation more often than stops. To test this hypothesis, Dutch and English speakers attempted to reproduce the VCCV stimuli from the AX discrimination experiment. The results of this study show that speakers of both languages exhibited less accuracy in reproducing nasals than stops, suggesting that articulatory difficulties might motivate nasals' cross-linguistic susceptibility to place assimilation.

This dissertation concludes by investigating how the frequency of particular places of articulation in English and Dutch could determine the targets of place assimilation in those languages. The results of this analysis suggest that cross-linguistic patterns in place assimilation are best understood as the product of various phonetic factors on the structure of phonology.

ACKNOWLEDGMENTS

Thanks go first and foremost to Keith Johnson, who not only provided direct intellectual inspiration and guidance for the work in this dissertation, but who also helped me survive through seven long years of grad school.

Thanks also go to Beth Hume for her guidance, advice and vision. Beth was also the person who introduced me to Jongho Jun's work and encouraged me to apply my thoughts on place perception to Jun's findings. The rest, as they say, is history.

Mary Beckman also helped by providing copious amounts of advice and an always challenging perspective on my basic assumptions about things linguistic.

Edgar Huitema, Annemarie Bodaar, Grant McGuire and Hope Dawson helped by recording the original set of stimuli for this experiment. Despite the fact that I couldn't use their recordings--due to my own stupidity--I'd like to thank them all for helping out, anyway.

I did use the recordings I made of Jos Bosch, Josja Eggen, Jeff Mielke and Julie McGory and would like to thank them for agreeing to let numerous strangers listen to their voices, again and again and again.

Allison Blodgett, Robin Dautricourt, Matt Makashay, Georgios Tserdanelis, Peggy Wong, Kiyoko Yoneyama, Tsan Huang and Laurie Maynell all took part in the stimulus quality pre-test, and deserve an extra special thanks for suffering through that uninspiring task. I hope you all enjoyed the doughnuts.

Tjeerd Dijkstra helped me find Dutch-speaking subjects in Columbus by putting me in contact with the entire Columbus Dutch Club.

I would like to thank John Nerbonne for coordinating my stay at the Rijksuniversiteit Groningen. Among other things, John put me in touch with Mariette Dijkstra, who probably did more than any other person to make sure this entire project was a success. Mariette not only shouldered the onerous task of scheduling two weeks' worth of subjects at Groningen--during exams--but also spent countless hours in that little sound booth room, making sure nothing blew up while subjects were running through the experiments. For all of her time and help, I would like to thank Mariette.

At Groningen, I also had the pleasure of getting to know Laurie Stowe, who helped right my experimental ship when things started falling apart.

And, speaking of things falling apart, I can't write an acknowledgments section without thanking that mousey, white-haired, middle-aged guy in the A/V department at Groningen who dug up and donated to me his home-made voltage transformer. I wouldn't have been able to do anything at Groningen without it. I'd like to thank him for his help, and apologize to him for never getting his name.

Back home, Jim Harmon and Jeff McCune provided more computer and technical help than I can even remember. Thanks, guys.

Pauline Welby also pitched in by helping me apply the Praat segmentation script she had written to my own data.

And, last but not least, my Mom, Suzanne Winters, made my trip to Holland feasible by loaning me the money I needed to pay subjects while I was at Groningen.

This material is based partly upon work supported under a National Science Foundation Graduate Fellowship.

On a personal level, I would also like to thank:

Robin Dautricourt, for the soccer games on Saturday mornings;

Jeff Mielke, for asking me to come back after my first practice with the band;

Paul Davis, for introducing me to Perl;

Patrick Arnold, for knowing more about baseball than I do;

Georgios Tserdanelis, for the nights of basketball at the Greek Church;

Jason Packer, for explaining to me the difference between *kelvar* and *olvar;*

Jenny Vannest, for always inviting me to Larry's;

Grant McGuire, for combining Axis & Allies with Eastwood one Saturday afternoon;

and my family, who, after all these years, are pretty much the only people who ever call anymore.

VITA

May 10, 1974....................................Born – Alexandria, Minnesota

1992...B.A., Linguistics, Pomona College

1998...M.A., Linguistics, Ohio State University

1996-present.......................................Graduate Teaching and Research Associate,
The Ohio State University

PUBLICATIONS

1. Hume, E. and Winters, S. 2002. Distinctive Features. In The Encyclopedia of Cognitive Science. London: MacMillan

2. Munroe, R.L., Munroe, R.H., and Winters, S. 1996. Cross-Cultural Correlates of the Consonant-Vowel (CV) Syllable. Cross-Cultural Research **30 (1)**, 60-83.

FIELDS OF STUDY

Major Field: Linguistics

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

PROLOGUE: THE IMPORTANCE OF PERCEPTION IN PHONOLOGY

There is a longstanding tradition in linguistic theory of emphasizing the importance of studying the auditory and perceptual aspects of speech sounds in phonology. Those phonological theories which have focused on perception, however, have often developed in a characteristic way. They generally begin by stressing the importance of certain phonological phenomena which articulatory factors alone cannot account for. Many of these phenomena can be readily explained by perceptual factors, which thenceforward form the basis of the theory. Over time, however, such approaches uncover other phonological facts which <u>perception</u> alone cannot account for. The theory is then generally revised so as to recognized the importance of <u>both</u> perceptual and articulatory factors in phonology, with equal theoretical status usually being given to both aspects of speech. However, the theories generally concede at this point that they will primarily use articulatory terms to describe phonological phenomena for practical reasons, despite the fact that perceptual or auditory descriptions would be equally appropriate. Characterizing phonolgical descriptions in this way thus opens the door for subsequent criticisms that the theory has an articulatory bias, and that it should more adequately reflect the influence perception and audition can have on the structure of speech sounds.

For example, Ferdinand de Saussure (1983) claimed in his Course in General Linguistics that,

> "Many phonologists limit themselves almost exclusively to the phonational act, i.e., the production of sound by the vocal organs (larynx, mouth, etc.) and neglect the auditory side. Their method is wrong. Not only does the auditory impression come to us just as directly as the image of the moving vocal organs, but it is also the basis of any theory." (63)

These "auditory impressions" were important within structuralist linguistic theory because they made it possible to perceive the significant differences between sounds.

> "We cannot tell articulatorily where one sound ends and another begins...It is the sequence the ear hears that enables us immediately to detect when one sound is replaced by another...The sequence of sounds we hear is not divided into segments of equal duration, but into segments identifiable as auditory units. This fact provides us with a natural starting point for the study of speech sounds." (64)

The auditory ability to perceive differences between sounds was, in turn, important for Saussure because language, itself, consisted of nothing but differences:

> "Just as the conceptual part of linguistic value is determined solely by relations and differences with other signs in the language, so the same is true of its material part. The sound of a word is not in itself important, but the phonetic contrasts which allow us to distinguish that word from any other. That is what carries the meaning." (163)

Despite having laid out this theoretical framework, Saussure (1983) based his own entirely on articulatory features. Saussure justified this approach on both practical...

> "The identification of sounds in a spoken sequence thus rests solely on auditory impressions. But their description is a different matter. This must be based upon articulatory considerations, since the units of an auditory sequence as such are unanalysable. We need to appeal to the corresponding sequence of movements in phonation." (65)

and theoretical grounds:

> "It then emerges that the same sound corresponds to the same movement: *b* (the auditory segment) = *b′* (the articulatory segment)...The speech

> sound is an aggregate of auditory impressions and articulatory movements, comprising what is heard and what is spoken, one delimiting the other. It is thus a complex unit, with a foot in each camp." (65)

Saussure used this belief in a direct correspondence between sounds and gestures to validate his decision to analyze speech in articulatory terms, even though it apparently contradicted his earlier comments about the fundamental importance of perception in phonology.

Such theoretical contradictions do not, however, seem to have had a detrimental effect on Saussure's subsequent influence on phonological theory. Some of those who followed him did, however, learn from his mistakes. N.S. Trubetzkoy (1969), for instance, considered neither audition nor articulation to be more important than the other in the study of phonology; he just considered both to be irrelevant:

> "The speech sounds that must be studied in phonetics possess a large number of acoustic and articulatory properties. All of these are important for the phonetician since it is possible to answer correctly the question of how a specific sound is produced only if all of these properties are taken into consideration. Yet most of these properties are quite unimportant for the phonologist. The latter needs to consider only *that aspect of sound which fulfills a specific function in the system of language*."

Like Saussure, however, Trubetzkoy cast all of his phonological descriptions in articulatory terms. Given that the phonological descriptions in Trubetzkoy (1969) are all articulatorily-based, however, it therefore appears that Trubetzkoy also believed that articulatory considerations provided all the facts necessary to descibe "those aspects of sound" which functioned in (and were therefore relevant to) phonology.

Joos (1948) (echoed by Hume & Johnson (2001)) correctly pointed out that the dependence of linguists like Trubetzkoy and Saussure on articulatory descriptions of phonological phenomena was an inevitable consequence of the fact that they had no technical means of analyzing the acoustic structure of speech in a meaningful way. Whatever their theoretical posture, they had to depend entirely on articulatory descriptions in order to make phonetic sense of phonological phenomena. Joos (1948) suggested that this traditional theoretical dependence on articulatory descriptions could make later phonologists resistant to future attempts to characterize speech sounds entirely in acoustic or auditory terms, as Saussure had once envisioned.

With the invention of the spectrograph in the years following World War II, however, a description of speech sounds along these lines suddenly became possible. Jakobson, Fant and Halle (1951) combined the progress made in the acoustic analysis of speech with Jakobson's theory of "distinctive features" to develop a paradigm in which all the linguistically relevant aspects of speech sounds had clear, acoustic definitions. This theory fell in line not only with the (unrealized) logic of Saussure (1983) but also with Jakobson's own conception of the communicative function of speech: "We speak in order to be heard in order to be understood." Moreover, Jakobson theorized that, in order to be able to understand anything a speaker might say, a listener had to glean information from the acoustic signal, which was coming to them as the result of the articulatory gestures that the speaker had made. This signal had to include all of the information the speaker hoped to convey; Jakobson conceived of this information as being encoded in a system of binary-valued acoustic features, to which the listener had direct access:

> "Any minimal distinction carried by the message confronts the listener with a two-choice situation. Within a given language each of these oppositions has a specific property which differentiates it from all the others. The listener is obliged to choose either between two polar qualities of the same category, such as grave vs. acute, compact vs. diffuse, or between the presence and absence of a certain quality, such as voiced vs. unvoiced, nasalized vs. non-nasalized, sharpened vs. non-

sharpened (plain). The choice between the two opposites may be termed distinctive feature." (p. 3)

Jakobson, Fant and Halle's system of distinctive features, then, placed acoustic features in the phonological foreground and relegated articulation to a secondary status in phonological descriptions. Articulations in this system only had relevance to phonology, in terms of the acoustic information they transmitted to a listener: "We cannot classify, nor even give a precise description, of the various articulations, unless we constantly hold in mind the question: what is the acoustic function of such and such a motor performance?" Jakobson, Fant and Halle's conception of distinctive features as representing only those acoustic properties of sound which functioned in a linguistic system thereby realized, to a certain extent, the earlier theoretical dreams of Saussure and Trubetzkoy.

However, the acoustically-based conception of distinctive features did not remain in the theoretical foreground for long after Jakobson, Fant and Halle (1951). As Joos (1948) had predicted, phonologists eventually returned to using articulatory terms to describe and explain phonological phenomena. One possible reason for this return to the articulatory status quo may have been that some early research on speech perception seemed to indicate that this process depended less on information in the acoustic signal than had originally been thought. Researchers at Haskins Labs at this time were searching for the "invariant" acoustic features of speech--that is, the acoustic cues which had a one-to-one correspondence with the speech sounds that listeners perceived. The approach that such studies as Cooper (1950), Cooper et al. (1951) and Cooper et al. (1952) used to find such invariant cues was to play simplified spectrograms to listeners through a Pattern Playback speech synthesizer and determine which acoustic cues (i.e., spectrographic frequencies, patterns and durations) always induced a particular phonemic percept. Through this approach, these researchers quickly found that invariant cues were hardest to isolate for the place of articulation of stops and nasals. Or, as Cooper et al. (1952) claimed, "It is precisely this kind of relationship [invariance] that we do not find, at least for these stripped-down stops and nasal resonants." Haskins Labs' inability to synthesize invariant acoustic cues for stop place of articulation led to a significant shift in their strategy for the synthesis of stop place cues, as well as a fundamental reappraisal of their intuitions about the nature of speech perception. Further Haskins Lab studies simply attempted to establish "sufficient" cues for stop place of articulation, which they eventually in the second formant transitions between consonant and vowel (as well as, to a lesser extent, the third formant transitions).

Liberman et al. (1954) noted the difficulty Cooper et al. (1952) had in synthesizing reliable, context-independent release burst cues for place and--following suggestions made in Joos (1948) and Potter, Kopp and Green (1947)--attempted to establish the "locus" of the second formant consonant-to-vowel transitions as the primary cue for stop place of articulation. Liberman et al. (1954) determined that the same second formant transitions sufficiently cued place of articulation for both voiced and voiceless stops, as well as nasals. Liberman et al. (1954) also determined that they could cue the manner distinctions between these classes of consonants with place-invariant manner cues (e.g., silence for voiceless stops and low-frequency resonances for nasal stops). Delattre et al. (1955) further established specific acoustic values for formant loci in synthesized speech. Delattre et al. (1955) found a second formant locus for /d/ at 1800 cps and a second-formant locus for /b/ at 720 cps. /g/ proved to be a more problematic case, however--"A locus for /g/ can be demonstrated only when the adjoining vowel has its second formant above about 1200 cps; below that level no /g/ locus was found."

This difficulty in establishing a second formant locus for /g/ helped lead to some interesting (and influential) conclusions by Haskins Lab researchers on the nature of speech perception. Cooper et al. (1952) had already noted, for instance, that,

> "...the perceived similarities and differences between speech sounds may correspond more closely to the similarities and differences in the *articulatory* domain than to those in the acoustic domain; that is to say, the relation between perception and articulation may be simpler than the relation between perception and the acoustic stimulus."

Liberman (1957), however, took this idea one step further:

> "We have been able to find no acoustic invariant to correspond to--or, if you will, to explain--the unchanging perception of /g/ in this series, and there are reasonable grounds for supposing that none exists. The important thing, of course, is that this discontinuity at the acoustic level is not paralleled by any corresponding discontinuity in articulation or in perception. The /g/ articulation is essentially the same throughout, and so also is the perception...All of this strongly suggests, as do other similar cases, that speech is perceived by reference to articulation--that is, that the articulatory movements and their sensory effects mediate between the acoustic stimulus and the event we call perception."

Within a span of about ten years after the publication of Jakobson, Fant and Halle (1951), then, researchers had concluded that the acoustic features of speech--rather than being the sole medium of linguistic contrast--had no relevance to the perception of speech whatsoever.

While Haskins Labs' conclusions did not translate directly into the realm of phonological analysis, the next incarnation of Distinctive Feature theory--in Chomsky and Halle's (1968) <u>The Sound Pattern of English</u>--defined all Distinctive Features only in terms of their articulatory properties.

> "In the succeeding pages we shall list the individual features that together represent the phonetic capabilities of man...We shall describe the articulatory correlate of every feature and illustrate the feature by citing examples of its occurrence in different languages of the world. We shall speak of the acoustical and perceptual correlates of a feature only occasionally, not because we regard these aspects as either less interesting or less important, but rather because such discussions would make this section, which is itself a digression from the main theme of our book, much too long." (p. 299)

Even though Chomsky & Halle were not troubled by the use of articulatory terms to define distinctive features, they did concede that their "abstract but not arbitrary" characterization of these features might be problematic. In the now (in)famous chapter nine of The Sound Pattern of English, Chomsky & Halle lament that "The entire discussion of phonology in this book suffers from a fundamental theoretical inadequacy...The problem is that our approach to features, to rules, and to evaluation has been overly formal...In particular, we have not made use of the fact that the features have intrinsic content." (p. 400) Specifically, Chomsky & Halle realized that their theory was unable to distinguish unlikely, or uncommon, phonological rules and phonemic systems from likely, or common ones. For instance, they point out that their theory would predict rule (1.2) to be more common than rule (1.1):

(1.1) $[+\text{nasal}] \rightarrow \begin{bmatrix} \alpha ant \\ \beta cor \end{bmatrix} / \text{---} \begin{bmatrix} \alpha ant \\ \beta cor \\ C \end{bmatrix}$

(1.2) $[+\text{nasal}] \rightarrow \begin{bmatrix} + ant \\ \alpha cor \end{bmatrix} / \text{---} [\alpha\text{cor}]$

However, rule (1.1) is found in many more languages than rule (1.2). Likewise, Chomsky & Halle's system makes no distinction in complexity between the following two vowel systems:

(1) i u
e o
a

(2) h ∨

É a

...even though the first is much more likely than the second.

Liljencrants & Lindblom (1972) claimed that one of the reasons Chomsky & Halle's system had these problems was that it did not recognize that speech was used for a specific purpose--namely, communication. Liljencrants & Lindblom hypothesized that the need to serve this function of communication would force phonemic systems to evolve in directions that would make linguistic/phonetic communication easier and more efficient. Liljencrants & Lindblom hypothesized that vowel systems such as (1.3) would evolve, then, so as to maximize the perceptual distance between all the vowels in the system. "The appeal of the principle of maximal contrast is no doubt based on the belief that vowels can serve as more efficient carriers of differences in meaning as they become more dissimilar, and the risk of confusing them decreases." (855) Liljencrants & Lindblom believed that the most likely vowel systems would have the greatest aggregate perceptual distance between all the vowels in the system, while systems which failed to maximize these perceptual distances would be less likely.

Liljencrants & Lindblom were able to test this hypothesis because the acoustic cues for vowels (unlike those for stop place of articulation) had been identified through spectrographic analysis: they are primarily encoded in the first and second formant values for each vowel. They therefore interpreted the perceptual distance between vowels as a Euclidean distance in a two-dimensional formant space, with one dimension for each of the two different formant values. For example, the formula for the perceptual distance between any two vowels, i and j, was:

(1.5) $r_{ij}^2 = (F1_i - F1_j)^2 + (F2_i - F2_j)^2$

Liljencrants & Lindblom assumed that any given vowel system would essentially gravitate towards an equilibrium which maximized the amount of contrast between all pairs of vowels in the system:

(1.6) $\sum_{i=1}^{m} 1 / r_i^2 \quad \rightarrow \quad \text{minimized}$

Liljencrants & Lindblom adapted this interpretation of perceptual contrast maximization into a computational simulation and discovered that it yielded results which fairly closely predicted the more common types of vowel systems in the languages of the world. For five-vowel systems, for instance, Liljencrants & Lindblom's model predicted:

(1.7) i u
ε Y
F/a

This system closely approximates the most common five-vowel systems in language, and is certainly far-removed (acoustically) from Chomsky & Halle's unlikely vowel system (1.4). Liljencrants & Lindblom therefore concluded from the results of this study that "...contrast plays an important role as a determinant of the structure of vowel systems. It is likely that it should be included among the variables in an explanatory phonological theory." (854)

Liljencrants & Lindblom's model of perceptual contrast did not, however, make perfect predictions about the structure of all vowel systems; it had an idiosyncratic predilection for high, central vowels such as [V] and [œ] and failed to generate [r] in any system. Liljencrants & Lindblom therefore suggsted that including articulatory factors might help their model make more accurate predictions: "It seems reasonable to suppose that a vowel system which has been optimized with respect to communicative efficiency consists of vowels that are not only 'easy to hear' but also 'easy to say.' Consequently a further improvement of the present theory might be obtained if we found a way of quantifying and incorporating 'ease of articulation'."

Some ten years later, Lindblom et al. (1983) came up with just such a system of quantifying "ease of articulation." Lindblom et al. combined this measure of articulatory ease with an adapted measure of perceptual discriminability in an attempt to derive a likely phonological inventory of CV syllables. Lindblom et al. measured the "articulatory ease" of a CV syllable by calculating the biomechanical cost of making the necessary articulatory gestures, as represented in their own production model (described in Lindblom and Sundberg 1971 and Lindblom, Pauli and Sundberg 1974). Their numbers for "biomechanical cost...represented 'penalty scores' that depended on the extent and rate of the given syllable's component movements." In order to determine the perceptual distance between CV syllables, Lindblom et al. (1983) developed a general measure of distance between auditory spectra:

(1.8) $$d(i,j) = [c \int_0^{24.5} |E(i,x) - E(j,x)|^p dt]^{1/p}$$

(where p=2, x is frequency in Barks, E(x) is an 'auditory spectrum' and 0 to 24.5 is the span of human auditory sensitivity in Barks). Lindblom et al. thus calculated the perceptual distance between two syllables by summing up this measure over the entire series of time slices sampled from both CV sequences. Their resultant measure of syllable discriminability--d(i,j)--became the mathematical basis for their perceptual optimization criterion:

(1.9) $$\sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{1}{(d(i,j))^2} \quad \rightarrow \quad \text{minimized}$$

Articulatory ease (a(i,j)) was then incorporated into this equation as an antagonistic force to perceptual discriminability:

(1.10) $$\sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{1}{(\frac{d(i,j)}{a(i,j)})^2} \rightarrow \quad \text{minimized}$$

Highly discriminable contrasts minimized the overall value of this equation, while correspondingly high scores of articulatory difficulty increased it. The ratio of (d(i,j)/(a(i,j)) was therefore a mathematical measure of how to produce the "maximal perceptual effect with the minimal articulatory means"--or, rather, how to get the most perceptual "bang" for one's articulatory "buck".

Lindblom et al. (1983) implemented this equation in a computational simulation to generate a set of fifteen optimal syllables in a developing phonological system, starting off with a different initial syllable in each of 133 runs. They found that their optimization algorithm selected the following fifteen syllables most often:

(38)	bi	bɛ	ba	bo	bu
	di	dɛ	da	do	du
	g_{+}i	g_{+}ɛ	ga	go	gi

Lindblom et al. also ran these simulations without including the ease of articulation factor and discovered that, under these conditions, the system preferred more marked consonantal onsets such as [ɖ] and [G]. Lindblom et al. also pointed out that onset-offset "assimilations" in the system were "caused by the fact that higher penalties were assigned to g_{+}u g_{+}o and gi, gɛ than to g_{+}i g_{+}ɛ and go gu." Since such results reflect common articulatory trends in the languages of the world, Lindblom et al. (1983) concluded that it was not unreasonable to suspect that both perceptual discriminability and articulatory ease played important roles in the development of linguistic sound systems.

The evolution of Liljencrants & Lindblom's (1972) model therefore provides one more example of how articulatory forces eventually develop an important role in theories which initially focus on the influences of perception on phonology. Hume & Johnson (2001), however, in attempting to make the case (yet again) for the importance of studying perception in phonology, suggest that the "articulatory bias" in phonological theory may have arisen for practical reasons alone. They endorse the independent study of perceptual influences on phonology and emphasize the advances in technology which can now offer new insight into the "interplay" of perception and phonology. This sentiment echoes the wave of interest in perception which emerged in phonological theory in the wake of Liljencrants & Lindblom (1972) and Lindblom et al. (1983) (in, e.g., Ohala 1981, Kohler 1990, Steriade 2001).

This dissertation suggests that, on the contrary, there are certain sound patterns in phonology which are best described in articulatory terms because they have no perceptual motivation whatsoever. Moreover, it shows that enthusiasm for perceptual interpretations of phonological phenomena may lead to incorrect assumptions about why certain sound patterns commonly occur in language. This dissertation specifically looks at the role that theories of perception in phonology have played in the exegesis of cross-linguistic patterns in place assimilation. A series of empirical studies follows which provides evidence that shows how articulatory factors influence the tendency of nasal stops to undergo place assimilation, regardless of the perceptual salience or distinctiveness of these particular segments. This evidence suggests that phonology cannot exist without articulatory descriptions or analyses--and that efforts to understand phonological processes entirely in terms of perception may lead to misguided assumptions about why these processes are occurring. Ironically, the history of phonological theory seems to show that it takes a shift of theoretical focus towards perception to make this seemingly obvious fact clear.

CHAPTER 2

BACKGROUND AND MOTIVATION: THEORETICAL AND EMPIRICAL APPROACHES TO UNDERSTANDING PLACE ASSIMILATION

2.1. Which kinds of place assimilation occur?

Nasals undergo place assimilation in more languages of the world than stops. While phonologists have long observed that nasals seemed to be particularly susceptible to place assimilation, this fact has only recently been confirmed by a series of cross-linguistic surveys of assimilatory processes (Cho 1990, Mohanan 1993, Jun 1995, Shin 2000). Three of these surveys--Mohanan (1993), Jun (1995) and Shin (2000)--point out that there is also an implicational relationship between nasals and stops as targets of place assimilation: there are no languages in which stops assimilate without nasals doing so as well, even though there are languages in which nasals but not stops assimilate (e.g., Chiyao, Ponapean, Yoruba), and languages in which both groups of sounds assimilate (e.g., Malayalam, Catalan, Korean.).

The research goal of determining which kinds of place assimilation are more or less likely to occur, cross-linguistically, has taken on a special theoretical significance in phonology ever since Chomsky & Halle (1968) pointed out that their own evaluation in SPE incorrectly predicted which kinds of assimilations ought to be common or rare. Formal treatments of place assimilation since Chomsky & Halle (1968) have, therefore, generally evolved so as to more gracefully represent the commonly occurring assimilations while making it impossible (or much more difficult) to represent assimilations which only occur rarely.

The treatment of place assimilation held no distinct formal status in the phonological framework first proposed by Chomsky and Halle (1968). As with all other phonological rules in SPE, assimilatory processes operated in the simple terms of changing values in the unordered feature matrices that were used to represent speech sounds. For instance, Chomsky and Halle accounted for a cross-morphemic process of place assimilation in southern Paiute (as described by Harms 1966) with the following "linear" rule:

$$(2.1.1)\,[+\text{cons}] \rightarrow \begin{bmatrix} \alpha ant \\ \beta cor \\ \gamma high \\ \delta back \end{bmatrix} \quad / \quad --- \quad + \quad \begin{bmatrix} +cons \\ \alpha ant \\ \beta cor \\ \gamma high \\ \delta back \end{bmatrix}$$

Chomsky and Halle deemed this Greek-variable notation a satisfactory account of assimilation, which they had loosely defined as "a process in which two segments are made to agree in the value assigned to one or more features." With respect to this process in Southern Paiute, they claimed that the Greek-variable notation captured "the fact that...the agreement of the consonants in terms of the features 'anterior,' 'coronal,' 'high,' and 'back' is linguistically significant," because it does not permit "arbitrary" rules such as the following:

$$(2.1.2)\ \ [+\text{cons}] \rightarrow \begin{bmatrix} +ant \\ -cor \\ -high \\ -back \end{bmatrix} \quad / \quad \text{---} \quad + \quad \begin{bmatrix} -ant \\ +cor \\ -high \\ +back \end{bmatrix}$$

Rule (2.1.2) is deemed "arbitrary" because the "relation between the change effected and the determining context is entirely fortuitous."

Chomsky & Halle's system for predicting which phonological processes were more likely to occur simply involved counting the number of distinctive features which were involved in each linear rule. The more features that were involved in the process, the less likely it was to occur. In applying this theory to actual data, however, Chomsky & Halle realized that their predictive mechanism was, in many cases, making the wrong predictions. For instance, the evaluation metric predicted that process (2.1.4) would occur more often than process (2.1.3), even though, in reality, just the opposite had been observed:

$$(2.1.3)[+\text{nasal}] \rightarrow \begin{bmatrix} \alpha ant \\ \beta cor \end{bmatrix} / \text{---} \begin{bmatrix} \alpha ant \\ \beta cor \\ C \end{bmatrix}$$

$$(2.1.4)\ [+\text{nasal}] \rightarrow \begin{bmatrix} +ant \\ \alpha cor \end{bmatrix} / \text{---}\ [\alpha\text{cor}]$$

Other criticisms of Chomsky & Halle's linear, feature-changing approach emerged soon after the publication of The Sound Pattern of English. Bach (1968), for instance, pointed out that the Greek-variable notation used for assimilatory processes permit apparently "arbitrary" rules such as:

$$(2.1.5)\ X \rightarrow \begin{bmatrix} \alpha cor \\ \beta ant \end{bmatrix} \quad / \quad --- \quad \begin{bmatrix} \alpha ant \\ \beta cor \end{bmatrix}$$

The relationship between the "determining context" for (2.1.5) and the features changed in the output is, once again, "fortuitous," even though Chomsky and Halle's Greek-notation formalism makes no explicit distinction between this rule and an assimilatory rule such as (2.1.1). It soon became apparent to phonologists that Bach's criticism had relevance not only because of its implications for the formal simplicity or arbitrariness of phonological rules, but also because it revealed that the formalisms of SPE permitted the expression of rules which were unattested among the phonologies of the world's languages.

Cressey (1974) pointed out another problem with the "naturalness" or "likelihood" of certain processes permitted by Chomsky and Halle's Greek-variable notation. Cressey noted that Greek-variable notation might ensure the appropriate assimilatory connection between context and output in both rules (2.1.6) and (2.1.7), but that (2.1.6) seemed to be a far more likely process than (2.1.7):

$$(2.1.6)[+nasal] \rightarrow \begin{bmatrix} \alpha coronal \\ \beta anterior \\ \gamma back \\ \delta distributed \end{bmatrix} \quad / \quad --- \quad (\#) \begin{bmatrix} +obstruent \\ \alpha coronal \\ \beta anterior \\ \gamma back \\ \delta distributed \end{bmatrix}$$

$$(2.1.7)\ [+\text{nasal}] \rightarrow \begin{bmatrix} \alpha anterior \\ \beta coronal \\ \gamma voice \\ \delta continuant \end{bmatrix} \quad / \quad \text{---} \quad (\#) \begin{bmatrix} +obstruent \\ \alpha anterior \\ \beta coronal \\ \gamma voice \\ \delta continuant \end{bmatrix}$$

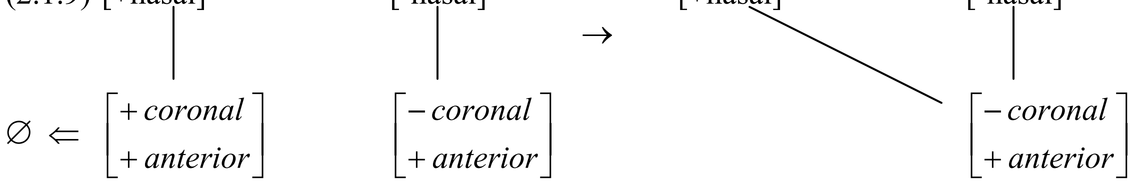
It is more likely, in other words, for the features [coronal, anterior, back and distributed] to participate together in a rule than it is for the features [anterior, coronal, voice and continuant]. As Cressey noted, "certain subsets of features (tend to, may) cluster together in rules involving the Greek variable notation, that is, assimilation and dissimilation rules. The evaluation metric proposed in SPE fails to achieve this result." Cressey proposed to solve this problem by incorporating a "quasi-feature" into the SPE matrix framework called [PA], which stood for "place of articulation." Operationally, [PA] became a cover feature for [high, back, coronal, anterior, and distributed]—all of the features in the SPE framework which together determined what would traditionally be called the "place of articulation" of some segment. With this new "quasi-feature," Cressey could simply abbreviate the rule in (2.1.5) with (2.1.8):

$$(2.1.8)\ [+\text{nasal}] \rightarrow [\alpha\text{PA}] \quad / \quad \text{---} \quad \begin{bmatrix} +obstruent \\ \alpha PA \end{bmatrix}$$

Since (2.1.8) was formally simpler than (2.1.6) (in terms of the SPE evaluation metric), its comparative "naturalness" or "likelihood" simply fell out from the newfound formalism.

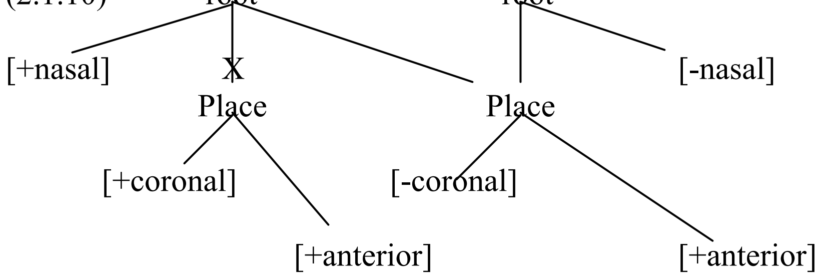
A distinct formal status for place assimilation in phonological theory emerged not long after Cressey (1974). Working on an analogy with the formalisms of "autosegmental phonology"—as first proposed by Goldsmith (1976)—works such as Halle & Vergnaud (1980), Goldsmith (1981) and Steriade (1982) conceived of place assimilation as a "spreading" process. Goldsmith (1981), for instance, illustrated place assimilation as spreading with rule (2.1.9), where "place of articulation" was determined by the combined features [coronal] and [anterior]:

(2.1.9)



Autosegmentalists would eventually incorporate Cressey's notion of a "Place" super-feature into this formalism to make the process of place assimilation explicit:

(2.1.10)



Clements (1985) proposed that the organization of features in assimilation processes set the methodological standard for understanding how sub-segmental features related to one another in autosegmental phonology: "If we find that certain sets of features consistently behave as a unit with respect to certain types of rules of assimilation or resequencing, we have good reason to suppose that they constitute a unit in phonological representation, independently of the actual operation of the rules themselves." This line of thinking would later be elaborated by Sagey (1986) and Hume (1994) (among others) to develop the highly-structured view of the segment proposed in, for instance, Clements and Hume (1996).

This organization of features in the segment not only reflected common (and attested) groupings of features in phonological rules but also constrained potential phonological processes by determining which groups of features could *not* function together in rules. Clements (1985) noted the theoretical consequences of such representations of the segment: "Such a theory of phonological representation offers a constrained theory of assimilation processes, according to which all assimilation rules involve the spreading of a single node: the root node, a class node, or a feature node." Restraining phonological representations in this way helped the "assimilation as spreading" view avoid the over-predictability engendered by the earlier SPE formalism. Representational restrictions also made the autosegmental view of assimilation more falsifiable (and therefore stronger) in that it explicitly predicted that certain sets of features should *never* function together in assimilatory processes.

The "assimilation as spreading" view also differed from the SPE framework in that it made the connection between the assimilatory output and the assimilatory context explicit. Whereas assimilation in SPE was a post-hoc interpretation of a rule involving Greek notation, "assimilation as spreading" resulted in a doubly-linked structure in the autosegmental view. As Sagey (1986) notes, (2.1.10)

> "...captures the fact that in assimilation, a segment changes to become identical to some segment in its environment with respect to certain features; the features of the trigger are simply realized on the target. There is no way for a feature not in the environment to end up on the target in a spreading assimilation. In contrast, changing feature values can, in principle, change neighboring segments to opposite values of the context feature, change the value of an unrelated feature, or even affect

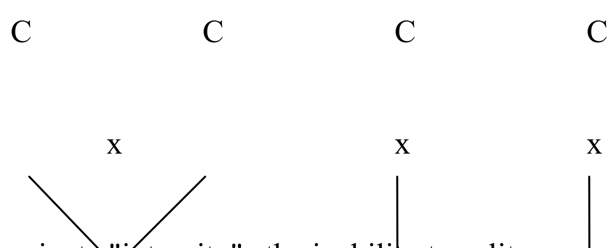
> segments not in the immediate environment. Such processes are extremely uncommon in comparison to assimilations where the target takes on some feature in the environment."

"Spreading" assimilations were thereby restricted (once again) by being limited to the immediate phonological environment; this restriction helped capture the fact that assimilatory rules essentially involve a connection between their output and their conditioning environment.

Representing the output of assimilation as a doubly-linked structure also helped account for certain inalterability effects that the linear framework could neither represent nor predict. Researchers such as Kaye (in Halle and Vergnaud, 1980) and Steriade (1982) had noted that "true" (i.e., underlying) geminates exhibit "integrity" and "inalterability" effects whereas their "fake" (i.e., derived) counterparts do not. This asymmetry easily fell out from the autosegmental formalism because underlying geminates were represented as doubly-linked structures while derived geminates were independently linked sequences of identical segments.

(2.1.11) (from Hayes 1986)

"True" geminates "Fake" geminates



True geminate "integrity"--the inability to split up a geminate consonant with epenthesis or metathesis--thereby followed from the universal representational restriction on crossing association lines. True geminate "inalterability"--the failure of geminate consonants to exhibit the expected effects of certain phonological rules--followed, in turn, from the doubly-linked geminate's inability to meet the structural description for the particular rules in question. Since the doubly-linked structure of geminates accounts for both their integrity and inalterability effects in autosegmental phonology, it was predicted that similar effects should be found in the doubly-linked output structures of assimilation. Steriade (1982) and Hayes (1986) did, indeed, find inalterability effects in Kolami and Toba Batak, respectively, thereby strengthening the argument for the representation of assimilation as a spreading process.

Spreading analyses of place assimilation such as Goldsmith (1981), Clements (1985) or Sagey (1986) typically assumed that the underlying place features of the target simply disappeared or "de-linked" for free on the way to reaching the phonetic surface. The framework of radical underspecification, however--as originally outlined by Kiparsky (1982) and Archangeli (1984)--provided phonologists with a subtly different option for representing place assimilation. In radical underspecification, segments are not specified for unmarked features in their underlying forms; these only first appear as the result of later "feature-filling" rules that ensure the well-formedness of a segment at the surface. Mascaro (1983) and Kiparsky (1985) interpreted such underspecification as an explanation for why only coronal nasals undergo place assimilation in Catalan. The radical underspecificationists had determined—for independent reasons—that coronals were "unmarked" for place of articulation at some point in the derivation of the output phonological forms in Catalan. Mascaro (1983) and Kiparsky (1985) therefore argued that place assimilation targeted only coronal nasals in Cataland precisely because these segments were unspecified for place of articulation.

> "..the outstanding question that we wish to answer is why only the coronals assimilate to all places of articulation, while the labials assimilate only in a limited way and the palatals and velars do not assimilate at all...The answer, again, is that the coronal nasals, being unmarked, are unspecified for place of articulation when Nasal Assimilation applies, and that Nasal Assimilation associates specified (and therefore marked) feature values in autosegments with segments that do not carry those feature values on autosegments."

Kiparsky (1985) specifically represented the process of coronal nasal assimilation in Catalan as:

(2.1.12)

Root Root

SLN SLN

Place

Articulation

This alternative autosegmental formalization of place assimilation had a number of interesting implications. Since only coronals were considered unmarked radical underspecification theory, it predicted, for one, that there should be languages in which only coronals assimilated to non-coronals but no languages in which only non-coronals assimilated to coronals and other non-coronals. It also seemed to account for the fact that assimilations targeting non-coronals--which involved both linking and delinking processes--were much less common than the feature-filling assimilations which targeted only coronals (Paradis and Prunet, 1991).

Cho (1988) elaborated on this underspecificationist approach while analyzing place assimilation in Korean. This (typically casual speech) process in Korean includes the following assimilations:

(2.1.13) Coronals assimilate to labials, palatals and dorsals

(2.1.14) Palatals and labials assimilate only to dorsals

(2.1.15) Dorsals do not assimilate

Since Cho (1988) specified place with only two features, [coronal] and [anterior], she was able to account for these asymmetries in Korean place assimilation by specifying place with only two features, [coronal] and [anterior]. Cho observed that, if the place specifications of an assimilatory target formed a subset of the place specifications of the assimilatory trigger, then assimilation would ensue. If this was not the case, then there would be no assimilation. For instance, labial ([-cor]) or palatal ([-ant]) segments assimilated to dorsals ([-cor, -ant]), but not vice versa--and neither palatals nor labials assimilated to each other. Coronals assimilated to all other places of articulation because their null set of place features formed a subset of all other place specifications. In Korean, therefore, place assimilation was not just a case of "unmarked" segments assimilating to "marked" segments but one of "less marked" segments assimilating to "more marked" segments. Cho's (1990) dissertation later accounted for the fact that no other language exhibits this asymmetry in assimilation by stipulating that Korean had a language-specific "parameter" which determined its targets and triggers in place assimilation.

2.2. Why does place assimilation occur?

After Cho (1990), some researchers returned to the hypothesis that functional forces external to formal phonology--such as the need for maximum discriminability and minimal articulatory effort--were motivating place assimilation to occur, whenever it did occur. These researchers therefore proposed that any adequate formal account of place assimilation needed to incorporate the phonetic forces which caused this process to happen.

Many of these researchers took their inspiration from the earlier work of Liljencrants & Lindblom (1972) and Lindblom et al. (1983). Mohanan (1991, 1993), for instance, claimed that phonology-external, phonetic forces—rather than underspecification or relative "markedness"--could best account for why certain types of segments were more likely to undergo place assimilation, cross-linguistically. Mohanan (1991) pointed out that only coronals undergo place assimilation in casual speech in English, even though this is supposedly at the level of post-lexical phonology, where all segments should be fully specified. Mohanan (1993) also pointed out that there are a number of cross-linguistic asymmetries in place assimilation, almost none of which can be accounted for in the framework of radical underspecification. In a brief survey of assimilatory processes in English, Korean, Malayalam and Hindi, Mohanan found that the following generalizations held true:

"(2.2.1) Asymmetries in Place Assimilation

a. Coronal Asymmetry

(i) If noncoronals undergo assimilation, so do coronals.

(ii) If coronals trigger assimilation, so do noncoronals.

b. Labial-velar Asymmetry

If labials trigger assimilation, so do velars.

c. Stop Asymmetry

(i) Nonstops do not undergo (the whole range of) assimilation.

(ii) If nonstops trigger assimilation, so do stops.

d. Sonorant Asymmetry

(i) If nonsonorants undergo assimilation, so do sonorants.

(ii) If sonorants trigger assimilation, so do nonsonorants."

Mohanan accounted for these generalizations not with any new representational device but rather by stipulating that they fell out from specific principles in Universal Grammar (UG). For instance, Mohanan claimed that the following principle existed in UG:

(2.2.2.) In the sequence [+stop][+cons], the two consonants must share a single place node.

Mohanan went on to say that "Individual languages must specify the domain in which the requirement [(2.2.2)] holds"--i.e., each language makes parametric decisions about how to implement this universal restriction.

Since Mohanan maintained that the radical underspecification approach--and its attendant notions of markedness--were unable to account for the generalizations in (2.2.1), he proposed the existence of a phonological quality known as "dominance" to account for the cross-linguistic patterns. Mohanan claimed that "dominance" is also specified in UG; his definition of it ran as follows: "The property that makes linguistic units survive I will refer to as <u>dominance</u>. Let us characterize this notion as follows:

(2.2.3) A dominant unit resists the forces that alter its properties."

Mohanan maintained that velar consonants were more dominant than palatals or labials--which, in turn, were more dominant than coronals. This dominance hierarchy would straightforwardly explained the patterns of place assimilation in Korean; however, Mohanan claimed that languages like Hindi and Malayalam were "insensitive to it"--the only restriction being that "no language...may reverse the dominance scale." Similarly appropriate dominance rankings accounted for the rest of the generalizations Mohanan found in (2.2.1).

Mohanan admitted that it was important to question how this dominance hierarchy had become a part of UG. His answer to this question suggested that dominance was part of the phonological residue of phonetic constraints on communication. "This question can only be answered only by appealing to nonlinguistic systems that shape the nature of UG, including the human articulatory and auditory systems, and the requirements of human communication." That is, dominance appears in UG as the result of interactions between forces which exist outside of the phonological component of UG itself--

> "We must acknowledge that assimilatory phenomena are archetypal instances of phonetically motivated phonological phenomena. Following Lindblom (1988) and others, we may suggest that the physical motivation for place assimilation is to be sought in the optimization of two competing requirements, namely, (i) the minimization of articulatory cost, and (ii) the maximization of discriminability."

In a series of papers, Kohler (1990, 1991, 1992) attempted to show just how perception and articulation might interact to induce synchronic reduction processes like place assimilation. Kohler (1990) constructed his theory on the basis of observations he made of reduction phenomena which occur often in casual speech in German. Kohler's interpretation of *why* certain reductions--but not others--were occurring in fast speech was derived from the older, Lindblomian idea that speakers try to get the "maximal perceptual effect with the minimal articulatory means." Kohler's twist on this notion was that "What is not very distinctive for a listener anyway may be reduced by a speaker more easily to yield to the principle of economy of effort." That is, instead of just making the minimal articulatory effort necessary to insure that listeners understood everything that was being said, speakers--in Kohler's view--were actively looking for opportunities to reduce the amount of articulatory effort they had to expend. Such opportunities were determined by their perceptual weakness; they arose in segments or phonological environments of low perceptual salience. Speakers--who only speak with a model of listeners' abilities and needs in mind--realize in these circumstances that extra articulatory effort will not help listeners perceive what they are saying. The speakers will therefore reduce the amount of articulatory effort they spend on such highly confusable segments, and instead produce a reduced or assimilated articulation that is easier to say.

> "Speakers not only control reduction with regard to the physiological and articulatory potentials combined in the dynamics of sound production but also take listeners into account and adapt to their needs in two ways :
> (a) Reduction processes are favoured that show a low degree of perceptual salience...
> (b) Different communicative situations put different demands on the perceiver of speech, and speakers have to tune their performance to these conditions to guarantee a successful language interaction (cf. Lindblom's H & H theory)."

Assimilatory processes, for instance, would occur in casual speech in German not only because assimilated forms are easier to say, but also because the assimilated and unassimilated forms are highly confusable with one another. Unlike Lindblom's earlier framework, then--where articulation and perception had an essentially antagonistic relationship with one another--Kohler's theory states that reduction processes occur whenever demands for ease of articulation intersect perceptual weakness.

Jun (1995) combined Kohler's (1990) ideas about the interaction of perception and articulation with an expanded version of Mohanan's (1993) cross-linguistic survey of place assimilation. However, unlike Mohanan--who considered perception and articulation to be external to formal phonology--Jun incorporated these phonetic influences directly into a formal treatment of place assimilation in Optimality Theory. Jun's approach followed Kohler's idea that speakers know what is difficult to perceive and easy to say, and that they use this knowledge to make on-line decisions about how to transmit a message to a listener--or, as the case may be, how to implement phonological reductions in casual speech.

Within this theoretical framework, then, Jun extended Mohanan's (1993) typology of place assimilation by cataloguing the process in some twenty-five different languages. Despite the greater extent of his survey, Jun reached similar conclusions as Mohanan about the cross-linguistic asymmetries therein:

(2.2.4) From Jun (1995):

"a. Target manner

(i) If fricatives or nonnasal sonorants are targets of place assimilation, so are stops.

(ii) If stops are targets of place assimilation, so are nasals.

b. Target place

(i) If velars are targets of place assimilation, so are labials.

(ii) If labials are targets of place assimilation, so are coronals.

c. Syllable position

If the onset is a target of place assimilation, so is the coda.

d. Trigger manner

(i) If nonnasal sonorants trigger place assimilation, so do nasals and fricatives.

(ii) If nasals or fricatives trigger place assimilation, so do stops.

e. Trigger place

If coronals are triggers, so are velars."

Jun proposed that such invariant cross-linguistic implicational relationships (i.e., which Mohanan located in UG) fell out from a number of invariant "harmonic" constraint rankings in Optimality Theory (in the sense of Prince and Smolensky, 1993). These universal constraint rankings reflect the influence of universal articulatory and perceptual forces. In order to motivate this phonetic interpretation of assimilatory phenomena, Jun cited results from a number of studies on the perception of specific cues for place of articulation to establish the following universal hierarchy of place salience:

(2.2.5) Dorsal¬ > Labial¬ > Coronal¬ (where > denotes "more salient than")

Jun decided to incorporate such universal salience hierarchies into the OT constraint pantheon in the form of harmonically-ranked "preservation" constraints. Following Kohler's (1991) contention that, "What is not very distinctive for a listener anyway may be reduced by a speaker more easily to yield to the principle of economy of effort," Jun maintained that less salient places of articulation would be less likely to be "preserved" (i.e., produced in the ouput) by speakers. Segments with less salient place of articulation cues would, therefore, be more likely to become targets of place assimilation. Perceptual salience could therefore translate directly into the formal, OT representation--"Constraints preserving acoustically more salient segments must be ranked above those preserving acoustically less salient segments."

Since Jun conceived of perceptual salience as being determined by universal, extra-linguistic forces, this model could only represent language-specific differences in the targets of place assimilation by showing how these "preservation" constraints interacted with an articulatorily-based "weakening" constraint. This weakening constraint competed with preservation constraints to reduce articulations past the point of perceptibility and, if properly ranked, ensured that assimilation took place.

(2.2.6) Malayalam: [sam giitam] → saŋgiitam 'GLOSS!'

Candidates	Pres (mnr(nas))	Pres (mnr(-cont)	Pres (pl(onset))	Weak	Pres (pl([nas]C))	Pres (pl(coda))
☞saŋ giitam				*	*	*
sam giitam				**!		
sa giitam	*!	*		*	*	*
sam piitam			*!	*		

Jun accounted for the complex pattern of place assimilation in Korean, then, by re-ranking a weakening constraint above the appropriate preservation constraints. Pres (dor ¬), for instance, would be ranked higher than the Weak constraint in Korean; hence, dorsals do not undergo assimilation, while labials and coronals (whose preservation constraints are ranked below Weak) do. The trigger asymmetry in Korean was also accounted for by a perceptually-motivated constraint ranking that differentiated between constraints preserving place cues before a coronal (Pres(pl(_cor)) and before a noncoronal (Pres(pl(__noncor)).

(2.2.7a) Korean: /ip + ko/ → [i{p} Ko]

Candidates	Pres (pl(onset))	Weak	Pres (pl(___C)	Pres (Pl(___noncor)	Pres(pl(lab¬))
ip ko		**!			
☞i{p} Ko		*	*	*	*
iP {k}o	*!	*			

Here, {p} represents a labial articulation reduced past the point of perceptibility, while K represents a dorsal articulation perceived as a geminate.

(2.2.7b) Korean: /ip + ta/ → [ipta]

Candidates	Pres(pl(onset))	Pres(pl(___cor))	Weak	Pres(pl(lab¬))
☞ ipta			**	
i{p} Ta		*!	*	*
iP {t}a	*!		*	

Jun's Optimality Theoretic approach was therefore capable of accounting for numerous language-specific differences in assimilation processes, but its reliance on language-universal harmonic rankings of preservation constraints ensured that it still predicted that certain assimilatory asymmetries would not be found in any language. For example, Jun's ranking of labials as perceptually more salient than coronals entailed that there could not be a language which ranked Pres(cor¬) higher than Pres(lab¬). No language could therefore have an assimilatory asymmetry whereby labials underwent assimilation even though coronals did not. Such predictions are in keeping with the cross-linguistic generalizations Jun found in (2.2.4).

Jun's emphasis on the reduction of segments past the point of perceptibility in place assimilation represents a significant departure from the interpretation of casual speech assimilations which had been proposed by Browman & Goldstein (1986, 1989, 1990), in developing their framework of Articulatory Phonology. This theoretical framework represented speech acts not as a sequence of features tied together into phonemic wholes, but rather, as a collection of gesture being implemented on a number of independent articulatory "tiers." In a basic description of this model, Browman & Goldstein (1990) portray each of these articulatory tiers as an aperture in the articulatory tract which may be independently opened or closed by a speaker in producing speech. Browman & Goldstein's (1990) represented the English word "palm" ([pa:m]), for instance, with a series of opening and closing gestures on four separate articulatory tiers:

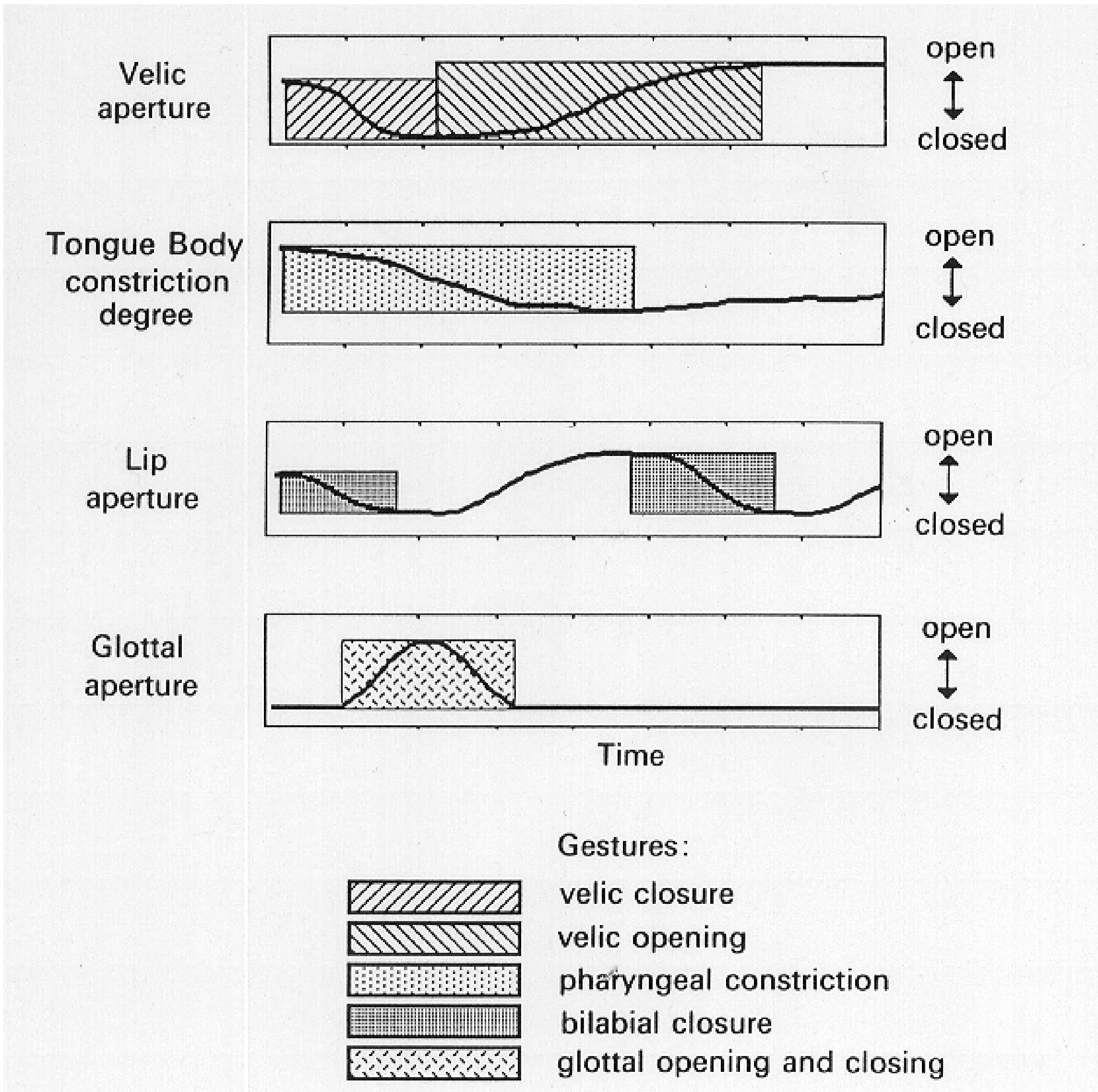


Figure 2.1: Hypothetical gestural scores and articulatory gestures for [pa:m], from Browman & Goldstein (1990)

Figure 2.1 shows that, while opening and closing gestures may be made on any one of these tiers independently of gestures on the other tiers, these gestures must nonetheless be highly coordinated with each other in order to produce the intended acoustic effect. Browman & Goldstein refer to such coordination between gestures as a “gestural score” and suggest that errors in the implementation of these gestural scores may account for certain reduction phenomena often found in casual speech.

Browman & Goldstein (1990), for instance, hypothesize that phase-shifting the onsets of certain articulatory gestures with respect to other gestures in their model may account for a variety of casual speech phenomena in English. They cite the following examples of casual speech processes in English from Brown (1977)—

(2.2.8a) “must be”	→	[ˈmʌsbi]	(consonant deletion)
(2.2.8b) “hundred pounds”	→	[hʌndrəb ˈpaʊndz]	(consonant assimilation)
(2.2.8c) “ground pressure”	→	[graʊm ˈprɛʃɚ]	(simultaneous deletion and assimilation)

--and claim that they all may be accounted for by the “gestural overlap” of the consecutive consonantal articulations. Assimilation may occur in (2.2.8b), for instance, if the labial closing gesture shifts forward in time so that it significantly obscures the closing gesture for the final /d/ in “hundred,” thereby rendering it imperceptible.
The deletion in (2.2.8a) may, on the other hand, occur if the labial closure for the /b/ in “be”—and its corresponding glottal aperture closing gesture—completely precede the closing gesture for the preceding /t/ in “must.” Browman & Goldstein stress the idea that both of these processes may be accounted for along the same continuum of gestural overlap: “...these examples of consonant assimilation and consonant deletion are all

hypothesized to occur as a result of increasing gestural overlap between gestures on separate oral tiers."

Importantly, Browman & Goldstein's gestural overlap analysis implies that the "assimilated" or "deleted" gestures are still articulated, even though their perceptual effects may be obscured. This interpretation of place assimilation in casual speech therefore differs significantly from that given in autosegmental phonological analyses such as Kiparsky (1985) or Hayes (1986) in that the place of articulation of the target is not replaced by the place of the trigger, but merely overlapped by it. Autosegmental analyses could also only represent place assimilation as an all-or-nothing process—either the target adopted the place of articulation of the trigger or it did not. Browman & Goldstein's interpretation of the process, however, allowed for a broad continuum of assimilations and deletions, depending on the amount of overlap between the target and the trigger gestures.

Browman & Goldstein (1990) presented evidence for their theory that gestures which had been "assimilated" or "deleted" in casual speech were perceptually "hidden" but actually still present, articulatorily, by citing x-ray microbeam data on the production of coronal stops in phrases such as "perfect memory" or "nabbed most." Similar evidence—from electropalatography (EPG) studies was also presented by Barry (1991) and Nolan (1992). Both Barry and Nolan found evidence that more than just the phase-shifting of gestures was involved in the casual speech assimilations/deletions of coronal stops. Both of these researchers found that, in some cases, speakers sometimes reduced the extent of the closure in the coronal gesture that was being assimilated. Barry (1991) reported, for instance, "incomplete" or "weakened" coronal gestures for sounds such as

/d/ in “hand grenade,” wherein contact was not made by the tongue at the alveolar ridge, even though there was contact further back, along both sides of the palate. Nolan (1992) reported a similar articulatory pattern—which he termed “residual alveolar”—for /t/s in phrases such as “late calls” or “boat covered.” The fact that such “residual alveolars” appeared in these studies primarily when these alveolars preceded velar consonants may be significant. Since velars are produced further back in the oral tract than alveolars, their production would not obscure that of preceding alveolars, even if their onsets were phase-shifted forward in time. Residual alveolars may indicate, therefore, that speakers do not rely on gestural overlap to induce assimilation or deletion in /t-l/ sequences, but rather, adopt an alternative articulatory strategy to achieve a similar assimilatory goal. Both Barry (1991) and Nolan (1992) therefore concluded that such reduction of alveolar gestures was under direct speaker control.

Zsiga (1994) similarly concluded that gestural overlap was also under speaker control, based on the results of a study she performed on the interaction of gestural overlap in English coronals with speaking rate. Zsiga (1994) investigated gestural overlap by having English speakers produce coronal stops before labial and dorsal consonants in basic phrases such as “bed pan” and “bad kick” in written sentences. Zsiga inferred the amount of gestural overlap on the /d/ productions by measuring the effect of the following formant transitions into the alveolar stop. For instance, early onset of a labial closure would decrease the second formant transition into the coronal, while early onset of a dorsal closure would increase the second formant transition. Zsiga had speakers produce these phrases at both a normal and a fast speaking rate, and calculated the formant transition values for all speaker productions. Zsiga found that a faster speaking rate did not necessarily yield greater gestural overlap. Rather, the amount of overlap seemed to be more directly related (for both speaking rates) to the ratio of initial vowel duration (e.g., the /ɛ/ in "bed pan") to the duration of the consonant cluster (e.g., /dp/). Zsiga interpreted this result as indicating that gestural overlap occurred when speakers reorganized the timing relationships between gestures in the consonant clusters, and was not simply the product of a rapid rate of speech.

Jun (1995), on the other hand, ran a production study in an attempt to show that place assimilation did not occur as the result of gestural overlap at all. Jun (1995) tested Korean speakers' production of labial /p/ before /k/ (as in /ipko/) in both "casual" and "formal" speech, since this segment had been reported to undergo place assimilation in casual speech in this context (Kim-Renaud 1974, Cho 1990). Jun's had his Korean speakers produced these sequences while he measured the air pressure in the portion of their oral tract beetween the lips and the velum. On the basis of his air pressure readings, Jun inferred the existence of gestural overlap between the /p/ and the /k/ articulations whenever oral air pressure decreased before the release of the /p/ (as it would for a bilabial click). Jun also inferred the existence of gestural reduction—i.e., a failure to make a labial closure—whenever there was no change in oral air pressure during the targeted closure interval for the /p/ in the consonant cluster. Jun found that both overlapped and reduced productions of /p/ predominated in his speakers' productions; in casual speech, the reduced forms were by far the most common.

After interpreting his production data in this way, Jun played sample overlapped and reduced tokens to Korean and English listeners and asked them to identify the consonant sequence as either /pk/ or /kk/. The listeners in this task consistently identified the overlapped sequences as /pk/ and the reduced sequences as /kk/. Jun concluded from this result that only articulatory reductions (and not gestural overlap) led to the perceptual assimilation of a targeted segment. Like Kohler (1991), Barry (1991) and Nolan (1992), Jun maintained that such articulatory reductions were under speaker control. Jun also argued, on the basis of his perceptual evidence, that Browman & Goldstein's gestural overlap analysis of casual speech place assimilation was untenable. Jun therefore rejected the possibility that gestural overlap in casual speech might constitute a qualitatively different phonological process than the traditional all-or-nothing conception of place assimilation. Jun's (1995) cross-linguistic survey of assimilatory processes does not, therefore, distinguish between casual speech assimilations (as in English, Korean or Malayalam) and non-optional assimilations (as in Toba Batak, Diola Fogny, or Yoruba). Interestingly, the cross-linguistic patterns than Jun unearths in this survey apparently hold for both the casual speech and non-optional assimilatory processes in these various languages.

Ohala (1990) sketched out a theory of the interaction of perception and place assimilation that differed radically from the line of thought which had developed in Kohler (1990), Mohanan (1993) and Jun (1995). Ohala (1990) suggested, firstly, that "articulatory ease" did not motivate particular patterns in place assimilation at all; these were, instead, the result of perceptual influences alone. Secondly, Ohala (1990) claimed that these perceptual influences in no way "optimized" the structure of speech for the purposes of communication. Instead, listeners' attempts to retrieve information from the speech signal simply resulted in "innocent misapprehensions," which, in turn, induced subsequent changes in phonological processes. For this reason, Ohala (1990) did not portray place assimilation as occurring because of any speaker knowledge of the communicative efficacy of certain articulations or perceptual contrasts; instead, the cross-linguistic patterns that could be found in this process--just emerged from listeners' tendencies in listeners to make certain perceptual mistakes.

2.3. Evidence for perceptual influences on place assimilation

One important contribution that Ohala (1990) made to the theoretical debate about how best to analyze place assimilation is his insistence that claims about "ease of articulation" or "ease of perceptibility" ought to be backed up with hard, empirical evidence. For instance, Ohala skewered the notion of "ease of articulation" precisely because it was so difficult to observe, empirically:

> "Unfortunately, the notion of 'ease' or 'simplicity' has never been satisfactorily defined. It is true that when a heterorganic cluster becomes a geminate (and necessarily homorganic) or when a nasal assimilates in place to the following consonant, there is one less articulator involved but it does not follow so straightforwardly that this yields an easier task."

Without a clear, empirical definition of "ease of articulation," Ohala claimed that it was easy to make misguided assumptions about what really is "easy" or "difficult" to say. Ohala, for instance, suggested that "ease of articulation" could be more plausibly used to argue for the likelihood of progressive assimilations, despite the fact that regressive assimilations occur much more often, cross-linguistically:

> "Finally, and this is the crucial defect, the notion of ease of articulation fails to explain why, in the above cases [C_1C_2 sequences], it is typically C_1 which assimilates to C_2 and not vice-versa. A priori, it seems more

plausible that if degree of effort really mattered, C_1 is the consonant that should prevail in these assimilations, i.e. after supposedly 'lazy' speakers adopt a given articulatory posture, one would expect them to maintain it during C_2. That the opposite happens is sufficient reason to be highly suspicious of such accounts."

Ohala's own account of place assimilation therefore focused on both the perceptual reasons (i.e., innocent misapprehensions) that motivated this process and the experimental evidence which showed that listeners actually made these perceptual mistakes in real-life perceptual situations. Specifically, Ohala theorized that the context-varying nature of place cues might account for the fact that place assimilation is usually regressive and not progressive. For instance, the first stop consonant in a VC_1C_2V sequence often lacks a release burst, in which case it only has a post-vocalic transition cue to its place of articulation. The C_2 consonant, however, always has 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), who showed that pre-vocalic stop transitions are perceptually stronger than post-vocalic stop transitions. On the basis of this evidence, then, Ohala hypothesized that the combined strength of place cues for C_2 might overwhelm the corresponding place cues for C_1 perceptually, and thereby result in the listener perceiving the entire 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 assimilation.

In order to demonstrate the ability of the burst and transition cues for C_2 place to overwhelm 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. For example, Ohala spliced together, a post-vocalic /p/ transition leading into an intervocalic consonant with a pre-vocalic /k/ transition (preceded by a /k/ burst) leading out. Ohala also constructed a similar set of VNCV stimuli, with post-vocalic nasal transition cues leading into the intervocalic consonant. In a small case study, Ohala played these stimuli to listeners and asked them to identify the intervocalic consonant. In the case of the example above, the listeners 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, 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 asymmetry between nasals and stops as targets of place assimilation, since they showed no significant perceptual differences between these two groups of sounds.

Hura et al. (1992) adapted Ohala's empirical approach towards understanding perceptual influences on place assimilation in order to test Kohler's (1990) hypotheses about which kinds of segments would make likely targets of place assimilation. Since Kohler claimed that "speakers are likely to sacrifice articulations that would be difficult for listeners to hear," Hura et al. reasoned that place assimilation would only target segments which were difficult to distinguish from one another. For example, Kohler (1990) had also 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 which had VC_1C_2V sequences embedded in pairings of first and last names, with the break between names situated 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 result 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.

Nonetheless, Hura et al. (1992) and Ohala (1990) both found empirical evidence that perception motivated two different asymmetries in place assimilation: one, a preference for regressive over progressive assimilations, and two, a tendency for nasals and stops to undergo assimilation more often than fricatives. Even more such assimilatory asymmetries had emerged in the cross-linguistic surveys of place assimilation in Mohanan (1993) and Jun (1995), however--and Jun (1995) hypothesized that they all had perceptual origins. Specifically, Jun had proposed the existence of the following harmonic rankings of preservation constraints:

(2.3.1) Place of target: Pres(pl(dor$^{\neg}$]) >> Pres(pl(lab$^{\neg}$)) >> Pres(pl(cor$^{\neg}$))

(2.3.2) Manner of target: Pres(pl([stop]C)) >> Pres(pl[nasal]C))

(2.3.3) Position of target: Pres(pl(onset)) >> Pres(pl(coda))

(2.3.4) Trigger: Pres(pl(__cor)) >> Pres(pl(__noncor))

Each of these rankings accounted for a particular cross-linguistic asymmetry in assimilatory processes. Ranking (2.3.3), for instance, accounted for the asymmetry between regressive and progressive assimilations that Ohala (1990) had investigated. Furthermore, the ranking of each preservation constraint was determined by aspects of universal perception, since speakers were least likely to "preserve" what was difficult for listeners to perceive, according to the Kohlerian (1990) production hypothesis. Ranking (2.3.1) implied, then, that place cues in unreleased dorsals were less confusable than the same cues in unreleased coronals. This perceptual difference therefore accounted for the phonological fact that coronals undergo place assimilation more often, cross-linguistically, than dorsals.

While Jun's harmonic rankings matched up neatly with the phonological facts he had found in his cross-linguistic survey of place assimilation, his claims about the perceptual salience of the various types of speech sounds in the rankings were not based on the results of perceptual experiments, as in Ohala (1990) and Hura et al. (1992). Instead, Jun derived these claims from analyses of the salience of place cues for each sound type, in the appropriate contexts. For this reason, Jun only considered the perceptual salience of place cues in <u>unreleased</u> stops in (2.3.1), since he assumed (following Ladefoged 1975) that stops were always unreleased before other consonants--i.e., in an appropriate context for place assimilation to occur. Under these conditions, the only cues for stop place of articulation are the preceding vowel-to-consonant formant transitions. Jun proposed that these cues would be stronger, perceptually, in dorsal and labial consonants than in coronal consonants, since the transitions for coronals are relatively short. "Tongue tip gestures are rapid; thus, they have rapid transition cues. In contrast, tongue dorsum and lip gestures are more sluggish; thus, they have long transitions. Consequently, noncoronals have more robust perceptual cues than coronals." Jun used a similar line of thought to justify the perceptual distinction that he proposed existed between dorsals and labials:

> "Unlike labials and coronals, velars have an acoustic attribute, i.e., compactness (Jakobson, Fant and Halle, 1951). Velars can be characterized by a noticeable convergence of F2 and F3 of a neighboring vowel. These two formants can form a prominence in the midfrequency range. As argued and discussed by Stevens (1989), such a midfrequency prominence of velars can form a robust cue for place of articulation...Based on Stevens' claim, we assume that velars have an

> additional acoustic cue, i.e., compactness, for place of articulation, compared to coronals and labials."

On the basis of this analysis, then, Jun claimed that the harmonic constraint ranking in (2.3.1) reflected both the universal facts of perceptual salience, as well as cross-linguistic tendencies in place assimilation. Jun developed similar analyses of acoustic cues to justify the corresponding rankings in (2.3.1)-(2.3.4).

Winters (2001) tested Jun's hierarchies in (2.3.1)-(2.3.4) to determine if they corresponded to the actual ability of listeners to identify place of articulation in an experimental paradigm similar to those used by Ohala (1990) and Hura et al. (1992). To test Jun's rankings, Winters (2001) played a series of VCCV stimuli to listeners and asked them to identify the place of articulation of both consonants. These consonants had either labial, coronal or dorsal places of articulation--in order to test rankings (2.3.1) and (2.3.4)--and the first consonant was either an unreleased stop or a nasal, in order to test ranking (2.3.2). The listeners in this task identified the places of articulation of both consonants in each VCCV stimulus by choosing one option out of a three-by-three forced-choice set.

The results of Winters' (2001) study only confirmed two of Jun's four harmonic constraint rankings. With respect to ranking (2.3.3), listeners were able to identify place more accurately in syllable onset position than in syllable coda. This result reflected Ohala's (1990) earlier experimental results, and also seemed to verify his analysis of why regressive assimilations occur more often than progressive assimilations. Winters' (2001) results also confirmed Jun's hypothesis that perceptual salience would be greater for stops and nasals preceding a coronal stop than it is before labial and dorsal stops. This result suggested that there is a perceptual motivation for the fact that place assimilation occurs more often before labial and dorsal stops--as in, for example, Korean.

The rest of the results of Winters' (2001) study indicated that the perceptual motivation for other patterns in place assimilation was less clear, however. For post-vocalic consonants, listeners identified place best in unreleased labials, and worst in dorsal consonants, with coronals somewhere in between. This result indicated that ranking (2.3.1) was an incorrect representation of the perceptual facts. Similarly, listeners showed no significant differences in their ability to perceive place in nasals and stops in the VCCV sequences. This result failed to confirm the constraint ranking in (2.3.2), and the perceptual motivation it provided for the phonological fact that nasals undergo place assimilation more often, cross-linguistically, than stops.

2.4. Refining the perceptual analysis of cross-linguistic patterns in place assimilation

Despite Jun's (1995) success in linking up comparative levels of perceptual salience with the cross-linguistic patterns in place assimilation in (2.3.3) and (2.3.4), Winters' (2001) results indicated that the similar cross-linguistic asymmetries in (2.3.1) and (2.3.2) could not have similar perceptual motivations. Evidence has also surfaced since Jun (1995) that the "harmonic" constraint ranking in (2.3.1) does not hold for all languages. For instance, Odden (1987) showed that, in Chukchi, only velar nasals undergo assimilation, and Marlett (1981) observed that, in Seri, only labials undergo assimilation. Hume & Tserdanelis (2003) also cited assimilatory processes in other languages which target all possible pairs of labial, coronal and dorsal places of articulation. Hume & Tserdanelis suggested that Optimality Theory might still be able to account for these unusual patterns of place assimilation, so long as preservation rankings for particular places of articulation were determined on a language-specific basis, rather than universally, as in Jun (1995). Along with experimental data from Winters (2001), this phonological evidence suggests that a universal, perception-based account of the likelihood of certain places of articulation to undergo place assimilation is no longer tenable.

No phonological evidence has yet surfaced, however, which contradicts Jun's ranking (2.3.2). The motivation behind the cross-linguistic asymmetry between nasals and stops as targets of place assimilation therefore remains an outstanding issue. Three studies which have investigated potential perceptual influences on place assimilation--Ohala (1990), Hura et al. (1992), and Winters (2001)--have all shown no significant differences in listeners' ability to perceive place in nasals vs. stops. Despite this apparent discrepancy between the perceptual and phonological facts, however, several researchers have claimed that nasals undergo place assimilation more often than stops precisely because of their perceptual weakness.

For instance, Ohala & Ohala (1993) claimed that "Although nasal consonants as a class are highly distinct from other consonants, their place cues are less salient than those for comparable obstruents." Ohala & Ohala (1993) therefore suggested that nasals undergo assimilation more often than stops for similar reasons to those put forth in Ohala (1990) to explain the cross-linguistic asymmetry between regressive and progressive assimilation.

Similarly, Boersma (1998)--in describing the limitation of place assimilation to just /n/ (and not /t/) in Dutch—stated 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's statement put an interesting twist on the traditional, Kohlerian view of perception's influence on phonology. Boersma seems to have claimed that listeners make less effort to distinguish between sounds which are highly confusable, and do not therefore notice if speakers pronounce such sounds incorrectly. Laziness on the part of the listener in this case--along with the perceptual weakness of nasals--allows nasals to undergo place assimilation, even when stops do not.

Steriade (2001) also developed an innovative interpretation of Kohler (1990). Instead of portraying speakers as trying to get the "maximal perceptual effect with the minimal articulatory means," Steriade proposed that speakers attempt to get away with articulatory reductions whenever listeners let them, because of perceptual confusions. Steriade suggested, in other words, that articulatory changes to highly salient segments would not be allowed by a speech community, simply because they would be too noticeable. Whatever the interpretation of speaker's motivations, however, both Steriade's and Kohler's theories predict that reductions and assimilatory processes should target the perceptually weakest classes of segments. Steriade thus presumed that perception could account for the cross-linguistic tendency of nasals to undergo place assimilation:

> “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 nasals and plosives, despite Steriade’s (2001) claims to the contrary. Interestingly, this assumption also surfaces in Ohala & Ohala (1993), who cited studies such as Singh & Black (1966) and Wang & Fillmore (1961) as demonstrating the perceptual weakness of nasals. However, upon closer examination of the results of Singh & Black (1966)—a cross-linguistic study of consonant identification—reveals almost perfect identification in nasals in all conditions. Singh & Black themselves remarked that, “the two nasal sounds /m/ and /n/ were highly intelligible among all groups of speakers and listeners.” Wang & Fillmore’s (1961) results also showed no clear trend towards nasals being perceptually weaker than stops. Wang & Fillmore’s listeners correctly identified place in nasals correctly about as often as they did place in stops. 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.

Despite the misguided assumptions behind some of these perceptual claims, there are some studies which do seem to provide genuinely compelling evidence for the notion that place cues in nasals are perceptually weaker than the same cues in stops. Jun (1995), for instance, cited the results of Malecot (1956) and Mohr & Wang (1968) in developing the motivation for ranking (2.3.2). Jun claimed that Malecot (1956) had shown that, in post-vocalic position, nasal place cues were less salient than stop place cues. Malecot's (1956) study of place perception included a condition in which post-vocalic stop and nasal transitions were cross-spliced with nasal murmurs of conflicting place (e.g., /ɛ{d}/ transitions + /m/ murmurs—or /ɛ{n}/ transitions + /m/ murmurs) and then played to listeners. Malecot's results showed that, for the stop transition + nasal murmur combinations, the place information in the stop 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 stop transitions must provide stronger place cues than do post-vocalic nasal transitions. Jun also argued for the greater confusability of nasals by appealing to the results of Mohr & Wang (1968). Mohr & Wang (1968) ran 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 stops, or sounds of any other type. Boersma (1998) also supported his contention that "the nasals /m/ and /n/ are perceptually much more alike than the plosives /p/ and /t/" by citing the results of Pols (1983), a consonant identification experiment in which Dutch listeners had to identify consonants in VC syllables in various conditions of reverberation and noise. The consonant confusion matrices that this study yielded seem to show that listeners <u>were</u>

worse at identifying the nasals /m/ and /n/ than they were at identifying the stops /p/, /t/ and /k/, under these listening conditions.

Winters (2002) considered the possibility that the nature of the listening conditions in Pols' (1983) study might account for the fact that it yielded a perceptual difference between nasals and stops even though Ohala (1990), Hura et al. (1992) and Winters (2001) had not. Specifically, Winters (2002) hypothesized that noise might significantly diminish the perceptual strength of nasals, while leaving stop cues unperturbed. To test this hypothesis, Winters (2002) replicated the earlier, Winters (2001) study by playing VCCV stimuli to listeners and asking them to identify the place of articulation of the two, intervocalic consonants in a three-by-three forced-choice task. As before, the initial consonant in each intervocalic cluster was either a nasal or an unreleased stop, and both consonants varied between the labial, coronal and dorsal places of articulation. Listeners in Winters' (2002) study, however, heard these stimuli in one of four separate listening conditions: at a comfortable listening level, in noise at a +6 dB signal-to-noise ratio (SNR), at a -6 dB SNR, and at speech reception threshold.

The results of this study showed that--contrary to what was expected--listeners identified nasals <u>better</u> than stops in both the clear and the noisy listening conditions. The nasals' perceptual advantage over stops did disappear at speech reception threshold, however. Winters (2002) speculated that the ability of place cues in the nasal murmurs to resist the masking effects of noise could account for this pattern of results. The combination of transitions and these murmur cues might provide nasals with a perceptual advantage over stops--which had only transition cues to their place of articulation--in both the clear and noisy listening conditions. At the low volumes of the speech reception threshold condition, however, the already quiet nasal murmur cues apparently disappeared, and the perceptual advantage of stop transition cues over nasal transition cues (as found in Malecot 1956) re-emerged.

2.5. Motivation for the following study

While Winters (2002) helped clarify the conditions under which nasals might have stronger place cues than stops, it failed to answer the question of why some earlier studies had shown no perceptual differences between nasals and stops while others had shown that stops had stronger place cues than nasals. One possible explanation for such discrepant results is that the studies in question might have included different acoustic cues for stop place of articulation. It has long been known, for instance, that release bursts provide salient place cue information for stops (Winitz et al., 1971; Stevens & Blumstein, 1978). Studies such as Ohala (1990), Winters (2001) and Winters (2002) did not include release burst cues in their stop stimuli, however, since they operated on the assumption that stops which preceded other consonants in a cluster generally lacked release bursts. The stops in Pols' (1983) VC stimuli, however, may have had release bursts, which might have made them more perceptually distinct from one another than their nasal counterparts.

Pols' (1983) study also differed from the others mentioned above in that it tested Dutch listeners instead of only English listeners. The fact that Pols' (1983) listeners identified place in stops more accurately than in nasals may, therefore, reflect language-specific differences in the perception of place; Dutch listeners may be worse at perceiving place in nasals, while English listeners perceive place in nasals and stops equally well.

Such language-specific perceptual differences exist despite the fact that the consonantal inventories of both languages are quite similar:

	Bilabial	Labio-dental	Interdental	Alveolar	Post-alveolar	Velar	Glottal
Stops	p b			t d		k g	
Fricatives		f v	θ ð	s z	ʃ ʒ		h
Nasals	m			n		ŋ	
Affricates				tʃ dʒ			
Liquids				l	ɹ		
Glides	w				j		

Table 2.1: English Consonant Phoneme Inventory

	Bilabial	Labio-dental	Interdental	Alveolar	Post-alveolar	Velar	Uvular	Glottal
Stops	p b			t d		k		
Fricatives		f v		s z	ʃ ʒ	x		h
Nasals	m			n		ŋ		
Affricates					dʒ			
Liquids				l			ʀ	
Glides		ʋ			j			

Table 2.2: Dutch Consonant Phoneme Inventory (Booij 1995)

The potential perceptual differences between English and Dutch listeners seem to bear, however, an interesting relationship to the different process of place assimilation found in casual speech in those two languages. In English, both the coronal nasal (/n/) and the coronal stops (/t/ and /d/) may undergo place assimilation when preceding obstruents with different places of articulation (Gimson 1962, Brown 1977, Nolan 1992):

(2.5.1) “green paint” → [gɹi:m peɪnt]

(2.5.2) “red car” → [ɹɛg ka:ɹ]

(2.5.3) “bad thoughts” → [bæd̪ θɔ:ts]

As was noted above, research in studies such as Browman & Goldstein (1990), Barry (1991) and Nolan (1992) has revealed that this process may yield partial as well as complete assimilations, due to gestural overlap or gesture reduction, which are not representable with the broad phonetic transcriptions given above.

While similar investigations have not been undertaken into the process of casual speech place assimilation in Dutch, Booij (1995), Boersma (1998) and Van Oostendorp (2001) report that this process may target only /n/, before any other consonant:

(2.5.4) “Wie de baan krijgt” → [ba:ŋ kʀɛɪxt] (“Who the job gets”)

(2.5.5) “Wie de boon pakt” → [bo:m pakt] (“Who the bean takes”)

(2.5.6) “in Parijs” → [ɪm parɛis] (“In Paris”)

(2.5.7) “onwaar” → [ɔɱʋa:ʀ] (“untrue”)

The fact that place assimilation in Dutch may target only a nasal may reflect the reported perceptual fact that Dutch listeners perceive place in nasals worse than they perceive place in stops (Pols 1983) English, however, listeners perceive place equally well (or, poorly, as the case may be) in both stops and nasals, and both of these kinds of segments may undergo place assimilation in casual speech in the language. The language-specific perceptual abilities of Dutch and English listeners, that is, may be reflected in language-specific rankings of "preservation constraints" in both languages:

(2.5.8) Dutch: Pres(pl([stop]C)) >> Weakening >> Pres(pl[nasal]C))

(2.5.9) English: Weakening >> Pres(pl([stop]C)), Pres(pl[nasal]C))

While (2.5.8) and (2.5.9) are proposed in the spirit of Jun's (1995) treatment of place assimilation, this analysis differs from Jun's conception of preservation constraints, because the perception of place apparently differs between Dutch and English listeners, rather than resulting from universal perceptual influences across both languages.

The experiments described in the following two chapters test whether these two factors--the presence of audible release bursts, and the potential perceptual differences between Dutch and English listeners--can account for the discrepant results of previous studies. These experiments include both a replication of Mohr & Wang's (1968) study and an AX discrimination experiment (following Tserdanelis 2001 and Huang 2001). Each of these studies use both Dutch and English listeners as participants.

The decision to use an AX discrimination experiment--instead of an identification experiment, as in Hura et al. (1992), Pols (1983), or Winters (2001)-- to test for language-specific differences in place perception was made because of the lower-level perceptual nature of the AX discrimination task. In a perceptual identification task, listeners may simply attach linguistic/phonemic labels to the stimuli they hear. The "perceptual distance"--or confusability--between any two stimuli may then be calculated by deriving a statistical measure of listener "sensitivity" to the contrast between those stimuli. In an AX discrimination experiment, on the other hand, the listener's task is to listen to two consecutive stimuli and decide whether the second (X) is the same as or different from the first (A). This task does not necessarily require the listener to attach linguistic labels to the stimuli and may, therefore, reflect tendencies in the lower, sub-linguistic levels of perceptual processing. The persistence of language-specific differences between Dutch and English listeners in this task would therefore provide a stronger argument that the language of the listener has a genuine effect on their perceptual abilities.

The decision to replicate Mohr & Wang (1968), on the other hand, was made in order to investigate the hypothesis that perceptual presumptions influence phonology more than perceptual abilities. Mohr & Wang (1968) found that listeners consistently ranked contrasting nasals as more similar to each other than contrasting stops, or sounds of any other type. This result conflicts with other studies (Ohala 1990, Hura et al. 1992, Winters 2001), which showed no significant differences between listener perception of nasals and stops. The phonological fact that nasals undergo place assimilation more often than stops may, however, reflect speakers' presumptions about how perceptually confusable these sounds are, rather than listeners' ability to actually distinguish them from one another. While this hypothesis may seem far-fetched, it does concur with the results of Mielke (2002), who found that Turkish speakers' tendency to delete /h/ in certain phonological environments does not correspond directly to the perceptual salience of /h/ in those environments. The influence of "perception" on phonology may, therefore, depend on a higher-order interpretation of "perceptibility" by speakers. Comparing the results of a replication of Mohr & Wang (1968) and an AX discrimination experiment should make any differences between these levels of perceptual processing clear.

As mentioned earlier, studies such as Pols (1983) may have found significant differences between the perception of nasals and stops because their stop stimuli may have included audible release bursts, unlike the corresponding stimuli in Ohala (1990) or Winters (2001, 2002). The AX discrimination experiment and the Mohr & Wang replication will therefore include stop stimuli both with and without audible release bursts, in order to gauge what effect these cues have on their perceptual salience in comparison to nasals.

It is also important to consider how relevant these acoustic cues are to the study of perception's interaction with place assimilation. While Jun (1995) suggested that stops preceding other consonants generally lack release bursts, the findings of Henderson & Repp (1982) indicate otherwise. Henderson & Repp (1982) ran a small production study in which they determined whether or not stops had release bursts in consonant clusters by simply asking people to read such clusters, embedded in words, in English sentences. Henderson & Repp's (1982) results showed that a significant percentage of these stops did, indeed, have audible release bursts. The presence of these release bursts in consonant clusters might, therefore, provide stops with a perceptual advantage over nasals in an assimilatory environment. Shin (2000), for instance, has even suggested that languages--like Dutch--in which stops never undergo place assimilation, always produce stops with audible release bursts in consonant clusters. Chapter 5 therefore describes a replication of Henderson & Repp (1982), using both Dutch and English speakers as subjects, which was undertaken to investigate the hypothesis that perception might interact with production in this language-specific way.

The fact that several studies which investigated perceptual influences on place assimilation--Ohala 1990, Hura et al. 1992, and Winters 2001--all found no significant differences between listeners' ability to identify place in nasals and in stops suggests, however, that the phonological asymmetry between nasals and stops as targets of place assimilation may not, in fact, have any perceptual origins whatsoever. Instead, it seems plausible that constraints on articulation alone may motivate this phonological asymmetry. In the most general terms, place assimilation involves the transformation of an underlyingly heterorganic consonant cluster into an overtly homorganic consonant cluster. The fact that this change involves the elimination of a place gesture suggests that it may be induced by a general, articulatory principle of least effort (Zipf 1949, Lindblom 1983, Kirchner 1998). Lindblom (1983), for instance, points out:

> "An assimilation--whether phonological (a historical fact or a grammatically significant pronunciation rule), or phonetic (a grammatically nonsignificant attribute of an individual utterance)--invariably implies shortened movement (glottal or supraglottal). If once more we compare speech production to a second-order mechanical system...and examine the efficiency of such a system in terms of its energy expenditure, we see that assimilation, defined as reduced distance between two sequentially timed articulatory targets, implies less work per unit time. In a mechanical system such a restructuring of a frequently used sequence of targets will obviously, in the long run, lower energy cost."

Assimilations thus save speakers from expending unnecessary articulatory effort and thereby wasting precious biomechanical energy. Logically, the articulatory pressure to make such assimilatory changes should be even greater for heterorganic nasal-stop clusters than it is for heterorganic stop-stop clusters. Heterorganic nasal-stop clusters require the coordination of two place gestures with a concomitant manner gesture (i.e., the opening of the velo-pharyngeal port), and speakers thus have to expend more energy to articulate these clusters properly than they do in producing stop-stop clusters. Any computational model of articulatory work or effort (e.g., Kirchner 1998, Lindblom 1983) would, correspondingly, associate a higher articulatory cost to heterorganic nasal-stop clusters, thereby making them even more dispreferred (theoretically) by the speakers who have to produce them.

Other theories of articulation's influence on phonology exist which may also have some bearing on the tendency of nasals to undergo place assimilation. Boersma (1998), for instance, hypothesized that the phonologies of specific languages may also disprefer articulatory sequences which involve the coordination of two or more gestures. He formalized this notion in Optimality Theoretic terms with a family of "*Coord" constraints--

(2.5.10) *Coord (gesture$_1$, gesture$_2$): "The two gestures gesture$_1$ and gesture$_2$ are not coordinated."

and explained that, "This is an example of language-specific effort. Several muscles can only be pulled as a group (at least when speaking). These coordinations are language-specific and reflect the organizational shortcuts that characterize experienced speakers." Phonological constraints against gesture coordination may be relevant to the tendency of nasals to undergo place assimilation, since nasals involve the coordination of a place gesture and a velic opening gesture.

While Ohala (1990) criticized the notion of "articulatory ease" as "never having been satisfactorily defined," he also championed the idea that "innocent misapprehensions" on the part of the listener might influence phonological patterns and processes. Hume & Johnson (2001)--in their model of external influences on phonology-

-extend this notion of "innocent" mistakes to the domain of articulation, as well. For example, their interpretation of perception's influence on phonology followed Ohala's:

> “To understand how perception filters p [a phonological system], suppose that p requires the perception of a distinction that is somewhat hard to hear. In some instances, the difficult distinction required by p will be missed, simply misheard, so p will undergo a change to p’.”

Hume & Johnson conception of articulation's influence on phonology is similar:

> “The filtering action imposed by production takes a similar form. The cognitive symbolic representation p requires that the speaker make a sound that is hard to say. In some instances the speaker will fail to produce the sound and say something else and in this way contribute to a change in p.”

This model of articulation's influence on phonology shares with Ohala's model of listener-based sound change the attractive aspect of being testable, without a specific model of what is "easy to say" or "difficult to perceive." What matters for phonology are the types of articulatory and perceptual mistakes that the speakers and listeners of a language might make.

Chapter 6 therefore describes a repetition/imitation task, which was undertaken in the spirit of Hume & Johnson (2001) in order to test the hypothesis that the difficulty of repeating heterorganic nasal+stop clusters consistently and accurately might be what motivates them to undergo place assimilation in so many languages of the world. The results of this study--along with the findings from the AX discrimination experiment, and the replications of Mohr & Wang (1968) and Henderson & Repp (1982)--indicate that articulation is, in fact, the motivating force behind nasals' cross-linguistic susceptibility to place assimilation.

CHAPTER 3

PERCEPTUAL INFLUENCES ON PLACE ASSIMILATION: FREE MAGNITUDE ESTIMATION

3.1. Free Magnitude Estimation: Introduction and Background

Even though the results of Mohr & Wang (1968) have been cited as providing evidence for the relationship between perception and nasals' susceptibility to place assimilation, Mohr & Wang did not run their study out of an interest in this phonological process. They were, instead, trying to determine what relationship, if any, existed between listeners' subjective ratings of the similarity of sounds and the markedness of the features which distinguished them. Mohr & Wang (1968) had inherited this interest from Greenberg & Jenkins (1964), who had first applied the magnitude estimation technique to phonological research in an earlier study. In this study, Greenberg & Jenkins investigated subjective ratings of differences between voiced and voiceless stop consonants (/b/, /d/, /g/ and /p/, /t/, /k/). They tried several different rating scales and magnitude estimation tasks in order to determine which task yielded the most consistent results. On the basis of their results, they concluded that a free magnitude estimation task, with both audio and visual (i.e., written) presentation of the stimuli correlated most consistently with the results of all the different rating tasks. They therefore suggested that "...future tests with larger samples should concentrate on the technique employed in test A, the only test involving aural-visual stimuli."

Greenberg & Jenkins' study also revealed that distinctive features did have some relevance to the comparative values of estimated magnitudes. Listeners consistently ranked sounds differing in only one feature as much more similar to each other than sounds which differed in more than one feature. Their results also indicated that agreement in voicing had the greatest effect on the magnitude of estimated differences. On the basis of this result, Greenberg & Jenkins claimed, "This perhaps not unexpected result presumably reflects the greater psychological weight of marked as against unmarked features and, it might be conjectured, would hold at least as strongly for agreement in nasality as against agreement in non-nasality."

Mohr & Wang (1968) essentially designed their follow-up study to Greenberg & Jenkins (1964) in order to test this hypothesis. Mohr & Wang's methodology was quite similar to the earlier Greenberg & Jenkins experiment; however, they used only a six-point rating scale, rather than a free magnitude estimation task. They also gave listeners no visual presentation of the two sound stimuli being compared; instead, listeners simply made their judgments by marking their scores on a generic rating scale:

1	2	3	4	5	6
extremely similar	→				not similar

Figure 3.1: Mohr & Wang's (1968) Similarity rating scale

Mohr & Wang also included nasal consonants—along with voiced and voiceless stops—in one portion of their experiment. While the voiced and voiceless stops were presented to listeners in CV syllables, though, the nasals were presented in VC syllables, in order to accommodate /ŋ/, which cannot appear in syllable onset position in English. Mohr & Wang (1968) reported raw averages for the results of this experiment, which were the following:

	m	n	ŋ	b	d	g	p	t	k
m	--								
n	**1.79**	--							
ŋ	**1.86**	**1.91**	--						
b	3.63	4.16	4.27	--					
d	3.99	3.88	4.25	**2.54**	--				
g	4.43	4.22	3.54	**2.93**	**2.82**	--			
p	3.71	4.39	4.70	2.05	2.96	3.02	--		
t	4.36	4.36	4.25	3.27	2.38	3.36	**2.89**	--	
k	4.79	4.67	4.01	3.16	3.17	2.10	**3.08**	**2.95**	--

Table 3.1: Raw means of estimated similarities in Mohr & Wang (1968)

Average estimated differences between sounds which differed only in their specification for place of articulation appear underlined in boldface. These numbers indicate that listeners ranked contrasting nasals as more similar to each other than contrasting voiced and voiceless stops.

Mohr & Wang (1968) also analyzed this data in terms of the markedness of features distinguishing one sound from another. The following presentation of the data also seemed to indicate that nasality had more of an impact on estimated magnitude values than did voicing, or any other feature:

Feature	Similarity by Specification		Difference
	+	-	
labial	2.90	3.31	0.41
alveolar	3.05	3.26	0.21
velar	3.22	3.36	0.14
nasal	1.85	2.95	1.10
voiced	2.63	3.05	0.42

Table 3.2: Estimated similarities in marked and unmarked sounds, Mohr & Wang (1968)

These results also confirmed Greenberg & Jenkins' observation that the members of a category of marked sounds were consistently ranked as more similar to each other than the corresponding members of a category of unmarked sounds.

Ohala (1990)—later echoed by Jun (1995)—cited the results of Mohr & Wang (1968) as showing that, "Although nasal consonants as a class are highly distinct from other consonants, their place cues are less salient than those for comparable obstruents." Since Mohr & Wang (1968) did not analyze their data statistically, however, it is uncertain whether the results in Table 3.1 actually indicate a significant trend in perceptual distances between nasals and stops. The apparent difference in estimated similarities between the two groups of sounds is, nonetheless, striking.

However, this apparent difference conflicts with the results of Ohala (1990), Hura et al. (1992), and Winters (2001), all of which found no significant differences between listeners' perception of nasals and stops. One possible reason for this discrepancy may be the fact that, in Mohr & Wang's (1968) study, the stops were presented to listeners in syllable onset position while the nasals were in syllable codas. Studies such as Ohala

(1990), Fujimura et al. (1978) and Winters (2001) have demonstrated that pre-vocalic place cues are more salient than post-vocalic place cues. One reason for this asymmetry is that pre-vocalic stops always have audible release bursts, while these cues are often missing from post-vocalic stops. Ohala (1990), Hura et al. (1992) and Winters (2001), in fact, removed all release bursts from their stop stimuli because these stops were in consonant clusters, where (it has generally been assumed) stops lack audible release bursts in real-life speech productions. The fact that Mohr & Wang's stop stimuli had audible release bursts—and were in pre-vocalic position—may have caused their listeners to rate them as less subjectively similar than their nasal stimuli, which lacked release bursts and were in post-vocalic position.

While the following study is essentially a replication of the earlier Mohr & Wang (1968) and Greenberg & Jenkins (1964) studies, it modifies some of their methodologies in order to determine the extent to which differences in subjective magnitude estimations may have relevance to possible perceptual influences on place assimilation. The listeners in this replication therefore heard both nasals and stops in syllable coda position. The listeners also heard stops both with and without audible release bursts, in order to gauge the effect these cues have on listener estimations of differences between stops in comparison to differences between nasals. The extent to which release burst cues may be found in stops and nasals which are in a likely context for place assimilation to occur—i.e., preceding another consonant—will be determined by a subsequent production study, which is described in detail in chapter five. Similarly, another potential reason for the discrepancy between Mohr & Wang (1968) and the studies which found no significant differences in listener perception of stops and nasals—namely, the possibility that such differences only emerge at a higher level of perceptual processing—will be tested by comparing the results of this study with a subsequent AX discrimination experiment.

The task in this magnitude estimation experiment also differed slightly from that used in Mohr & Wang (1968). Instead of a rating scale, listeners estimated the differences between sounds using a free magnitude estimation task, as endorsed by Greenberg & Jenkins (1964). This task did not, however, include both an audio and a visual presentation of the stimuli, because there is no way to represent the presence or absence of audible release bursts in standard English orthography.

The working hypothesis for this study was that listeners would judge nasals to be significantly less different than stops with release bursts, but not significantly different than stops without these cues to place of articulation.

3.2. Magnitude Estimation Materials: Stimulus Construction

The stimuli for this experiment were constructed from items produced by two native speakers (male and female) each of English and Dutch. These original production items were recorded in a sound-attenuated booth, as spoken over a Shure SM10-A head-mounted microphone. This microphone was connected via an Symetrix SX202 Dual Mic Pre-amplifier (gain ≈ 50 dB) to a Sony DTC-790 Digital Audio Tape Deck, sitting outside the booth, with which the speakers' utterances were digitally recorded onto Sony PDP 65-C DAT cassettes.

The original production items for all speakers were VC and CV syllables, embedded in carrier phrases. The consonants in the VC syllables included both nasals and stops, of three different places of articulation (labial, coronal and dorsal). The vowels in both the VC and CV syllables were the three corner vowels, /i/, /u/ and /a/. The CV syllables only included stops of three different places of articulation, since the velar nasal, /ŋ/, is not found in syllable-onset position in either English or Dutch. Also, all stops in both kinds of syllables were voiceless, since Dutch lacks a voiced velar stop (/g/).

There were five unique carrier contexts for each syllable, as exemplified below:

English Script	Dutch Script
1a) Say OPP.	1a) Zeg AAP.
2a) Don’t say OPP, say OTT.	2a) Zeg niet AAP, zeg AAT.
2b) Don’t say OPP, say OCK.	2b) Zeg niet AAP, zeg AAK.
3a) Don’t say OTT, say OPP.	3a) Zeg niet AAT, zeg AAP.
3b) Don’t say OCK, say OPP.	3b) Zeg niet AAK, zeg AAP.

The production items were written pseudo-phonetically (as shown) in the orthography of each language. “Unique” contexts 2a and 2b—and 3a and 3b—only differed from each other in terms of the contrast between the two syllables in the carrier sentence. The entire scripts for both languages can be found in Appendices A and B.

The English speakers who produced these items were both phonetically trained linguists, while the Dutch speakers had neither phonetic training nor any background in linguistics. Both Dutch speakers were, however, fluent speakers of English. Both groups of speakers were recruited through e-mail contacts, and were paid $10 for their time and participation.

All digital recordings from each speaker were transferred to PC via a HOSA model ODL-276 Optical Data Link into a Creative SoundBlaster Live! Drive. This digital transfer was done on the PC with CoolEdit 2000 Software (© Syntrillium

Software Corporation) at a 44100 Hz sample rate, on a mono channel, with 16-bit resolution. Once on the PC, all CV and VC syllables were excised out of their original carrier contexts by using the sound editor in the Praat software package (version 4.0). The excised CV and VC syllables were saved individually in .aiff format for later use in a stimulus quality pre-test.

3.3 Stimulus Quality Pre-Test

All of the individual CV and VC syllables were presented to five trained phoneticians in a stimulus quality pre-test, in order to determine if they would be perceived as they had been intended to be uttered. This pre-test took the form of a perceptual identification task, implemented in a customized SuperCard (version 2.0) program running on a Macintosh. The listeners' task in this pre-test, therefore, was to listen to each individual CV or VC syllable and identify the place of articulation of the consonant in each syllable from among three pre-set choices (labial, coronal, or dorsal). In order for the listeners to be able to identify solely the place of articulation of the consonant in each stimulus they heard, all 540 stimuli were broken down into 9 separate groups according to syllable type (VC or CV), vowel type (/a/, /u/ or /i/) and manner of consonant (stop or nasal). The sixty separate stimuli in each block, then, varied among the four different speakers, the three different places of articulation for each consonant, and the five unique carrier contexts.

The listeners in this pre-test heard each item while sitting in a sound-attenuated booth. The items were played by the customized SuperCard program, running on a

Power Macintosh 7100/80, over Sony MDR-7502 Dynamic Stereo Headphones. The task itself was self-paced. After hearing each stimulus, the SuperCard program would present listeners with the on-screen question, “What did the speaker say?” The listeners could respond to this question by simply clicking on the appropriate response button on the screen. (see Figure 3.2 below) Three response options were given to the listeners in quasi-phonetic transcription (e.g., “ut,” “ta,” “ing”); these options only differed in the place of articulation of the consonant. The listeners also had the option of selecting a fourth response button for “none of the above.” In case any listener responded with this option, they were instructed to mark down the stimulus trial number and an alternative phonetic transcription for that stimulus on a separate response sheet. The listeners had the option of changing their responses before moving on to the next sound, but could only hear each individual stimulus once.

The participants in this pre-test had all undergone training in phonetic transcription. Their native languages included French, Cantonese, German, Greek and English. They were recruited from the population of the phonetics lab at Ohio State University, and were collectively compensated with a dozen mixed doughnuts for their participation in this pre-test.

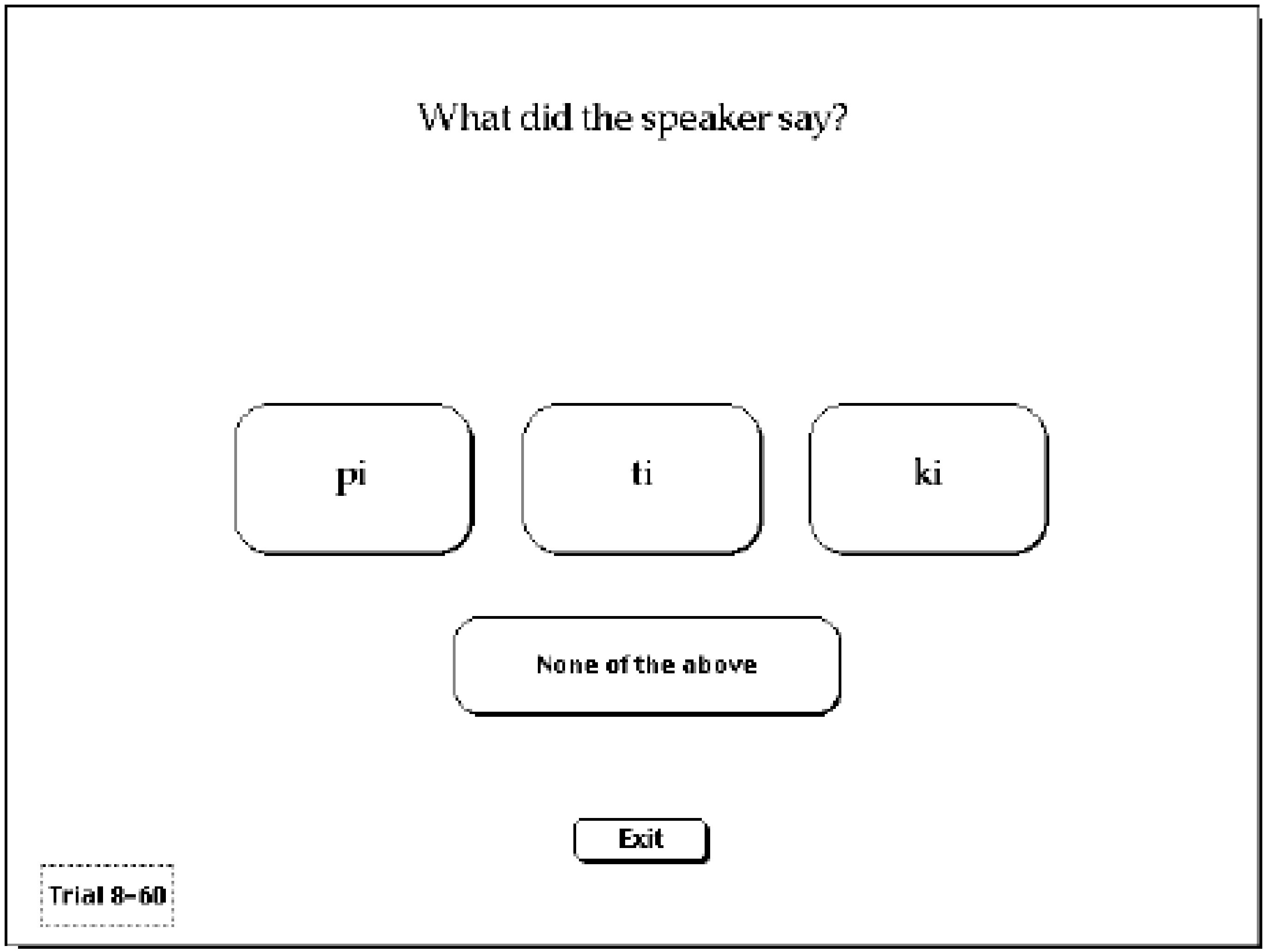


Figure 3.2: Example response screen for stimulus quality pre-test

The intention of the stimulus quality pre-test was to identify which of the original production items were not suitable for use in the subsequent magnitude estimation test. Suitable items were supposed to include only those original production items which had been correctly identified by all five listeners in the pre-test. However, there were certain syllable types for which there were no original productions which all five trained phoneticians identified correctly.

Tables 3.3-3.10 (see below) provide the tallies of raw responses for all of the original syllables in the pre-test. Each set of tables presents the responses to one speaker's original productions. These tables are broken down by individual syllable types. The five rows underneath each syllable type indicate the carrier context in which that syllable was uttered, while the four columns indicate the total number of responses--broken down by response option--that the five pre-test listeners gave for that particular production of the syllable. These response options include: L for labial, C for coronal, D for dorsal and N for "none of the above." The production which was ultimately chosen to represent each syllable type in the perceptual studies is indicated by the bolded carrier context number in each syllable production table. Ideally, these productions would only include the original utterances which had been identified correctly by all five listeners in the pre-test. In certain cases, however, productions which had been identified correctly by all five listeners could not be used in the perceptual experiment because of transient noises or other, distracting acoustic problems in the recording of that stimulus.

Tables 3.7 and 3.9 also reveal a more serious problem regarding the selection of appropriate stimuli for Dutch coronal nasals following /a/. In most cases, the pre-test listeners identified these stimuli as having a labial place of articulation. These results suggested that none of the recorded Dutch /an/ stimuli were suitable for use in the magnitude estimation task.

It had been noted during the recording sessions that the Dutch coronals had a dental place of articulation--rather than alveolar, as in English. In order to test the hypothesis that problems had arisen with these segments because of listener unfamiliarity with the dental place of articulation—rather than problems in the production of these syllables--the two Dutch speakers participated in the same stimulus quality pre-test as the original five, trained phoneticians. The Dutch speakers, however, only worked through the a-vowel, nasal-consonant, VC syllable block of this test, in order to gauge their ability to perceive the /an/ stimuli as they had been intended to be uttered.

The combined results of this follow-up study are given in Table 3.11. The results show that the Dutch speakers had near perfect identification of all of their own productions of the coronal nasal. On the basis of this result, then, these segments were included without reservation in the magnitude estimation study.

AP	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

AT	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

AK	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	1	0	4	0

IP	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

IT	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

IK	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

UP	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

UT	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

UK	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

AM	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

AN	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

ANG	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

IM	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

IN	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	4	1	0
3b	0	5	0	0

ING	L	C	D	N
1a	0	0	5	0
2a	0	2	3	0
2b	0	1	4	0
3a	0	0	5	0
3b	0	0	5	0

UM	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

UN	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

UNG	L	C	D	N
1a	0	0	5	0
2a	1	0	4	0
2b	2	0	3	0
3a	0	0	5	0
3b	0	0	5	0

Table 3.3: Pre-test identification responses for VC syllables, male English speaker

PA	L	C	D	N	TA	L	C	D	N	KA	L	C	D	N
1a	5	0	0	0	**1a**	0	5	0	0	1a	0	0	5	0
2a	5	0	0	0	2a	0	5	0	0	2a	0	0	5	0
2b	5	0	0	0	2b	0	5	0	0	2b	0	0	5	0
3a	5	0	0	0	3a	0	5	0	0	**3a**	0	0	5	0
3b	5	0	0	0	3b	0	5	0	0	3b	0	0	5	0

PI	L	C	D	N	TI	L	C	D	N	KI	L	C	D	N
1a	5	0	0	0	1a	0	5	0	0	1a	0	0	5	0
2a	4	0	1	0	2a	0	5	0	0	2a	0	0	5	0
2b	5	0	0	0	2b	0	5	0	0	2b	0	0	5	0
3a	5	0	0	0	3a	0	5	0	0	3a	0	0	5	0
3b	4	0	1	0	**3b**	0	5	0	0	**3b**	0	0	5	0

PU	L	C	D	N	TU	L	C	D	N	KU	L	C	D	N
1a	5	0	0	0	**1a**	0	5	0	0	**1a**	0	0	5	0
2a	5	0	0	0	2a	0	5	0	0	2a	0	0	5	0
2b	5	0	0	0	2b	0	5	0	0	2b	0	0	5	0
3a	5	0	0	0	3a	0	5	0	0	3a	0	0	5	0
3b	5	0	0	0	3b	0	5	0	0	3b	0	0	5	0

Table 3.4: Pre-test identification responses for CV syllables, male English speaker

AP	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

AT	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

AK	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

IP	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

IT	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

IK	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

UP	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

UT	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

UK	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

AM	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

AN	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	4	1	0
3a	0	5	0	0
3b	0	5	0	0

ANG	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

IM	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	2	2	1	0
3a	4	0	1	0
3b	4	1	0	0

IN	L	C	D	N
1a	0	4	1	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	4	1	0
3b	3	2	0	0

ING	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

UM	L	C	D	N
1a	1	0	4	0
2a	3	1	1	0
2b	4	0	1	0
3a	5	0	0	0
3b	5	0	0	0

UN	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	0	5	0

UNG	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

Table 3.5: Pre-test identification responses for VC syllables, female English speaker

PA	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

TA	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	4	1	0

KA	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

PI	L	C	D	N
1a	4	0	1	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

TI	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

KI	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

PU	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	3	0	2	0
3b	5	0	0	0

TU	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

KU	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

Table 3.6: Pre-test identification responses for CV syllables, female English speaker

AP	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

AT	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	1	4	0	0
3a	0	5	0	0
3b	0	5	0	0

AK	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

IP	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

IT	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

IK	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

UP	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	4	0	1	0
3b	5	0	0	0

UT	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

UK	L	C	D	N
1a	0	0	5	0
2a	4	0	1	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

AM	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

AN	L	C	D	N
1a	4	0	1	0
2a	4	0	1	0
2b	3	0	2	0
3a	5	0	0	0
3b	5	0	0	0

ANG	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	N

IM	L	C	D	N
1a	4	0	1	0
2a	5	0	0	0
2b	2	1	2	0
3a	5	0	0	0
3b	0	1	4	0

IN	L	C	D	N
1a	0	2	3	0
2a	0	0	5	0
2b	0	1	4	0
3a	0	4	1	0
3b	0	2	3	0

ING	L	C	D	N
1a	0	1	4	0
2a	0	0	5	0
2b	0	1	4	0
3a	0	1	4	0
3b	0	0	5	0

UM	L	C	D	N
1a	3	0	2	0
2a	5	0	0	0
2b	5	0	0	0
3a	4	0	1	0
3b	5	0	0	0

UN	L	C	D	N
1a	0	3	2	0
2a	0	2	3	0
2b	0	4	1	0
3a	5	0	0	0
3b	4	1	0	0

UNG	L	C	D	N
1a	5	0	0	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	3	0	2	0

Table 3.7: Pre-test identification responses for VC syllables, male Dutch speaker

PA	L	C	D	N	TA	L	C	D	N	KA	L	C	D	N
1a	5	0	0	0	**1a**	0	5	0	0	1a	0	0	5	0
2a	5	0	0	0	2a	0	5	0	0	2a	0	0	5	0
2b	5	0	0	0	2b	0	4	1	0	2b	0	0	5	0
3a	4	0	0	1	3a	0	4	0	1	**3a**	0	0	5	0
3b	5	0	0	0	3b	0	5	0	0	3b	0	0	5	0

PI	L	C	D	N	TI	L	C	D	N	KI	L	C	D	N
1a	5	0	0	0	**1a**	0	5	0	0	1a	0	0	5	0
2a	5	0	0	0	2a	0	5	0	0	2a	0	1	4	0
2b	5	0	0	0	2b	0	5	0	0	2b	0	0	5	0
3a	4	1	0	0	3a	0	5	0	0	**3a**	0	0	5	0
3b	4	0	0	1	3b	0	5	0	0	3b	0	0	5	0

PU	L	C	D	N	TU	L	C	D	N	KU	L	C	D	N
1a	5	0	0	0	**1a**	0	5	0	0	**1a**	0	0	5	0
2a	5	0	0	0	2a	0	5	0	0	2a	0	0	5	0
2b	5	0	0	0	2b	0	5	0	0	2b	0	0	5	0
3a	5	0	0	0	3a	0	5	0	0	3a	0	0	5	0
3b	5	0	0	0	3b	0	5	0	0	3b	0	0	5	0

Table 3.8: Pre-test identification responses for CV syllables, male Dutch speaker

AP	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

AT	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

AK	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

IP	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

IT	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

IK	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

UP	L	C	D	N
1a	5	0	0	0
2a	4	0	1	0
2b	4	0	1	0
3a	5	0	0	0
3b	5	0	0	0

UT	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

UK	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

AM	L	C	D	N
1a	5	0	0	0
2a	5	0	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

AN	L	C	D	N
1a	4	1	0	0
2a	4	1	0	0
2b	4	1	0	0
3a	4	1	0	0
3b	4	0	1	0

ANG	L	C	D	N
1a	0	0	5	0
2a	0	0	5	0
2b	0	0	5	0
3a	0	0	5	0
3b	0	0	5	0

IM	L	C	D	N
1a	5	0	0	0
2a	4	1	0	0
2b	5	0	0	0
3a	5	0	0	0
3b	5	0	0	0

IN	L	C	D	N
1a	4	1	0	0
2a	3	2	0	0
2b	5	0	0	0
3a	2	3	0	0
3b	0	3	2	0

ING	L	C	D	N
1a	2	1	2	0
2a	3	0	2	0
2b	1	1	3	0
3a	0	1	4	0
3b	0	1	4	0

UM	L	C	D	N
1a	5	0	0	0
2a	1	1	3	0
2b	2	0	3	0
3a	3	0	2	0
3b	5	0	0	0

UN	L	C	D	N
1a	0	5	0	0
2a	0	5	0	0
2b	0	5	0	0
3a	0	5	0	0
3b	0	5	0	0

UNG	L	C	D	N
1a	2	0	3	0
2a	0	1	4	0
2b	0	0	5	0
3a	0	1	4	0
3b	0	0	4	1

Table 3.9: Pre-test identification responses for VC syllables, female Dutch speaker

PA	L	C	D	N	TA	L	C	D	N	KA	L	C	D	N
1a	4	0	0	1	1a	1	4	0	0	**1a**	0	0	5	0
2a	5	0	0	0	2a	0	3	1	1	2a	0	0	4	1
2b	4	0	0	1	2b	0	4	1	0	2b	0	0	5	0
3a	5	0	0	0	3a	0	4	1	0	3a	0	0	4	1
3b	4	0	0	1	**3b**	1	4	0	0	3b	0	0	5	0

PI	L	C	D	N	TI	L	C	D	N	KI	L	C	D	N
1a	4	0	0	1	**1a**	0	5	0	0	1a	0	0	5	0
2a	4	0	0	1	2a	0	5	0	0	**2a**	0	0	5	0
2b	5	0	0	0	2b	0	5	0	0	2b	0	0	4	1
3a	4	0	0	1	3a	0	5	0	0	3a	0	0	5	0
3b	5	0	0	0	3b	0	5	0	0	3b	0	0	5	0

PU	L	C	D	N	TU	L	C	D	N	KU	L	C	D	N
1a	4	0	0	1	**1a**	0	5	0	0	1a	0	0	5	0
2a	4	0	0	1	2a	0	5	0	0	2a	0	0	5	0
2b	4	0	0	1	2b	0	5	0	0	2b	0	0	5	0
3a	4	0	0	1	3a	0	5	0	0	**3a**	0	0	5	0
3b	5	0	0	0	3b	0	5	0	0	3b	0	0	5	0

Table 3.10: Pre-test identification responses for CV syllables, female Dutch speaker

AM	L	C	D	N	AN	L	C	D	N	ANG	L	C	D	N
1a	2	0	0	0	1a	0	2	0	0	**1a**	0	0	2	0
2a	2	0	0	0	2a	0	1	1	0	2a	0	0	2	0
2b	2	0	0	0	**2b**	0	2	0	0	2b	0	0	2	0
3a	2	0	0	0	3a	2	0	0	0	3a	0	0	2	0
3b	2	0	0	0	3b	0	1	0	1	3b	0	0	2	0

Table 3.11a: Dutch Male Speaker

AM	L	C	D	N	AN	L	C	D	N	ANG	L	C	D	N
1a	1	0	0	1	1a	0	2	0	0	**1a**	0	0	2	0
2a	2	0	0	0	2a	0	2	0	0	2a	0	0	2	0
2b	2	0	0	0	**2b**	0	2	0	0	2b	0	0	2	0
3a	2	0	0	0	3a	0	2	0	0	3a	0	0	2	0
3b	2	0	0	0	3b	0	2	0	0	3b	0	0	2	0

Table 3.11b: Dutch Female Speaker

AM	L	C	D	N	AN	L	C	D	N	ANG	L	C	D	N
1a	2	0	0	0	**1a**	0	2	0	0	1a	0	0	2	0
2a	2	0	0	0	2a	0	2	0	0	2a	0	0	2	0
2b	2	0	0	0	2b	0	2	0	0	2b	0	0	2	0
3a	2	0	0	0	3a	0	2	0	0	**3a**	0	0	2	0
3b	2	0	0	0	3b	0	2	0	0	3b	0	0	2	0

Table 3.11c: English Male Speaker

AM	L	C	D	N	AN	L	C	D	N	ANG	L	C	D	N
1a	2	0	0	0	**1a**	0	2	0	0	**1a**	0	0	2	0
2a	2	0	0	0	2a	0	2	0	0	2a	0	0	2	0
2b	2	0	0	0	2b	0	2	0	0	2b	0	0	2	0
3a	2	0	0	0	3a	0	2	0	0	3a	0	0	2	0
3b	2	0	0	0	3b	0	2	0	0	3b	0	0	2	0

Table 3.11d: English Female Speaker

Table 3.11: Second pre-test identification responses for A-nasal syllables, by speaker

3.4 Magnitude Estimation Task: Stimulus Presentation

The listeners in the magnitude estimation task heard all stimuli in a sound-attenuated booth, as played to them over Sony MDR-7502 Dynamic Stereo headphones by a customized SuperCard (version 2.0) program running on a Macintosh PowerPC 6500/275 inside the booth. The SuperCard program was designed to play a pair of VC syllables—separated by a one-second inter-stimulus interval--to the listeners and then present them with the question, "How different were the two sounds?" Listeners responded to this question by simply keying in whatever number they felt was appropriate. In estimating the magnitudes of difference between sounds, the listeners were instructed to freely choose their own scale of comparison; however, the instructions gave them an example scale of 1 to 100. The complete instructions for this experiment—adapted from the instructions in Greenberg & Jenkins (1964)—are given below:

In this task, you will be listening to a series of syllables in order to give a subjective judgment as to how different the sounds in the syllables are. Specifically, the computer will play two syllables to you at a time. After hearing each pair of syllables, you will then be asked to give a number which represents how different you think the sounds in those two syllables are. For example, you might give a large number (such as 100) if you think the sounds are very different, or a small number (such as 1) if you think the sounds are very similar to each other. Click on the button to the right if you care for a brief demonstration.

In performing this task, you should use the judgment of difference that you give for the first pair as a basis for later judgments. That is, if the second pair of sounds sound twice as different to you than the first pair, you should give a number for the second pair that's twice as large as the number you gave for the first. During the experiment, you may worry that your scale is changing or that you have forgotten what you said before, and so on. Don't let that worry you; you can do this task better than you think you can. And you will get better and better as you go along. Once you are ready, click on the ***start*** *button. A speaker will then say the first pair of syllables. Listen to these syllables carefully, give a numerical estimate of how different you think they are, and then use this estimate as a basis for later judgments. Please contact the experimenter if you have any questions.*

Before making any official magnitude estimations in this experiment, listeners were given the option of running through a demonstration of the task, which they could repeat as many times as they wished. After making each magnitude estimation in the test itself, the listeners also had the option of revising their estimated score before moving on to the next pair of sounds, but they only heard each pair of sounds once.

The sounds the listeners heard consisted of VC syllables which had been selected for the experiment, based on the results of the stimulus quality pre-test (see Tables 3.3-3.11). The vowel in all stimuli was /a/--as it had been in all of Mohr & Wang's (1968) and Greenberg & Jenkins' (1964) stimuli. The consonants in the stimuli, on the other hand, varied between three places of articulation (labial, coronal, dorsal) and three manners of articulation (nasals, stops with bursts, and stops without bursts). There were thus nine basic stimuli for each of the four speakers. The listeners heard all possible <u>different</u> combinations of these stimuli, in both orders of presentation. These combinations included pairs which contrasted in both manner and place of articulation, as well as in each of these features individually. However, the speaker of both stimuli remained the same in all pairs. Each listener heard only two speakers in the experiment—one speaker of Dutch and one speaker of English. These speakers were also cross-matched by gender, so that every listener heard one male and one female voice. Presenting the speakers to the listeners in this way effectively split the entire study into two experiments—one with the Dutch male and the English female as speakers, and the other with the English male and the Dutch female as speakers.

Eight native Dutch-speaking listeners and eight native English-speaking listeners participated in this experiment. An equal number of listeners for both language groups participated in both halves of the experiment. All of the Dutch speakers were fluent speakers of English, having been recruited through e-mail contacts from the Columbus, Ohio Dutch Club and the student population at the Ohio State University. All Dutch speakers were paid $8 for participating in this experiment. All of the native English-speaking participants were recruited from introductory linguistics classes at the Ohio State University, and participated for course credit.

All of these listeners also participated in the burst production study (described in Chapter 5) immediately before participating in the magnitude estimation task.

3.5 Magnitude Estimation: Analysis

Since the listeners in this task were free to choose their own scales for making magnitude estimations, it would be inappropriate to simply compare raw scores or means of estimated magnitudes across listeners. Instead, a statistical analysis of the response data was only performed after all estimated magnitudes for each individual listener had been converted into a z-score of standard deviations away from the mean magnitude estimation score given by that particular listener. Calculating this measure essentially entailed determining the mean (M_i) score of estimated magnitudes for a particular listener, along with the standard deviation of all estimated magnitude scores for that same listener (SD_i). Each estimated magnitude score (EM_{ij}) given by that listener was then converted into a z-score (Z_{ij}) representing the number of standard deviations it was distant from the mean score for that listener:

$$(3.1)\quad Z_{ij} = \frac{EM_{ij} - M_i}{SD_i}$$

A repeated measures analysis of variance (ANOVA) was then run on the resultant z-scores in order to determine the extent to which the factors of manner of articulation (nasal, stop with burst, stop without burst), place contrast (labial-coronal, labial-dorsal, coronal-dorsal), language of speaker (English, Dutch) and language of listener (English, Dutch) influenced the magnitude estimation scores. In order to simplify this analysis of variance, the order of presentation of the stimuli in each pair was not considered as a factor. The score for each particular pair in the ANOVA was therefore an average of the z-scores for both orders of presentation.

Tables 3.12-3.15 show the z-scores by pair, broken down by language of speaker and listener, but averaged over order of presentation. The z-scores for those pairs which differed only in place of articulation are bolded in these tables. The analysis of variance was run over the combination of data in these four tables, excluding those stimuli pairs which differed in both manner and place of articulation. Tables 3.16-3.17 provide a view of only those z-scores which were used in the ANOVA, and Table 3.18 lists the results of the ANOVA itself.

	AKb	ATb	APb	AKn	ATn	APn	ANG	AN	AM
AKb	---	**0.025**	**0.136**	-1.335	-0.305	0.022	0.609	0.732	0.736
ATb		---	**-0.288**	-0.186	-1.140	-0.511	0.900	0.259	0.608
APb			---	-0.057	-0.868	-1.558	0.621	0.766	0.783
AKn				---	**-1.195**	**-0.444**	0.465	0.289	0.282
ATn					---	**-1.143**	0.239	0.176	0.589
APn						---	0.590	0.371	0.195
ANG							---	**-0.290**	**-0.314**
AN								---	**-1.596**
AM									---

Table 3.12: Average z-scores for Dutch stimuli, given by English listeners
(AKb = /ak/ with burst; AKn = /ak/ without burst, etc.)

	AKb	ATb	APb	AKn	ATn	APn	ANG	AN	AM
AKb	---	**0.158**	**0.550**	-1.187	0.058	0.269	0.766	0.788	0.855
ATb		---	**-0.236**	-0.089	-0.771	0.448	0.700	0.724	0.661
APb			---	0.096	-1.020	-1.549	0.627	0.457	0.403
AKn				---	**-0.833**	**-0.890**	0.244	0.101	0.083
ATn					---	**-1.372**	0.184	0.405	-0.075
APn						---	0.259	0.202	-0.051
ANG							---	**-0.805**	**-0.399**
AN								---	**-1.691**
AM									---

Table 3.13: Average z-scores for Dutch stimuli, given by Dutch listeners
(AKb = /ak/ with burst; AKn = /ak/ without burst, etc.)

	AKb	ATb	APb	AKn	ATn	APn	ANG	AN	AM
AKb	---	**-0.086**	**0.244**	-1.534	0.019	0.171	0.728	0.982	0.503
ATb		---	**0.442**	-0.050	-1.563	-0.258	0.403	0.872	0.876
APb			---	-0.035	-0.121	-1.695	0.765	0.379	0.280
AKn				---	**-0.926**	**-0.498**	0.507	0.552	0.413
ATn					---	**-0.227**	0.696	0.591	0.253
APn						---	0.492	0.548	0.279
ANG							---	**-1.017**	**-0.861**
AN								---	**-0.613**
AM									---

Table 3.14: Average z-scores for English stimuli, given by English listeners (AKb = /ak/ with burst; AKn = /ak/ without burst, etc.)

Total	AKb	ATb	APb	AKn	ATn	APn	ANG	AN	AM
AKb	---	**0.324**	**0.468**	-1.565	-0.168	0.397	0.449	1.120	0.639
ATb		---	**0.266**	0.176	-1.148	0.428	1.129	0.829	1.033
APb			---	0.192	-0.027	-1.644	0.686	0.706	0.735
AKn				---	**-1.040**	**-0.755**	0.071	0.291	0.330
ATn					---	**-0.587**	0.203	0.219	0.411
APn						---	0.226	0.398	-0.090
ANG							---	**-1.030**	**-1.019**
AN								---	**-0.726**
AM									---

Table 3.15: Average z-scores for English stimuli, given by Dutch listeners (AKb = /ak/ with burst; AKn = /ak/ without burst, etc.)

Manner	Bursts			No Bursts			Nasals		
Place Contrast	Dor-Cor	Dor-Lab	Cor-Lab	Dor-Cor	Dor-Lab	Cor-Lab	Dor-Cor	Dor-Lab	Cor-Lab
English Listeners	-1.099	0.262	0.416	-0.914	-0.202	-0.357	-0.759	-0.295	0.726
	-0.342	0.756	-1.195	-0.342	-0.463	-2.171	-0.342	-0.220	-1.561
	0.103	-0.441	0.647	-0.986	-0.115	-0.659	0.974	0.321	-2.401
	0.847	0.261	-0.715	-0.324	-0.715	-0.793	-0.832	-0.520	-1.418
	0.148	-0.291	-0.272	-2.007	-1.263	-1.778	-0.157	-0.100	-1.778
	0.627	0.627	-0.289	-3.009	0.016	-0.901	0.627	0.933	-3.284
	0.028	0.028	-1.213	-0.504	0.028	-1.213	-1.213	-1.213	-1.390
	-0.115	-0.115	0.320	-1.476	-0.839	-1.274	-0.622	-1.419	-1.665
Dutch Listeners	-0.316	0.778	-0.762	-0.803	-0.843	-0.965	-0.843	-0.722	-0.965
	0.835	0.522	0.210	-0.884	-1.041	-2.228	-0.728	-1.041	-1.197
	0.297	0.079	0.297	-0.576	-1.013	-1.013	-1.231	-0.576	-1.886
	0.870	0.339	0.870	-0.616	-1.041	-1.359	-0.404	0.021	-2.845
	-0.616	1.004	-0.385	-0.847	-1.079	-0.709	-1.079	-0.616	-1.542
	-0.263	0.190	0.643	-1.395	-0.942	-1.169	-0.489	0.643	-1.622
	0.235	0.621	-1.691	-0.150	0.235	-2.461	-0.920	-0.150	-2.076
	0.220	0.865	-1.072	-1.395	-1.395	-1.072	-0.749	-0.749	-1.395

Table 3.16: Raw z-score values for Analysis of Variance, Dutch stimuli only

Manner	Bursts			No Bursts			Nasals		
Place Contrast	Dor-Cor	Dor-Lab	Cor-Lab	Dor-Cor	Dor-Lab	Cor-Lab	Dor-Cor	Dor-Lab	Cor-Lab
English Listeners	-1.378	-0.450	-0.048	-0.635	-0.512	0.014	-0.512	-0.233	-0.883
	0.390	0.756	-0.098	-0.707	-0.707	-0.585	-0.707	-0.951	0.756
	-0.224	0.212	1.083	-0.550	-0.333	0.538	-0.659	-1.203	-0.550
	-0.129	0.457	-0.129	-0.520	-0.324	-0.715	-0.754	-0.715	-0.793
	0.071	-0.062	0.529	-0.577	-1.187	0.110	-1.740	-1.454	-0.806
	0.627	0.627	0.322	-2.979	-0.595	-1.206	-0.595	0.016	-0.289
	-0.149	0.383	1.269	-1.035	-0.504	-0.149	-1.390	-0.858	-1.213
	0.103	0.030	0.610	-0.404	0.175	0.175	-1.781	-1.491	-1.129
Dutch Listeners	0.373	-0.235	-0.033	-0.884	-0.681	0.129	-0.884	-0.722	-0.762
	-0.103	0.522	0.835	-0.572	0.366	-1.353	-0.728	-0.572	-0.103
	0.297	1.170	-0.576	-0.358	-0.576	-0.358	-1.231	-1.231	-1.231
	0.763	0.445	0.551	-1.783	-1.677	-0.722	-1.147	-1.465	-0.192
	-0.385	-0.060	0.541	-1.079	-1.079	-0.616	-1.125	-0.616	-0.847
	0.870	0.417	0.417	-0.942	-1.169	-0.942	-0.036	-1.169	-0.942
	0.235	0.621	-0.150	-1.306	-0.150	0.235	-1.691	-1.306	-1.306
	0.543	0.865	0.543	-1.395	-1.072	-1.072	-1.395	-1.072	-0.426

Table 3.17: Raw z-score values for Analysis of Variance, English stimuli only

	F	df	Sig.
Speaker	11.215	1,14	**0.005**
Speaker * Listener	0.027	1,14	0.872
Manner	60.677	2,13	<**0.001**
Manner * Listener	1.710	2,13	0.219
Pair	5.532	2,13	**0.018**
Pair * Listener	0.654	2,13	0.536
Speaker * Manner	2.513	2,13	0.119
Speaker * Manner * Listener	0.388	2,13	0.686
Speaker * Pair	12.727	2,13	**0.001**
Speaker * Pair * Listener	0.179	2,13	0.838
Manner * Pair	1.588	4,11	0.246
Manner * Pair * Listener	0.854	4,11	0.521
Speaker * Manner * Pair	0.910	4,11	0.491
Speaker * Manner * Pair * Listener	3.699	4,11	**0.038**
Listener	0.816	1,14	0.382

Table 3.18: Analysis of Variance for results of Magnitude Estimation Test

(speaker = language of speaker, listener = language of speaker, manner = manner of articulation, pair = contrast in place of articulation)

The analysis of variance reveals that there are significant main effects for manner of articulation, language of speaker and place contrast. (Tables 3.19-3.23 show the results of post-hoc, two-tailed t-tests which were conducted to determine the direction of significance for all significant effects in the ANOVA.) Of particular interest here is the manner of articulation effect; the post-hoc tests show that the stops with bursts have significantly higher z-scores than both stops without bursts ($p < .001$) and nasals ($p < .001$). There is, however, no significant difference between the z-scores for nasals and stops without bursts ($p = .81$). This pattern of results corresponds exactly to what had been predicted, and seems to indicate that the presence or absence of auditory release bursts in stops can account for the discrepancy between the results of earlier studies such as Mohr & Wang (1968) and Ohala (1990), Hura et al. (1992), and Winters (2001).

Post-hoc t-tests for the place contrast factor also reveal a similar pattern of results to those found by Mohr & Wang (1968) and Greenberg & Jenkins (1964). The raw numbers indicate that the contrast between dorsal and labial places of articulation is largest (mean = -.315), across all factors, while dorsal-coronal is second (-.56), followed closely by coronal-labial (-.648). The same ranking of place contrasts was found in both of the earlier studies; however, the post-hoc t-tests for the results of this study indicate that only the differences between dorsal-labial and the other two place pairs are significant ($p = .014$ for dorsal-coronal and $p = .019$ for coronal-labial), while there is no significant difference between the dorsal-coronal and coronal-labial pairs ($p = .410$).

Manner		T-Tests	
	Mean Z	vs. No Burst	vs. Nasal
Burst	0.167	**< <u>.001</u>**	**< <u>.001</u>**
No Burst	-0.826	---	0.810
Nasal	-0.864	---	---

Speaker		
	Mean Z	T-Test
Dutch	-0.591	**<u>0.025</u>**
English	-0.425	---

Pair		T-Tests	
	Mean Z	vs. Dor-Lab	vs. Cor-Lab
Dor-Cor	-0.560	**<u>0.014</u>**	0.410
Dor-Lab	-0.315	---	**<u>0.019</u>**
Cor-Lab	-0.648	---	---

Speaker * Pair			T-Tests	
Speaker	Pair	Mean Z	vs. Dor-Lab	vs. Cor-Lab
Dutch	Dor-Cor	-0.462	0.095	**<u>0.035</u>**
	Dor-Lab	-0.172	---	**<u>0.003</u>**
	Cor-Lab	-1.058	---	---
English	Dor-Cor	-0.616	**<u>0.022</u>**	**<u>0.004</u>**
	Dor-Lab	-0.400	---	0.105
	Cor-Lab	-0.244	---	---

Pair * Speaker			
Pair	Speaker	Mean Z	T-Tests
Dor-Cor	Dutch	-0.462	0.175
	English	-0.616	
Dor-Lab	Dutch	-0.172	**<u>0.038</u>**
	English	-0.400	
Cor-Lab	Dutch	-1.058	**<u>0.003</u>**
	English	-0.244	

Table 3.19: Results of post-hoc, two-tailed t-tests for significant effects in ANOVA

These broad patterns, however, generalize over an interesting interaction between the language of the speaker of the stimuli and the place contrast in question. Figure 3.3 represents this interaction graphically. It shows that, for the English stimuli, the ranking of place pair distinctiveness differs from the pattern found in the earlier studies. For these stimuli, dorsal-coronal pairs had consistently smaller z-scores than both dorsal-labial (p = .022) and coronal-labial (p = .004) pairs. There was no significant difference between z-scores for coronal-labial and dorsal-labial pairs (p = .105).

The Dutch stimuli reflected a pattern of results more similar to those found in the earlier studies. For these stimuli, coronal-labial pairs consistently showed smaller z-scores than either their dorsal-coronal (p = .035) or dorsal-labial (p = .003) counterparts, while there was no significant difference between dorsal-coronal and dorsal-labial pairs (p = .095). The overall interaction essentially amounts to a difference between the place pair that listeners judge to be most similar in the two different languages: for the English stimuli, coronals and dorsals are most similar, while, in Dutch, coronals and labials are perceptually closest to one another. Small z-scores for the coronal-labial pairs in Dutch is perhaps not surprising, given the results of the stimulus quality pre-test, in which most of the Dutch /an/ stimuli were mis-categorized as /am/. In fact, the consistently small z-scores for these pairs accounts, in great part, for the broader language of speaker effect, in which the Dutch stimuli were given consistently lower z-scores than the English stimuli (p = .025). The small z-scores for the English coronal-dorsal pairs in this study, however, do not correspond to similar confusions in the stimulus quality pre-test.

In interpreting the comparative perceptual distances between place pairs in their own magnitude estimation study, Mohr & Wang (1968) suggested that they seemed to reveal a relationship between perception and production: "In all of these cases, perceptual distances reflect physiological distances, in that a class whose members are articulatorily closer is judged to be perceptually closer." "Physiological differences" seem to play a role in the results of this study, as well. The Dutch dental coronals are considerably closer, perceptually, to labials than the English alveolar coronals, while these alveolars are perceptually closer to velar consonants than the Dutch dentals. This shift in perceptual proximity corresponds to a shift in articulatory proximity between coronal consonants and their place of articulation counterparts in the two languages. Alveolars are articulatorily closer to velars than dentals, while dentals are articulatorily closer to labials.

This observation cannot, however, account for the fact that alveolars are not articulatorily closer to velars than to labials, despite the apparent perceptual proximity of these places of articulation in English. Thus, in both Greenberg & Jenkins (1964) and Mohr & Wang (1968)—in which the connection between articulatory and perceptual distances was first observed—listeners rated alveolars as being closer to labials than to velars. Why the opposite effect was observed in this study for English alveolars is unclear. One possible explanation is that all the stop stimuli in both Mohr & Wang (1968) and Greenberg & Jenkins (1964) were played to listeners in CV syllables, rather than VC syllables, as was the case in this experiment. The perceptual organization of stop place of articulation may, therefore, differ according to the position of the stop in the syllable. This may reflect differences in the acoustic cues for stop place of articulation in VC and CV syllables. Second formant transitions between vowel and consonant, for instance, are more similar between alveolars and velars than they are between alveolars and labials. For alveolars and velars, these transitions both rise into the stop closure, whereas, for labials, they fall. Listeners might associate more weight to these cues in estimating the magnitude of differences between VC syllables. Likewise, the release bursts of labials and alveolars are perceptually similar in that they are both diffuse, rather than compact, as in velars. These cues might play more of an important role in determining the magnitude of estimated differences in CV syllables, which always contain these cues for stop place of articulation.

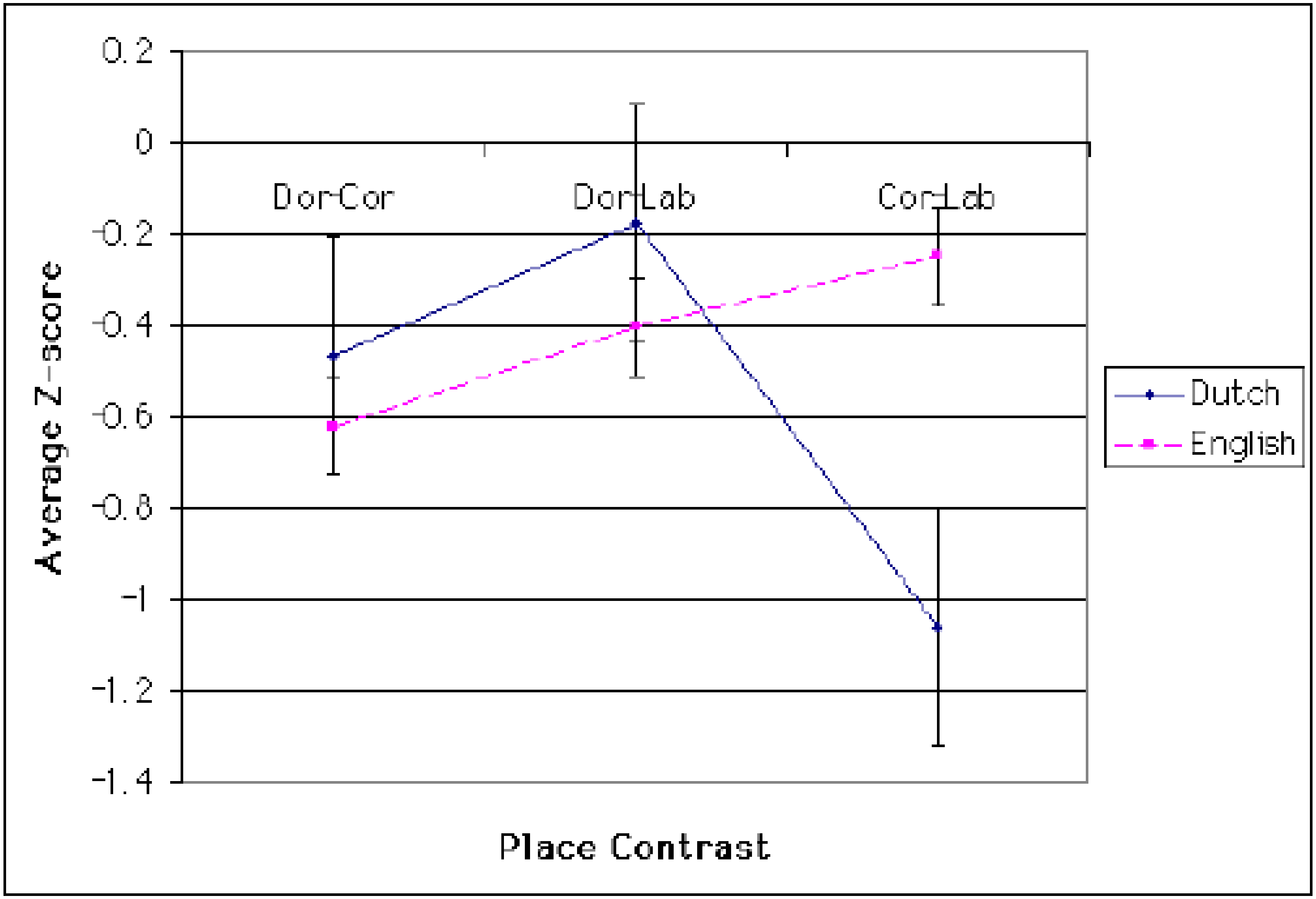


Figure 3.3: Average z-scores of estimated magnitudes, by place contrast and language of speaker

Tables 3.20-3.23 show the results of post-hoc t-tests for the complicated, but significant, four-way speaker * manner * pair * listener interaction in the Analysis of Variance. The results of these tests are listed primarily for reference, as they reveal little of significance in the data beyond what emerged in the considerations of the manner of articulation and language of speaker – place contrast effects. Some interesting interactions between the language of speaker and language of listener factors do, however, surface in these tests. These interactions become clearest, perhaps, in Table 3.23, which lists the results of t-tests for the data as broken down in the order manner of articulation, place contrast, language of listener, and language of speaker. For the stop stimuli, there are only three significant effects. The English listeners rated the Dutch coronal-labial pairs as significantly closer to each other than the English coronal-labial pairs, both with release bursts ($p = .044$) and without release bursts ($p = .013$). Likewise, the Dutch listeners rated Dutch coronal-labial stop pairs without bursts as more similar to each than the burstless English coronal-labial stop pairs ($p = .037$). This echoes the earlier finding that the Dutch dentals are very close, perceptually, to their labial counterparts.

In the nasal stimuli, however, an interesting cross-language effect emerges. Only the Dutch listeners rated the Dutch coronal-labial nasal pairs as more similar to each other than the English coronal-labial pairs ($p = .007$). Likewise, only the English listeners rated the English coronal-dorsal pairs as more similar to each other than the corresponding pairs in Dutch ($p = .030$). In both of these cases, the listeners rated the stimuli from their native language as more similar to each other than stimuli from a second or unfamiliar language. In one sense, this result is counterintuitive—one might expect listeners to be able to perceive the differences between two sounds in their native language better than they can perceive the differences between two sounds in another language. The Dutch listeners in the stimulus quality pre-test, for instance, were able to identify the place of articulation of Dutch /an/ syllables correctly, while trained phoneticians who did not speak Dutch natively could not. However, the listeners in this test were not simply trying to identify place of articulation; they were estimating how different or similar two given sounds were. Since these estimations were lower for native-language stimuli, these listeners apparently perceived a language similarity for stimuli from their native language which they did not perceive in non-native language stimuli. The fact that a pair of native-language stimuli might be less perceptually confusable to a listener while still seeming to be less different from a pair of non-native stimuli indicates that these two perceptual tasks—identification and magnitude estimation--are mutually independent of each other.

The following experiment will investigates the perceptual confusability of the stimuli that were used in this magnitude estimation task, and also addresses the issue of how confusability might relate to the cross-linguistic tendency of some sounds to undergo place assimilation.

Speaker*Manner*Pair*Listener

Speaker	Manner	Pair	Listener	Mean Z	T-Tests
Dutch	Burst	Dor-Cor	English	0.025	0.607
			Dutch	0.158	
		Dor-Lab	English	0.136	0.086
			Dutch	0.550	
		Cor-Lab	English	-0.288	0.902
			Dutch	-0.236	
	No Burst	Dor-Cor	English	-1.195	0.196
			Dutch	-0.833	
		Dor-Lab	English	-0.444	**0.025**
			Dutch	-0.890	
		Cor-Lab	English	-1.143	0.371
			Dutch	-1.372	
	Nasal	Dor-Cor	English	-0.290	0.137
			Dutch	-0.805	
		Dor-Lab	English	-0.314	0.755
			Dutch	-0.399	
		Cor-Lab	English	-1.596	0.816
			Dutch	-1.691	
English	Burst	Dor-Cor	English	-0.086	0.152
			Dutch	0.324	
		Dor-Lab	English	0.244	0.202
			Dutch	0.468	
		Cor-Lab	English	0.442	0.603
			Dutch	0.266	
	No Burst	Dor-Cor	English	-0.926	0.758
			Dutch	-1.040	
		Dor-Lab	English	-0.498	0.399
			Dutch	-0.755	
		Cor-Lab	English	-0.227	0.143
			Dutch	-0.587	
	Nasal	Dor-Cor	English	-1.017	0.943
			Dutch	-1.030	
		Dor-Lab	English	-0.861	0.532
			Dutch	-1.019	
		Cor-Lab	English	-0.613	0.603
			Dutch	-0.726	

Table 3.20: Results of post-hoc, two-tailed t-tests for significant four-way interaction in Analysis of Variance

Speaker*Manner*Listener*Pair

Speaker	Manner	Listener	Pair	Mean Z	T-Tests vs. Dor-Lab	vs. Cor-Lab
Dutch	Burst	English	Dor-Cor	0.025	0.682	0.428
			Dor-Lab	0.136		0.273
			Cor-Lab	-0.288		
		Dutch	Dor-Cor	0.158	0.175	0.253
			Dor-Lab	0.550		0.092
			Cor-Lab	-0.236		
	No Burst	English	Dor-Cor	-1.195	0.076	0.900
			Dor-Lab	-0.444		**0.009**
			Cor-Lab	-1.143		
		Dutch	Dor-Cor	-0.833	0.647	0.137
			Dor-Lab	-0.890		0.222
			Cor-Lab	-1.372		
	Nasal	English	Dor-Cor	-0.290	0.889	0.069
			Dor-Lab	-0.314		0.061
			Cor-Lab	-1.596		
		Dutch	Dor-Cor	-0.805	**0.041**	**0.010**
			Dor-Lab	-0.399		**0.007**
			Cor-Lab	-1.691		
English	Burst	English	Dor-Cor	-0.086	**0.040**	0.090
			Dor-Lab	0.244		0.437
			Cor-Lab	0.442		
		Dutch	Dor-Cor	0.324	0.474	0.810
			Dor-Lab	0.468		0.471
			Cor-Lab	0.266		
	No Burst	English	Dor-Cor	-0.926	0.207	**0.013**
			Dor-Lab	-0.498		0.264
			Cor-Lab	-0.227		
		Dutch	Dor-Cor	-1.040	0.158	0.127
			Dor-Lab	-0.755		0.584
			Cor-Lab	-0.587		
	Nasal	English	Dor-Cor	-1.017	0.295	0.093
			Dor-Lab	-0.861		0.387
			Cor-Lab	-0.613		
		Dutch	Dor-Cor	-1.030	0.956	0.199
			Dor-Lab	-1.019		0.135
			Cor-Lab	-0.726		

Table 3.21: Results of post-hoc, two-tailed t-tests for significant four-way interaction in Analysis of Variance

Speaker*Pair*Listener*Manner					T-Tests	
Speaker	Pair	Listener	Manner	Mean Z	vs. No Burst	vs. Nasal
Dutch	Dor-Cor	English	Burst	0.025	**0.027**	0.317
			No Burst	-1.195		0.129
			Nasal	-0.290		
		Dutch	Burst	0.158	**0.002**	**0.001**
			No Burst	-0.833		0.896
			Nasal	-0.805		
	Dor-Lab	English	Burst	0.136	**0.016**	0.144
			No Burst	-0.444		0.650
			Nasal	-0.314		
		Dutch	Burst	0.550	<**0.001**	**0.010**
			No Burst	-0.890		0.060
			Nasal	-0.399		
	Cor-Lab	English	Burst	-0.288	**0.005**	**0.023**
			No Burst	-1.143		0.300
			Nasal	-1.596		
		Dutch	Burst	-0.236	**0.012**	**0.012**
			No Burst	-1.372		0.291
			Nasal	-1.691		
English	Dor-Cor	English	Burst	-0.086	0.098	**0.020**
			No Burst	-0.926		0.827
			Nasal	-1.017		
		Dutch	Burst	0.324	**0.001**	<**0.001**
			No Burst	-1.040		0.960
			Nasal	-1.030		
	Dor-Lab	English	Burst	0.244	**0.007**	**0.002**
			No Burst	-0.498		0.181
			Nasal	-0.861		
		Dutch	Burst	0.468	**0.002**	**0.001**
			No Burst	-0.755		0.238
			Nasal	-1.019		
	Cor-Lab	English	Burst	0.442	**0.009**	**0.020**
			No Burst	-0.227		0.315
			Nasal	-0.613		
		Dutch	Burst	0.266	**0.042**	<**0.001**
			No Burst	-0.587		0.686
			Nasal	-0.726		

Table 3.22: Results of post-hoc, two-tailed t-tests for significant four-way interaction in Analysis of Variance

Manner*Pair*Listener*Speaker

Manner	Pair	Listener	Speaker	Mean Z	T-Tests
Burst	Dor-Cor	English	Dutch	0.025	0.541
			English	-0.086	
		Dutch	Dutch	0.158	0.464
			English	0.324	
	Dor-Lab	English	Dutch	0.136	0.461
			English	0.244	
		Dutch	Dutch	0.550	0.748
			English	0.468	
	Cor-Lab	English	Dutch	-0.288	**0.044**
			English	0.442	
		Dutch	Dutch	-0.236	0.158
			English	0.266	
No Burst	Dor-Cor	English	Dutch	-1.195	0.307
			English	-0.926	
		Dutch	Dutch	-0.833	0.384
			English	-1.040	
	Dor-Lab	English	Dutch	-0.444	0.784
			English	-0.498	
		Dutch	Dutch	-0.890	0.562
			English	-0.755	
	Cor-Lab	English	Dutch	-1.143	**0.013**
			English	-0.227	
		Dutch	Dutch	-1.372	**0.037**
			English	-0.587	
Nasal	Dor-Cor	English	Dutch	-0.290	**0.030**
			English	-1.017	
		Dutch	Dutch	-0.805	0.195
			English	-1.030	
	Dor-Lab	English	Dutch	-0.314	0.059
			English	-0.861	
		Dutch	Dutch	-0.399	0.065
			English	-1.019	
	Cor-Lab	English	Dutch	-1.596	0.093
			English	-0.613	
		Dutch	Dutch	-1.691	**0.007**
			English	-0.726	

Table 3.23: Results of post-hoc, two-tailed t-tests for significant four-way interaction in Analysis of Variance

CHAPTER 4

PERCEPTUAL INFLUENCES ON PLACE ASSIMILATION: AX DISCRIMINATION

Huang (2001) and Tserdanelis (2001) both used AX discrimination experiments to investigate the relationship between contrast perceptibility and the structure of a particular phonological process. Huang (2001), for instance, was interested in perceptual influences on a process of tone sandhi in Putonghua Chinese. In this process, a sequence of underlying 214-214 tones changes into a sequence of 35-214 tones on the surface. Huang hypothesized that the first tone in the sequence changed to 35—and not to one of the other two tones in the tone inventory—because the 35 tone was the most similar, perceptually, to the 214 tone. This hypothesis followed the line of thinking in Steriade (2001), who suggested that the output of phonological changes (e.g., 35-214) should be maximally similar, perceptually, to the input of the process (e.g., 214-214) in order to evade notice by the listening community, which is resistant to change.

Huang (2001) tested this hypothesis by testing the ability of both Chinese and English listeners to distinguish between tone sequences in an AX discrimination task. Listeners in an AX discrimination task hear a sequence of two sound stimuli and decide whether or not the second stimulus (X) was identical to or different from the first (A). Listeners make this decision by pressing a button marked either "same" or "different" on a button box, which can register the amount of time that passes in between the presentation of the stimuli and the listeners' registration of a response. This interval—commonly referred to as "reaction time" or "RT"—is generally assumed to have an inverse relationship with the perceptual distance between the A and the X stimuli. Smaller perceptual distances between stimuli, that is, makes them more confusable, which, in turn, leads to longer reaction times from listeners who are trying to determine whether these stimuli are the same or different.

In Huang's (2001) experiment, the pairs of sound stimuli were sequences of tone pairs: 55-315, 51-214, etc. Huang (2001) therefore hypothesized that the 35-214 and 214-35 tone sequences should induce longer reaction times (and more same/different judgment errors) than all other tone pairs, if perceptual similarity truly motivated the aforementioned tone sandhi process in Putonghua Chinese. Huang (2001) did, indeed, find that perceptual distance was smallest between 35 and 214 tones for both English and Chinese listeners, suggesting that perceptual similarity did, in fact, shape the output of Putonghua Chinese tone sandhi.

Tserdanelis (2001) took a similar approach to investigating potential perceptual influences on a process of manner dissimilation in Greek. This process affects both stop-stop and fricative-fricative sequences and by changing both such sequences into fricative-stop clusters. Underlying /pt/ sequences, for instance, change to [ft] on the surface, and underlying /fθ/ sequences also change to [ft] on the surface. Tserdanelis hypothesized that perceptual factors might account for why both of these processes yielded identical fricative-stop clusters on the surface. Specifically, Tserdanelis suggested that fricative-stop sequences maximize the possibility that listeners will perceive the place of articulation of both elements in the sequence correctly. In Tserdanelis' view, a phonological process would be motivated by perception into transforming certain sound patterns so that they would become easier to perceive or identify. While this analysis of perception's influence on phonology differs from that proposed by Kohler (1990) or Steriade (2001), it falls in line with the thinking of Liljencrants & Lindblom (1972)--or Hume (1998), who suggested that metathesis processes shift segments into positions where the perceptual cues to their identity become more salient.

Tserdanelis' test of his hypothesis therefore involved playing intervocalic consonant clusters, of different manners of articulation, to both Greek and English listeners in an AX discrimination task. Tserdanelis' sound pairs were matched by manner of articulation (e.g., stop-fricative, stop-stop, etc.) and only differed from each other in terms of the place of articulation of the individual consonants (e.g., /ft/ ~ /θt/, /pk/ ~ /pt/, etc.). Tserdanelis found that reaction times to different fricative-stop pairs in this study were faster (for Greek listeners) than reaction times to different stop-stop or fricative-fricative pairs. This result indicated that dissimilated, fricative-stop sequences were perceptually more distinct from one another than either stop-stop or fricative-fricative sequences. Dissimilation in Greek therefore changed such sequences into forms that were more distinct from one another, perceptually, for the listener. Tserdanelis also found that reaction times for different fricative-stop sequences were shorter than the reaction times for stop-fricative sequences. This result suggested that perception also influenced the direction of the dissimilatory change, so that it affected the first consonant in the sequence and thereby yielded a maximally discriminable consonant cluster.

Tserdanelis (2001) and Huang (2001) were therefore successful in relating the results of AX discrimination experiments to patterns in phonological processes, even though their individual interpretations of the relationship between perception and phonology differed. The following study also attempted to relate the results of an AX discrimination experiment to cross-linguistic patterns in the process of place assimilation. Its interpretation of the interaction between perception and phonology followed that which has been proposed by previous thinkers on this topic—e.g., Ohala (1990), Kohler (1990), Hura et al. (1992), Jun (1995)—all of whom claimed that segments with low perceptual salience or distinctiveness are the most likely to become the targets of place assimilation. This study therefore examined the distinctiveness of common targets of place assimilation—nasals and stops—in a likely environment for place assimilation to occur: directly before another stop, in an intervocalic consonant cluster. According to some theoreticians (Jun 1995, Boersma 1998, Shin 2000), nasals should have lower perceptual distinctiveness than stops in this environment; this would account for their comparatively greater cross-linguistic susceptibility to place assimilation. This experiment also tested the distinctiveness of stops both with and without audible release bursts in this environment, in order to determine whether the absence or presence of these acoustic cues can account for the discrepancies in the results of previous studies on the perception of place in both stops and nasals. Both Dutch and English listeners also participated in this experiment, in order to address the possibility that the ability to perceive place of articulation in nasals and stops might differ between these two groups of listeners.

4.1 Experimental Design

As in Tserdanelis (2001), listeners in this task only discriminated between sounds which had identical manners of articulation; each pair of A-X stimuli differed only in terms of the place of articulation of the two sounds. This experimental design therefore made it possible to test directly the hypothesis that place cues in nasals are less distinct than place cues in stops. As in Winters (2001, 2002), all target stimuli also took the form of vowel-consonant-consonant-vowel sequences. Although this experiment was primarily designed to test listeners' ability to discriminate between stops and nasals of varying places of articulation in the <u>first</u> consonant in these V-C-C-V sequences, it also required them to discriminate between an equivalent number of stimuli pairs which differed only in the place of articulation of the <u>second</u> consonant of the V-C-C-V sequence. The pairs of different stimuli were balanced in this way in order to prevent listeners from focusing solely on potential place differences in the <u>first</u> consonant in the intervocalic sequence. Likewise, listeners heard equal numbers of same and different pairs, in order to prevent them from being biased towards one response option over another.

One consequence of balancing the experiment in this way is that it required listeners to work through a very large number of stimuli pairs. The basic set of test stimuli, for instance, included only VCCV pairs which differed in terms of the place of articulation of the C_1 consonant. These consonants had three different manners of articulation: nasal, stop with an audible release burst, and stop without a release burst. These consonants were also at either the labial, coronal, or dorsal places of articulation;

however, the vowels in all stimuli were only /a/, in order to keep the number of stimuli in the experiment to a reasonable minimum. Nonetheless, the pairs of stimuli which met all of these design specifications alone included:

C1=Plosive	C1=Nasal
appa - appa	**ampa - ampa**
appa - atpa	ampa - anpa
appa - akpa	ampa - aŋpa
apta - apta	**amta - amta**
apta - atta	amta - anta
apta - akta	amta - aŋta
apka - apka	**amka - amka**
apka - atka	amka - anka
apka - akka	amka - aŋka
atpa - appa	anpa - ampa
atpa - atpa	**anpa - anpa**
atpa - akpa	anpa - aŋpa
atta - apta	anta - amta
atta - atta	**anta - anta**
atta - akta	anta - aŋta
atka - apka	anka - amka
atka - atka	**anka - anka**
atka - akka	anka - aŋka
akpa - appa	aŋpa - ampa
akpa - atpa	aŋpa - anpa
akpa - akpa	**aŋpa - aŋpa**
akta - apta	aŋta - amta
akta - atta	aŋta - anta
akta - akta	**aŋta - aŋta**
akka - apka	aŋka - amka
akka - atka	aŋka - anka
akka - akka	**aŋka - aŋka**

Table 4.1: VCCV pairs which differed only in the C_1 consonant, along with same pairs (in boldface)

For each of these pairs, there was a corresponding pair of stimuli which differed only in the place of articulation of the C_2 consonant. These C_2 consonants were voiceless stops with bursts, in all cases. These pairs of stimuli therefore included the following:

C1=Plosive	C1=Nasal
appa - appa	**ampa - ampa**
appa - apta	ampa - amta
appa - apka	ampa - amka
apta - appa	amta- ampa
apta - apta	**amta - amta**
apta - apka	amta - amka
apka - appa	amka - ampa
apka - apta	amka - amta
apka - apka	**amka - amka**
atpa - atpa	**anpa - anpa**
atpa - atta	anpa -anta
atpa - atka	anpa - anka
atta - atpa	anta - anpa
atta - atta	**anta - anta**
atta - atka	anta - anka
atka - atpa	anka - anpa
atka - atta	anka - anta
atka - atka	**anka - anka**
akpa - akpa	**aŋpa - aŋpa**
akpa - akta	aŋpa - aŋta
akpa - akka	aŋpa - aŋka
akta - akpa	aŋta - aŋpa
akta - akta	**aŋta - aŋta**
akta - akka	aŋta - aŋka
akka - akpa	aŋka - aŋpa
akka - akta	aŋka - aŋta
akka - akka	**aŋka - aŋka**

Table 4.2: VCCV pairs which differed only in the C_2 consonant, along with same pairs (in boldface)

Listeners also heard an equivalent number of same pairs for all of the different pairs listed in the two tables above. These pairs are also listed in Tables 4.1 and 4.2 in boldface; each of them needed to be presented to listeners twice in order to match the total number of different pairs in the experiment.

One of the more intriguing results of the magnitude estimation study was that the place cues for coronal consonants differed between English and Dutch, with Dutch's dental coronals being perceptually closer to labial consonants while English's alveolar coronals were perceptually closer to dorsals. This experiment therefore included a language of speaker factor, as well, whereby listeners heard all of the above stimuli pairs as spoken by both native-Dutch and native-English speakers.

The total number of stimuli pairs in this AX discrimination experiment therefore included the following:

3 places of articulation (labial, coronal, dorsal)
2 places which contrast with this place of articulation
(e.g., labial ~ coronal, labial ~ dorsal)
3 places of articulation for the other consonant in the cluster (labial, coronal, dorsal)
3 manners of articulation for the C_1 consonant (nasal, stop with burst, stop without burst)
2 positions in which consonants can contrast (C_1, C_2)
2 types of stimuli pairs (same, different)
2 languages of speakers (Dutch, English)
----
432 basic stimuli pairs

The listeners heard each of these stimuli pairs 6 times, for a total of 2592 repetitions in the AX discrimination task. In order to make this long listening task bearable, it was split into six separate listening sessions, each of which was approximately one hour long.

The results of previous studies (Ohala 1990, Hura et al. 1992, Winters 2001) indicated that there should be no significant differences between the comparative distinctiveness of nasals and stops in this experiment; however, the results of Pols (1983) suggested that Dutch listeners might be able to distinguish place in stops better than place in nasals. Given the results of the magnitude estimation experiment, it was also expected that all listeners would be able to distinguish between stops with bursts better than between both nasals and stops without bursts. The results of the magnitude estimation also laid the foundation for another perceptual expectation: that Dutch coronals would be perceptually closer to labials, while English coronals would be perceptually closer to dorsals. However, it was not assumed that the results of this study would agree perfectly with the results of the magnitude estimation task. Any differences which might emerge between the results of these two studies would be expected to provide an empirical glimpse into whatever differences exist between the corresponding levels of perceptual processing. The results were also, of course, expected to shed light on the relevance of the discriminability of nasal and stop sounds to the cross-linguistic patterns these sounds exhibit in place assimilation.

4.2 Methods: Stimulus Presentation

The stimuli for the AX discrimination experiment were constructed from individual VC and CV syllables which had passed the stimulus quality pre-test described in detail in Chapter 3. The exact same VC stimuli—for all three manners of articulation--therefore, that were used in the magnitude estimation task were also used in the AX discrimination study, in conjunction with some of the original CV productions from all four speakers. These VC and CV syllables were spliced together using the sound editor in the Praat software package (version 4.0). The timing of the intervocalic duration in the spliced-together VC-CV sequences was determined by the results of a consonant cluster production study, which was run before this experiment and is described in detail in Chapter 5. The following table lays out in detail the editing specifications for the duration of all acoustic cues for nasals and stops in both Dutch and English stimuli.

1. Burst Stimuli

1.1. Intervocalic duration
=140 milliseconds for Dutch
=120 milliseconds for English

1.2. Burst Beginning
=65 milliseconds for both languages

1.3. Burst duration
=25 milliseconds for Dutch
=15 milliseconds for English

1.4. Remainder
=50 milliseconds for Dutch
=40 milliseconds for English

2. Non-burst stimuli

2.1. Intervocalic duration
=110 milliseconds for Dutch
=105 milliseconds for English

3. Nasal stimuli

3.1. Intervocalic duration
=120 milliseconds for both languages

3.2. Nasal murmur duration
=65 milliseconds for Dutch
=90 milliseconds for English

Table 4.3: Timing specification for the construction of V-C-C-V stimuli

These durations were only specified according to the manner of articulation of the first consonant in the cluster, and the language of the speaker. Although intervocalic consonant cluster durations did vary according to the place of articulation of the first consonant in the cluster (see chapter 5), place of articulation was not included as a factor in determining the durations for the VC-CV stimuli in the AX discrimination task. Tuning consonant clusters to particular C_1 places of articulation would have made duration into a consistent-—but not necessarily natural—cue to the places of articulation that the listeners were trying to discriminate.

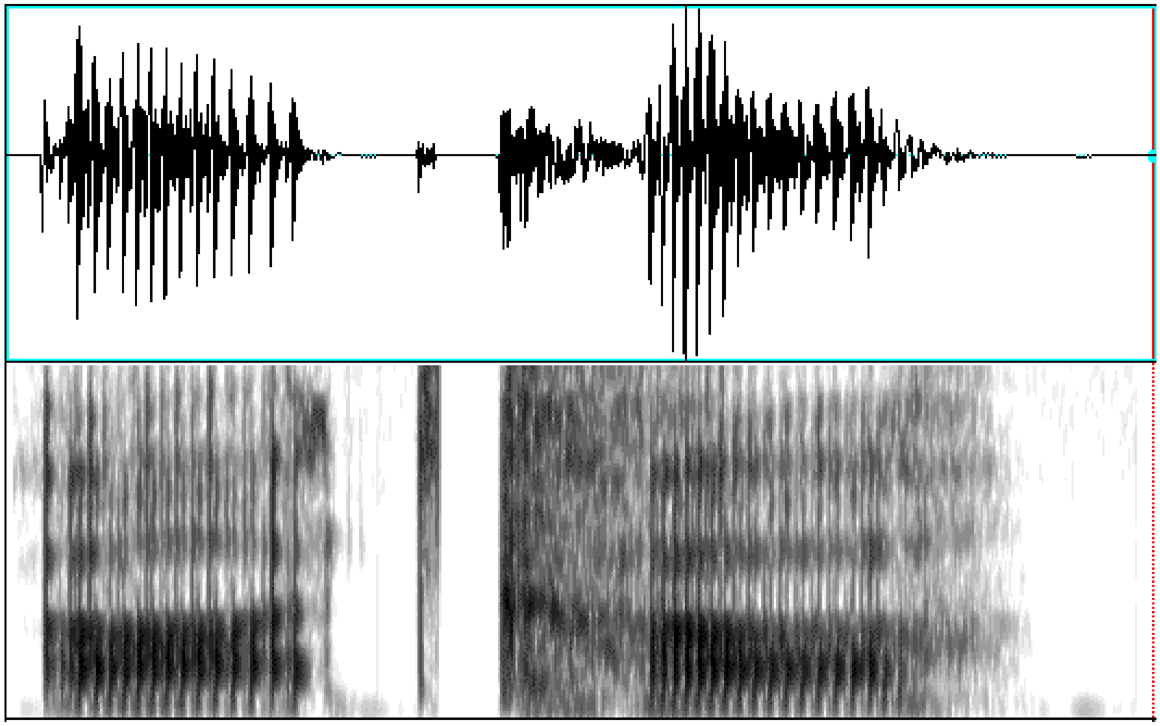


Figure 4.1: Example waveform and spectrogram for English male [atka] stimulus, with burst

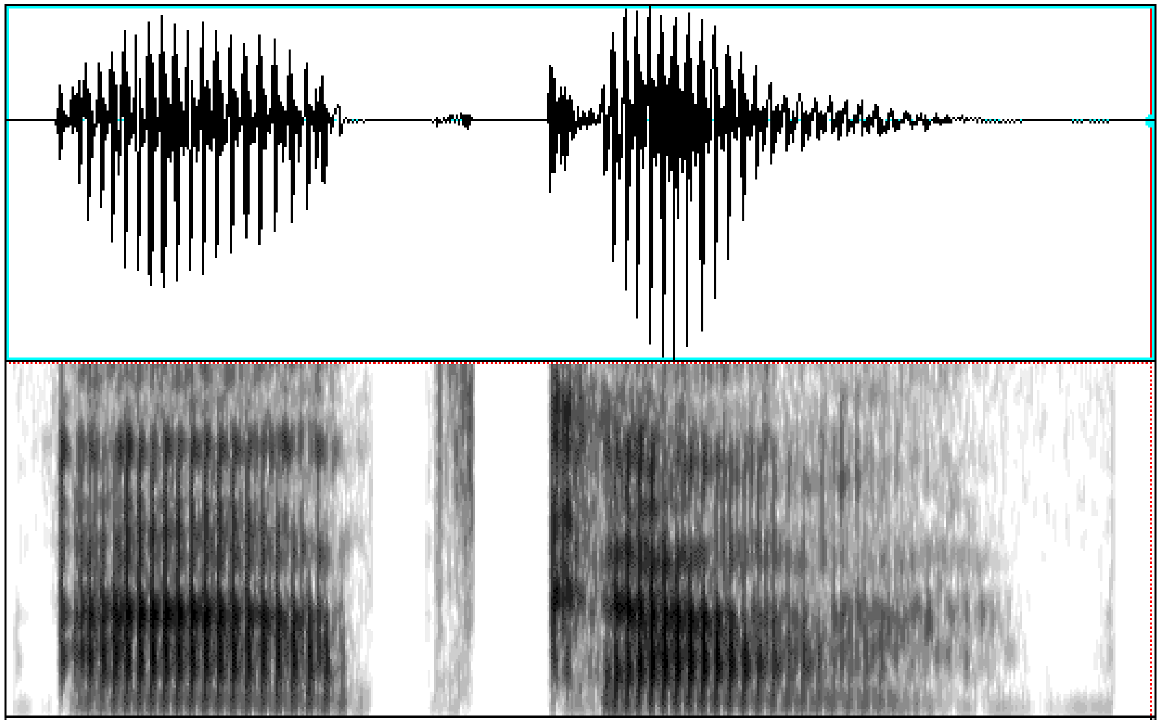


Figure 4.2: Example waveform and spectrogram for Dutch male [atka] stimulus, with burst

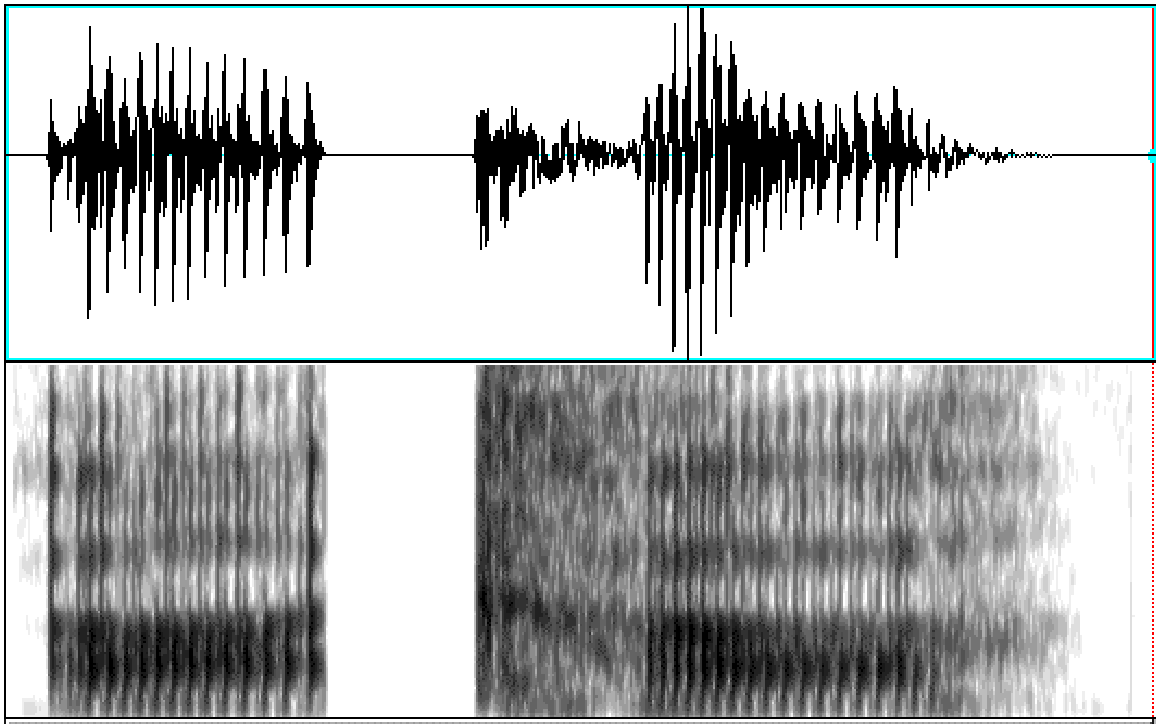


Figure 4.3: Example waveform and spectrogram for English male [atka] stimulus, without burst

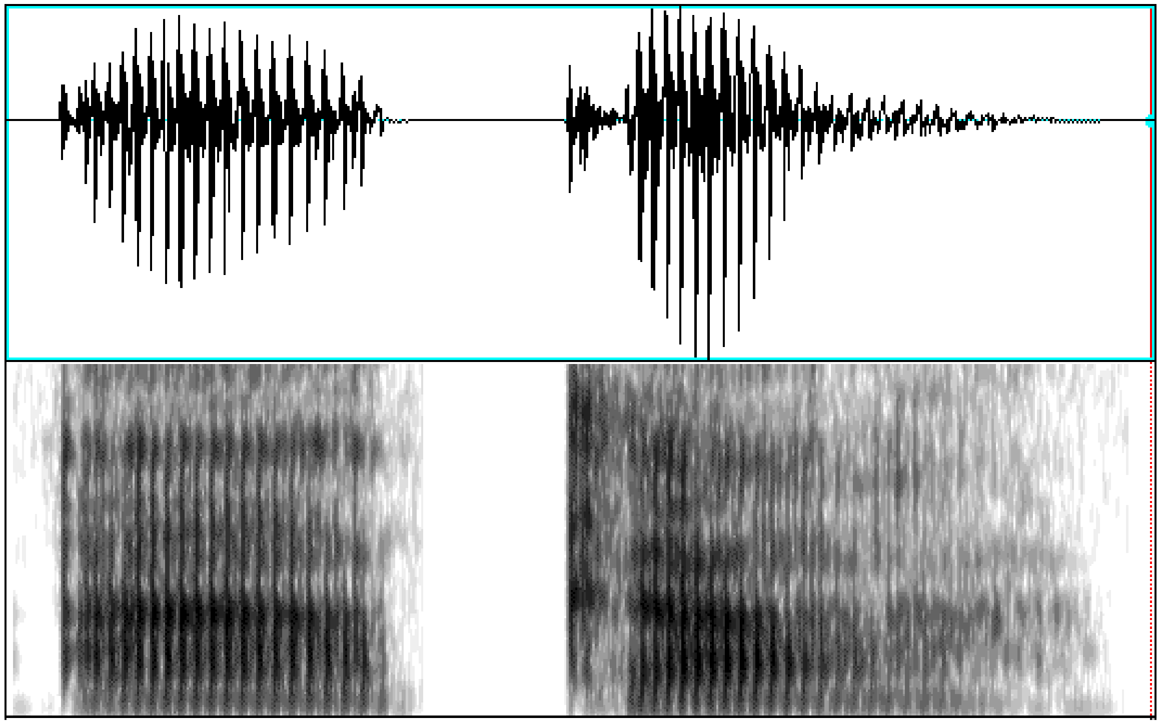


Figure 4.4: Example waveform and spectrogram for Dutch male [atka] stimulus, without burst

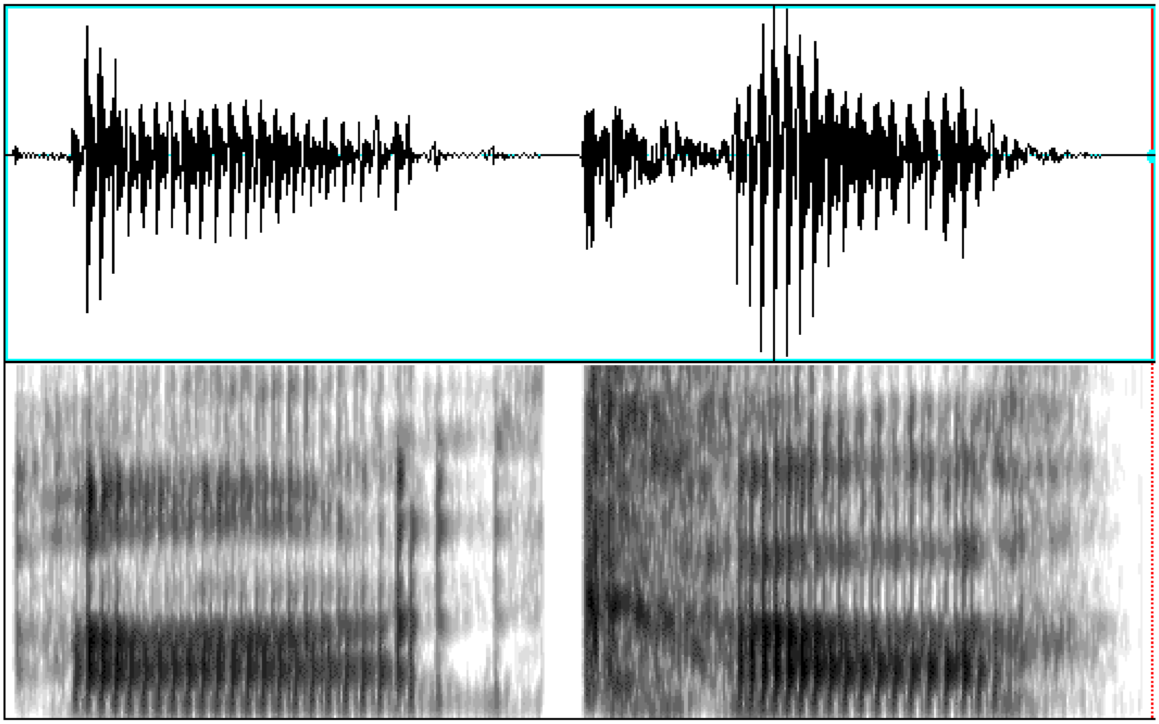


Figure 4.5: Example waveform and spectrogram for English male [anka] stimulus

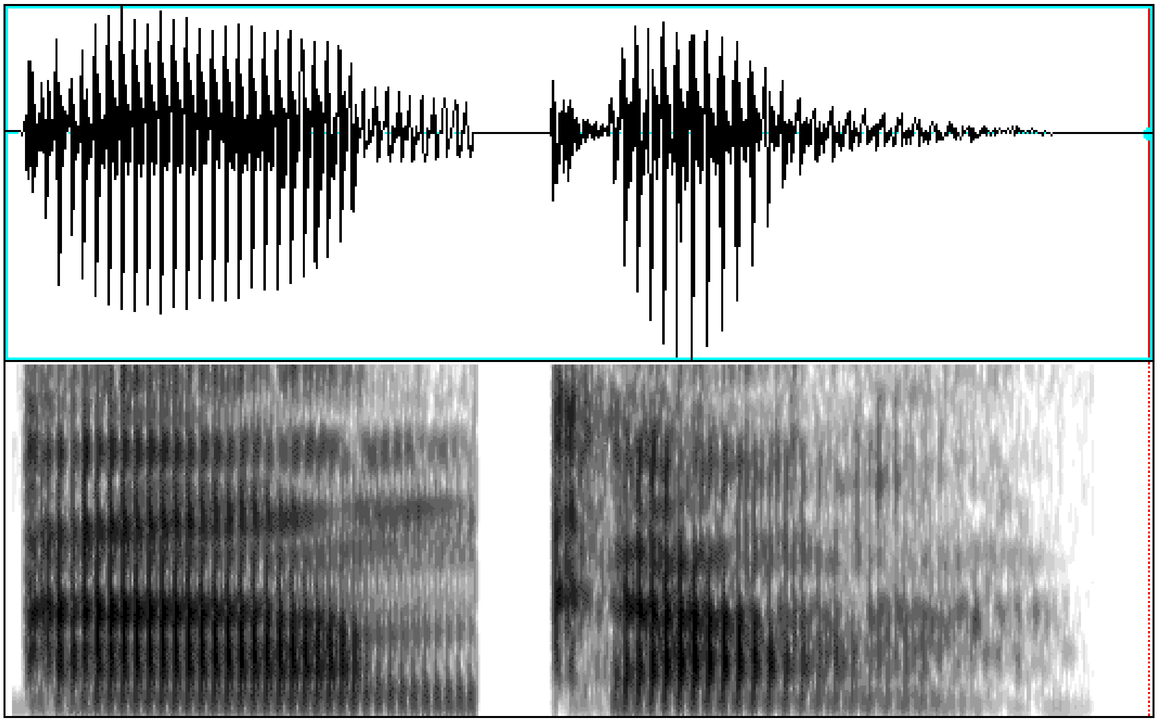


Figure 4.6: Example waveform and spectrogram for Dutch male [anka] stimulus

Great care was taken while splicing together VC and CV syllables to cut all waveforms at zero crossings, in order to avoid any artificial transients or spikes in the waveform. This approach became especially important when splicing off acoustic information at the end of stop release bursts or nasal murmurs in the original VC productions.

Once all such VC-CV sequences had been spliced together, they were down-sampled to a rate of 22050 Hz and saved as .wav files for later presentation in the AX discrimination test. The listeners heard all sound stimuli as played to them by a customized E-Prime (version 1.0) program, running on a PC in a quiet room. The native-English speaking participants heard all stimuli over Optimus Nova 71 headphones, while the Dutch speakers heard the same stimuli over Sennheiser HMD 25-l headphones. The customized E-prime program played pairs of VCCV stimuli to listeners with a 1000 millisecond interval between each stimulus. Listeners were instructed to decide whether the second of these stimuli was the same as or different from the first, as quickly and as accurately as possible. Listeners made this decision by pressing buttons marked either "same" or "different" on a Psychology Software Tools Serial Response Box Model #200A. The full text of the instructions to the listeners is given below:

THE PATTERNS IN PLACE PERCEPTION EXPERIMENT

This experiment is designed to test your ability to discriminate between some of the sounds of the world's languages.

The format of the experiment is simple: the computer will play two sequences of sounds to you in succession, and your job is to determine whether or not the second sequence of sounds was the SAME as the first, or DIFFERENT from the first sequence.

Press any key on the keyboard to continue.

Once you've decided whether the two sequences were the same or different, you should respond appropriately by pressing either the SAME or the DIFFERENT button on the *button box in front of you. You should try to do this as QUICKLY and as ACCURATELY as possible.*

Press any key on the keyboard, and the computer will run you through a short demonstration of this process.

The customized E-Prime software recorded the amount of time that elapsed between the onset of the second stimulus and the registration of a response button press by the listener. This amount of time was logged by the program as the “reaction time” for the response to each stimulus pair. The E-Prime software also recorded which button (i.e., “same” or “different”) the listener pressed, and whether or not this was the correct response for the stimulus pair. The program could thus keep a running total of the number of correct responses that the listener had registered. After each stimulus-response trial, the program presented the listener with both the reaction time for that trial and the percentage of correct responses for all trials. In this way, the program provided listeners with feedback on how quickly and how accurately they were responding to each stimulus pair.

In all, the listeners heard a total of 2592 VCCV-VCCV pairs. As in the magnitude estimation experiment, half of these pairs were spoken by a native Dutch speaker, while the other half were produced by a native English speaker. Since there were two speakers (male and female) for each language, they were cross-matched by gender for each listener. This effectively split the experiment into two sub-experiments, each with distinct sets of stimuli: one with the Dutch male and English female as speakers, and the other with the Dutch female and the English male as speakers.

All 2592 stimuli pairs—for both combinations of speakers—were independently randomized by a customized Perl script. Each independent randomization was then split into six, sequential parts, each of which became the set of stimuli for the six separate blocks of the AX discrimination experiment. The customized E-prime program then randomized each of these individual sets of stimuli once again at run time, before presenting them to the listener.

Ten native Dutch-speaking and ten native English-speaking listeners participated in this experiment; each group of listeners was evenly split between the Dutch Male-English Female and English Male-Dutch Female halves of the experiment. Dutch listeners were recruited from the student population at the Rijksuniversiteit Groningen, in the Netherlands; all of them had some ability to speak English, and many were fluent English speakers. All Dutch listeners were paid 50 Euros for their participation in the experiment. English listeners were recruited from the student population at the Ohio State University, in Columbus, Ohio. None of these listeners reported any previous experience with the Dutch language. All English listeners were paid $48 for participating in the experiment.

The entire experiment consisted of six separate blocks of 432 stimuli pairs each; all listeners were expected to work through each block in about an hour. Listeners were not expected to finish the entire experiment in one day; instead, they were given the option of working through two or three blocks per listening session, and thereby completed the entire experiment over the course of two or three days.

4.3 Analysis

Where this study differs from previous works such as Tserdanelis (2001) and Huang (2001) is in its interpretation of the relationship between reaction time in an AX discrimination task and "perceptual distance." Both Tserdanelis (2001) and Huang (2001) followed Shepherd (1978) in assuming that reaction time has an inversely proportional relationship to perceptual distance: the longer the reaction time, the shorter the perceptual distance between A and X stimuli, and the shorter the reaction time, the larger the perceptual distance. Both Tserdanelis (2001) and Huang (2001) also assumed that this inverse relationship only held for correct "different" responses to different A-X pairs. They therefore analyzed the mistakes in listeners' same/different responses independently of the corresponding reaction time measurements.

Podgorny & Garner (1979), however, suggested that mistakes and reaction time are not wholly independent in an AX discrimination task. Podgorny & Garner (1979) observed that pairs of different stimuli which induced longer reaction times also tended to induce comparatively larger proportions of incorrect "same" responses than did different pairs with relatively short reaction times. Podgorny & Garner developed this observation into a theory that a <u>directly</u> proportional relationship exists between reaction time and perceptual distance for stimuli pairs which are categorized as the "same" by listeners. In other words, long reaction times for "same" responses indicates a greater perceptual distance between the A and X stimuli than exists in "same" responses to which listeners have shorter reaction times. Podgorny & Garner were suggesting, in effect, that reaction time measurements revealed that there were different levels of "self-similarity" between stimuli which listeners perceived as being the same. The less “self-similar” two stimuli were, in other words, the more different they were—or, as the case may be, the greater the perceptual distance between them. For “same” responses, then, greater perceptual distance corresponded to less self-similarity, which induced longer reaction times. Conversely, smaller perceptual distances corresponded to greater self-similarity, which induced shorter reaction times. Contrary to what is generally assumed for “different” responses, then, the relationship between perceptual distance and reaction time was directly proportional for “same” responses.

One way of interpreting Podgorny & Garner’s (1979) theory is to consider the same/different discrimination task to be dependent on both the perceptual distance between the A and X stimuli and a same/different decision threshold. This threshold would be located somewhere along the perceptual distance continuum; it represents a hypothetical response strategy on the part of the listener to respond “different” to any stimulus pair whose perceptual distance exceeds that threshold, and “same” to any stimulus pair whose perceptual distance does not exceed the threshold. Figure 4.7 schematizes the likely effects this response strategy would have on reaction time:

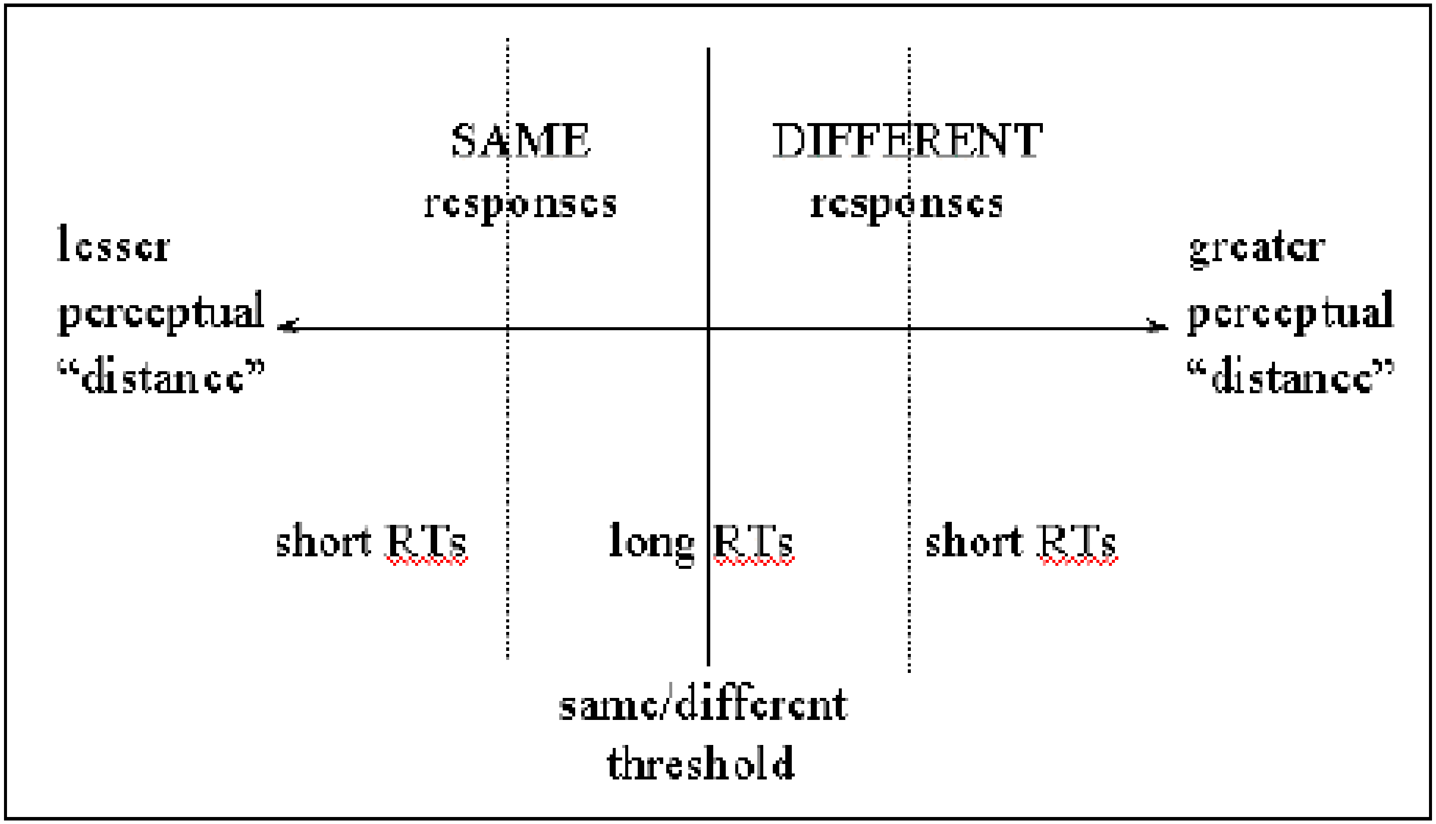


Figure 4.7: Analysis of response to different pairs in AX discrimination task

”Confusability” in making the same/different decision results from hearing two stimuli whose perceptual distance is close to the same/different decision threshold. Such stimuli pairs therefore inspire longer reaction times, regardless of whether they fall on the “same” or “different” side of the threshold. Theoretically, a perceptual distance which matched the decision threshold exactly would induce an infinitely long reaction time, as the listener would be unable to determine whether the perceptual distance between the A and the X stimuli fell on either side of the decision threshold. On the other hand, reaction times should become shorter for perceptual distances which are further away from the threshold, on either the “same” or “different” side.

The most important implication of this model is that it eliminates the need to analyze reaction times and categorization mistakes independently in an AX discrimination task. Instead, both “same” and “different” responses to <u>different</u> stimuli ought to be analyzed along the same continuum of perceptual distance. The appropriate measure to consider, in fact, is not raw perceptual distance but, rather, perceptual distance from the same/different decision threshold. This value is negative for all “same” responses and positive for all “different” responses; it also increases in magnitude (for both response types) as reaction times get smaller. Following Takane & Sargent (1983) and Nosofsky (1992), it will be assumed that reaction times increase logarithmically as they get closer and closer to the same/different decision threshold.

For any given pair of different stimuli, i and j, then, the appropriate equation for the perceptual distance between the pair, given a “different” response, is:

(4.1) $$PD_{ij} = \frac{1}{\ln(RT(D_{ij}))}$$ (where PD = perceptual distance, RT = reaction time)

In this model, perceptual distance is assumed to represent the perceptual distance from the same/different decision threshold. Since such perceptual distances always fall short of the same/different decision threshold for "same" responses, the corresponding equation for "same" responses simply requires the introduction of a negative sign before equation (4.1):

$$(4.2) \quad PD_{ij} = -\frac{1}{\ln(RT(S_{ij}))}$$

Listeners heard each different pair in the experiment a total of six times. Some of these presentations induced "same" responses from the listeners, while others induced "different" responses. Equation (4.1) was used to calculate perceptual distances individually for all "different" responses, while Equation (4.2) was used to make the corresponding calculations for all "same" responses. For each individual pair, i and j, all of the resultant "same" and "different" perceptual distances were then averaged across all six responses; Equation (4.3) shows the mathematical representation of this calculation:

$$(4.3) \quad PD_{ij} = \frac{\sum_{k=1}^{m} \frac{1}{\ln(RT(D_{ij})_k)} - \sum_{l=1}^{n} \frac{1}{\ln(RT(S_{ij})_l)}}{m+n}$$

($(D_{ij})_k$ represents the kth "different" response, while $(S_{ij})_l$ represents the lth "same" response)

Tables 4.4-4.15 present the results of this experiment in terms of this interpretation of perceptual distance. These perceptual distance values are based on responses to only those VCCV pairs which differed in the place of articulation of the C_1 consonant. Pairs which differed in the place of articulation of the C_2 consonant were, therefore, treated as filler items in the experiment. Tables 4.4 through 4.15 thus list perceptual distance values averaged across both orders of presentation for each pair. They are also averaged across all following C_2 place contexts. Each table does, however, break the perceptual distance means down by three factors: place contrast (coronal-labial, dorsal-labial, dorsal-coronal), language of listener (English, Dutch) and language of speaker (English, Dutch). These three tables differ from each other along a fourth dimension: manner of articulation (stops with bursts, stops without bursts, nasals).

			Same Responses			Different Responses			
Subj.	1	2	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	k	p	---	0	---	761.5	36	0.151	0.152
2	k	p	---	0	---	869.5	36	0.148	0.148
3	k	p	821.8	3	-0.149	695.1	33	0.153	0.128
4	k	p	---	0	---	894.1	36	0.147	0.148
5	k	p	---	0	---	747.6	36	0.151	0.152
6	k	p	---	0	---	1196.4	36	0.141	0.142
7	k	p	921.0	1	-0.147	816.0	35	0.149	0.142
8	k	p	837.0	2	-0.149	876.8	33	0.148	0.132
9	k	p	---	0	---	686.6	36	0.153	0.153
10	k	p	1594.0	1	-0.136	940.7	35	0.146	0.139
1	k	t	---	0	---	687.4	36	0.153	0.154
2	k	t	874.5	2	-0.148	1043.9	34	0.144	0.129
3	k	t	755.0	2	-0.151	695.5	34	0.153	0.137
4	k	t	1594.0	1	-0.136	931.7	35	0.146	0.139
5	k	t	---	0	---	792.6	36	0.150	0.150
6	k	t	1833.0	2	-0.133	1254.4	34	0.140	0.126
7	k	t	---	0	---	970.8	36	0.145	0.147
8	k	t	613.3	3	-0.156	852.0	33	0.148	0.124
9	k	t	1085.0	2	-0.143	669.3	34	0.154	0.138
10	k	t	1361.5	2	-0.139	987.8	34	0.145	0.130
1	t	p	564.0	1	-0.158	694.8	35	0.153	0.145
2	t	p	1168.5	28	-0.142	1176.1	8	0.141	-0.075
3	t	p	870.0	2	-0.148	621.6	34	0.155	0.139
4	t	p	1112.7	24	-0.143	1137.7	12	0.142	-0.043
5	t	p	1214.3	14	-0.141	1028.4	22	0.144	0.031
6	t	p	---	0	---	1155.1	36	0.142	0.143
7	t	p	1078.5	22	-0.143	1332.6	14	0.139	-0.027
8	t	p	3393.0	1	-0.123	794.7	35	0.150	0.143
9	t	p	---	0	---	623.9	36	0.155	0.156
10	t	p	927.4	12	-0.146	1033.8	24	0.144	0.048

Table 4.4: Average perceptual distance values, based on responses by Dutch listeners to stop-with-burst stimuli produced by Dutch speakers

			Same Responses			Different Responses			
Subj.	1	2	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	k	p	1109.0	1	-0.143	858.5	35	0.148	0.141
2	k	p	985.2	5	-0.145	713.6	31	0.152	0.111
3	k	p	---	0	---	756.1	36	0.151	0.151
4	k	p	1278.0	1	-0.140	740.3	35	0.151	0.144
5	k	p	---	0	---	885.4	36	0.147	0.148
6	k	p	1342.5	2	-0.139	816.5	34	0.149	0.135
7	k	p	---	0	---	713.1	36	0.152	0.153
8	k	p	1122.9	7	-0.142	950.4	29	0.146	0.090
9	k	p	---	0	---	894.1	36	0.147	0.148
10	k	p	1278.5	3	-0.140	1034.2	33	0.144	0.121
1	k	t	1625.5	4	-0.135	854.1	32	0.148	0.117
2	k	t	1052.3	3	-0.144	680.3	33	0.153	0.129
3	k	t	---	0	---	870.7	36	0.148	0.148
4	k	t	607.0	1	-0.156	626.1	35	0.155	0.147
5	k	t	---	0	---	924.3	36	0.146	0.147
6	k	t	1424.0	1	-0.138	860.9	35	0.148	0.142
7	k	t	---	0	---	746.4	36	0.151	0.152
8	k	t	---	0	---	890.3	36	0.147	0.149
9	k	t	---	0	---	932.9	36	0.146	0.147
10	k	t	1540.2	5	-0.136	1023.4	31	0.144	0.106
1	t	p	989.8	6	-0.145	934.8	29	0.146	0.097
2	t	p	---	0	---	653.4	36	0.154	0.155
3	t	p	1636.1	12	-0.135	1181.4	24	0.141	0.050
4	t	p	---	0	---	649.0	36	0.154	0.155
5	t	p	1702.8	4	-0.134	1058.1	32	0.144	0.112
6	t	p	---	0	---	846.9	36	0.148	0.150
7	t	p	864.9	16	-0.148	784.4	20	0.150	0.025
8	t	p	---	0	---	814.0	36	0.149	0.150
9	t	p	960.0	8	-0.146	1055.4	28	0.144	0.087
10	t	p	1223.0	3	-0.141	938.6	33	0.146	0.123

Table 4.5: Average perceptual distance values, based on responses by English listeners to stop-with-burst stimuli produced by Dutch speakers

			Same Responses			Different Responses			
Subj.	1	2	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	k	p	---	0	---	769.8	36	0.150	0.151
2	k	p	---	0	---	931.6	36	0.146	0.147
3	k	p	785.5	2	-0.150	766.5	34	0.151	0.134
4	k	p	1545.0	1	-0.136	884.6	35	0.147	0.140
5	k	p	---	0	---	818.7	36	0.149	0.150
6	k	p	1176.3	5	-0.141	1187.8	31	0.141	0.103
7	k	p	1554.0	2	-0.136	785.3	34	0.150	0.135
8	k	p	854.0	1	-0.148	943.5	35	0.146	0.139
9	k	p	---	0	---	685.5	36	0.153	0.153
10	k	p	---	0	---	964.4	36	0.146	0.146
1	k	t	1021.4	10	-0.144	884.9	26	0.147	0.068
2	k	t	1447.3	5	-0.137	985.9	31	0.145	0.106
3	k	t	873.8	13	-0.148	866.0	23	0.148	0.042
4	k	t	---	0	---	981.4	36	0.145	0.146
5	k	t	1679.0	3	-0.135	946.3	33	0.146	0.124
6	k	t	2035.5	13	-0.131	1418.4	23	0.138	0.041
7	k	t	1586.8	6	-0.136	1041.3	30	0.144	0.098
8	k	t	1097.1	7	-0.143	898.5	29	0.147	0.092
9	k	t	863.5	9	-0.148	812.1	27	0.149	0.076
10	k	t	1298.5	9	-0.139	1073.1	27	0.143	0.073
1	t	p	1111.4	5	-0.143	876.0	31	0.148	0.108
2	t	p	1083.0	1	-0.143	957.3	35	0.146	0.138
3	t	p	875.5	2	-0.148	791.1	34	0.150	0.134
4	t	p	1111.5	2	-0.143	1066.5	34	0.143	0.128
5	t	p	---	0	---	822.4	36	0.149	0.149
6	t	p	1073.0	1	-0.143	1202.7	35	0.141	0.134
7	t	p	1566.0	2	-0.136	900.4	34	0.147	0.133
8	t	p	---	0	---	847.9	36	0.148	0.149
9	t	p	1222.0	1	-0.141	700.8	35	0.153	0.145
10	t	p	1360.0	1	-0.139	935.0	35	0.146	0.139

Table 4.6: Average perceptual distance values, based on responses by Dutch listeners to stop-with-burst stimuli produced by English speakers

			Same Responses			Different Responses			
Subj.	1	2	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	k	p	---	0	---	878.8	36	0.148	0.149
2	k	p	863.3	6	-0.148	774.9	30	0.150	0.101
3	k	p	1454.0	1	-0.137	862.1	35	0.148	0.140
4	k	p	876.0	2	-0.148	727.1	34	0.152	0.136
5	k	p	---	0	---	943.8	36	0.146	0.147
6	k	p	3245.0	1	-0.124	922.1	35	0.146	0.141
7	k	p	---	0	---	715.9	36	0.152	0.153
8	k	p	1151.0	1	-0.142	897.0	35	0.147	0.140
9	k	p	---	0	---	845.6	36	0.148	0.149
10	k	p	---	0	---	1080.1	36	0.143	0.144
1	k	t	925.9	7	-0.146	847.8	29	0.148	0.091
2	k	t	978.3	8	-0.145	770.8	28	0.150	0.085
3	k	t	1410.8	12	-0.138	1089.0	24	0.143	0.049
4	k	t	929.7	9	-0.146	786.4	27	0.150	0.076
5	k	t	1323.8	3	-0.139	1086.2	33	0.143	0.120
6	k	t	1595.5	2	-0.136	947.2	34	0.146	0.132
7	k	t	1036.0	1	-0.144	754.0	35	0.151	0.143
8	k	t	1018.4	9	-0.144	974.9	27	0.145	0.074
9	k	t	---	0	---	877.0	36	0.148	0.148
10	k	t	1208.4	13	-0.141	1218.0	23	0.141	0.040
1	t	p	1066.3	3	-0.143	968.9	33	0.145	0.122
2	t	p	1017.3	5	-0.144	753.5	31	0.151	0.110
3	t	p	3224.0	1	-0.124	944.3	35	0.146	0.140
4	t	p	997.5	2	-0.145	772.1	34	0.150	0.134
5	t	p	---	0	---	957.9	36	0.146	0.146
6	t	p	---	0	---	913.2	36	0.147	0.149
7	t	p	1007.0	1	-0.145	756.6	35	0.151	0.143
8	t	p	1214.0	3	-0.141	977.7	33	0.145	0.122
9	t	p	---	0	---	957.8	36	0.146	0.147
10	t	p	1175.3	5	-0.141	1246.6	31	0.140	0.102

Table 4.7: Average perceptual distance values, based on responses by English listeners to stop-with-burst stimuli produced by English speakers

			Same Responses			Different Responses			
Subj.	1	2	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	k	p	793.5	3	-0.150	796.2	33	0.150	0.126
2	k	p	1726.3	3	-0.134	961.6	33	0.146	0.123
3	k	p	743.5	2	-0.151	742.3	34	0.151	0.135
4	k	p	---	0	---	937.9	36	0.146	0.147
5	k	p	---	0	---	831.8	36	0.149	0.149
6	k	p	1862.5	2	-0.133	1130.7	34	0.142	0.127
7	k	p	680.0	1	-0.153	968.9	35	0.145	0.138
8	k	p	825.8	3	-0.149	905.3	33	0.147	0.123
9	k	p	1298.0	1	-0.139	660.9	35	0.154	0.146
10	k	p	1434.0	11	-0.138	1001.9	25	0.145	0.059
1	k	t	888.5	2	-0.147	659.9	34	0.154	0.138
2	k	t	1133.5	6	-0.142	1053.6	30	0.144	0.097
3	k	t	982.0	2	-0.145	608.3	34	0.156	0.140
4	k	t	1477.0	1	-0.137	895.8	35	0.147	0.140
5	k	t	1132.8	3	-0.142	864.6	33	0.148	0.125
6	k	t	1382.0	1	-0.138	1172.5	35	0.142	0.134
7	k	t	1311.0	8	-0.139	1075.8	28	0.143	0.081
8	k	t	1515.0	1	-0.137	887.8	35	0.147	0.141
9	k	t	882.9	6	-0.147	690.1	30	0.153	0.104
10	k	t	1068.0	4	-0.143	1046.8	32	0.144	0.113
1	t	p	---	0	---	646.3	36	0.155	0.155
2	t	p	977.6	23	-0.145	1320.5	13	0.139	-0.043
3	t	p	725.5	7	-0.152	624.6	29	0.155	0.096
4	t	p	986.0	23	-0.145	1164.5	13	0.142	-0.042
5	t	p	1154.3	10	-0.142	1002.3	26	0.145	0.066
6	t	p	1655.6	7	-0.135	1426.9	29	0.138	0.085
7	t	p	1083.8	22	-0.143	1225.1	14	0.141	-0.033
8	t	p	1213.8	4	-0.141	995.9	32	0.145	0.114
9	t	p	892.5	3	-0.147	677.4	33	0.153	0.129
10	t	p	1126.7	21	-0.142	1291.6	15	0.140	-0.025

Table 4.8: Average perceptual distance values, based on responses by Dutch listeners to stop-without-burst stimuli produced by Dutch speakers

			Same Responses			Different Responses			
Subj.	1	2	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	k	p	1047.8	4	-0.144	809.4	32	0.149	0.117
2	k	p	1149.9	7	-0.142	708.0	29	0.152	0.095
3	k	p	---	0	---	772.5	36	0.150	0.151
4	k	p	747.5	2	-0.151	762.2	34	0.151	0.134
5	k	p	---	0	---	889.6	36	0.147	0.148
6	k	p	---	0	---	788.1	36	0.150	0.151
7	k	p	966.3	3	-0.145	705.1	33	0.152	0.128
8	k	p	1309.8	5	-0.139	986.8	31	0.145	0.107
9	k	p	986.2	7	-0.145	954.7	29	0.146	0.090
10	k	p	1174.5	8	-0.141	999.7	28	0.145	0.082
1	k	t	1061.8	10	-0.144	879.3	26	0.148	0.067
2	k	t	838.0	2	-0.149	670.6	34	0.154	0.138
3	k	t	---	0	---	807.5	36	0.149	0.150
4	k	t	---	0	---	665.1	36	0.154	0.155
5	k	t	1983.0	2	-0.132	940.8	34	0.146	0.131
6	k	t	---	0	---	875.5	36	0.148	0.149
7	k	t	853.0	2	-0.148	777.3	34	0.150	0.135
8	k	t	---	0	---	786.4	36	0.150	0.150
9	k	t	1141.5	4	-0.142	955.2	32	0.146	0.114
10	k	t	1232.5	6	-0.141	1010.4	30	0.145	0.097
1	t	p	892.4	14	-0.147	927.8	22	0.146	0.033
2	t	p	918.8	3	-0.147	619.2	33	0.156	0.131
3	t	p	1506.0	19	-0.137	813.4	17	0.149	-0.005
4	t	p	---	0	---	639.1	36	0.155	0.155
5	t	p	1339.6	12	-0.139	1229.5	24	0.141	0.049
6	t	p	607.0	1	-0.156	1008.8	35	0.145	0.138
7	t	p	850.5	26	-0.148	932.3	10	0.146	-0.066
8	t	p	1168.0	3	-0.142	808.4	32	0.149	0.125
9	t	p	955.8	22	-0.146	1251.7	14	0.140	-0.035
10	t	p	1140.0	4	-0.142	975.6	32	0.145	0.115

Table 4.9: Average perceptual distance values, based on responses by English listeners to stop-without-burst stimuli produced by Dutch speakers

			Same Responses			Different Responses			
Subj.	1	2	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	k	p	812.0	2	-0.149	845.2	34	0.148	0.132
2	k	p	1170.0	1	-0.142	928.3	35	0.146	0.139
3	k	p	967.7	6	-0.145	772.3	30	0.150	0.102
4	k	p	1609.0	1	-0.135	873.1	35	0.148	0.140
5	k	p	1754.0	1	-0.134	749.6	35	0.151	0.144
6	k	p	1627.3	3	-0.135	1222.6	33	0.141	0.118
7	k	p	---	0	---	829.3	36	0.149	0.150
8	k	p	1081.0	2	-0.143	1021.4	34	0.144	0.130
9	k	p	807.0	1	-0.149	730.5	35	0.152	0.144
10	k	p	1695.5	2	-0.134	911.5	34	0.147	0.132
1	k	t	1030.0	4	-0.144	831.9	32	0.149	0.117
2	k	t	986.6	6	-0.145	1268.3	30	0.140	0.094
3	k	t	834.4	12	-0.149	869.7	24	0.148	0.049
4	k	t	1101.0	1	-0.143	966.5	35	0.145	0.138
5	k	t	1090.6	7	-0.143	944.3	29	0.146	0.090
6	k	t	1410.8	11	-0.138	1254.0	25	0.140	0.055
7	k	t	1207.8	16	-0.141	1005.9	20	0.145	0.016
8	k	t	1215.0	13	-0.141	1005.7	23	0.145	0.042
9	k	t	811.7	11	-0.149	717.3	25	0.152	0.060
10	k	t	1350.2	20	-0.139	1295.3	16	0.140	-0.016
1	t	p	1066.2	7	-0.143	866.1	29	0.148	0.092
2	t	p	597.0	1	-0.156	961.5	35	0.146	0.138
3	t	p	812.1	6	-0.149	840.2	30	0.149	0.100
4	t	p	999.0	2	-0.145	1056.8	34	0.144	0.128
5	t	p	---	0	---	820.6	36	0.149	0.150
6	t	p	1655.2	14	-0.135	1247.8	22	0.140	0.033
7	t	p	---	0	---	896.2	36	0.147	0.148
8	t	p	1056.0	12	-0.144	976.1	24	0.145	0.049
9	t	p	829.0	2	-0.149	697.2	34	0.153	0.136
10	t	p	1056.0	1	-0.144	998.3	35	0.145	0.138

Table 4.10: Average perceptual distance values, based on responses by Dutch listeners to stop-without-burst stimuli produced by English speakers

			Same Responses			Different Responses			
Subj.	1	2	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	k	p	1199.0	2	-0.141	819.4	34	0.149	0.134
2	k	p	946.8	11	-0.146	826.1	25	0.149	0.058
3	k	p	---	0	---	926.9	36	0.146	0.148
4	k	p	821.4	10	-0.149	759.5	26	0.151	0.068
5	k	p	---	0	---	926.8	36	0.146	0.147
6	k	p	---	0	---	918.9	36	0.147	0.148
7	k	p	1162.0	1	-0.142	735.1	35	0.152	0.145
8	k	p	978.5	2	-0.145	892.3	34	0.147	0.131
9	k	p	---	0	---	855.3	36	0.148	0.149
10	k	p	1334.9	9	-0.139	1105.2	27	0.143	0.073
1	k	t	1059.4	8	-0.144	975.4	28	0.145	0.082
2	k	t	903.9	13	-0.147	727.7	23	0.152	0.044
3	k	t	1922.6	12	-0.132	1531.9	24	0.136	0.048
4	k	t	821.0	14	-0.149	979.6	22	0.145	0.032
5	k	t	1386.0	2	-0.138	1024.1	34	0.144	0.129
6	k	t	1110.2	5	-0.143	975.2	31	0.145	0.107
7	k	t	---	0	---	722.3	36	0.152	0.153
8	k	t	899.6	9	-0.147	955.6	27	0.146	0.074
9	k	t	915.4	5	-0.147	1017.8	31	0.144	0.106
10	k	t	1273.2	11	-0.140	1102.3	25	0.143	0.057
1	t	p	---	0	---	922.0	36	0.146	0.147
2	t	p	860.6	10	-0.148	792.4	26	0.150	0.067
3	t	p	2092.3	3	-0.131	1053.9	33	0.144	0.122
4	t	p	820.7	17	-0.149	746.7	19	0.151	0.009
5	t	p	---	0	---	928.9	36	0.146	0.147
6	t	p	813.0	3	-0.149	959.1	33	0.146	0.123
7	t	p	1961.0	1	-0.132	776.3	35	0.150	0.143
8	t	p	955.4	13	-0.146	925.0	23	0.146	0.042
9	t	p	---	0	---	882.2	36	0.147	0.148
10	t	p	1072.1	14	-0.143	1239.7	22	0.140	0.031

Table 4.11: Average perceptual distance values, based on responses by English listeners to stop-without-burst stimuli produced by English speakers

			Same Responses			Different Responses			
Subj.	1	2	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	ŋ	m	---	0	---	576.0	36	0.157	0.158
2	ŋ	m	---	0	---	874.1	36	0.148	0.148
3	ŋ	m	628.0	1	-0.155	634.2	35	0.155	0.147
4	ŋ	m	---	0	---	860.1	36	0.148	0.149
5	ŋ	m	---	0	---	720.3	36	0.152	0.153
6	ŋ	m	1272.0	1	-0.140	1047.4	35	0.144	0.137
7	ŋ	m	1513.0	1	-0.137	884.4	35	0.147	0.141
8	ŋ	m	994.5	2	-0.145	716.3	34	0.152	0.137
9	ŋ	m	881.0	1	-0.147	610.9	35	0.156	0.148
10	ŋ	m	1173.5	2	-0.141	1000.0	34	0.145	0.130
1	ŋ	n	---	0	---	624.4	36	0.155	0.156
2	ŋ	n	---	0	---	922.8	36	0.146	0.147
3	ŋ	n	1479.0	1	-0.137	625.6	35	0.155	0.148
4	ŋ	n	992.0	1	-0.145	961.7	35	0.146	0.138
5	ŋ	n	1165.0	1	-0.142	861.7	35	0.148	0.140
6	ŋ	n	1488.4	6	-0.137	1204.6	30	0.141	0.096
7	ŋ	n	1451.5	2	-0.137	860.4	34	0.148	0.133
8	ŋ	n	931.0	1	-0.146	919.8	35	0.147	0.140
9	ŋ	n	---	0	---	605.0	36	0.156	0.156
10	ŋ	n	1514.0	1	-0.137	811.3	35	0.149	0.142
1	n	m	894.1	7	-0.147	738.6	29	0.151	0.094
2	n	m	1462.0	1	-0.137	895.4	35	0.147	0.140
3	n	m	750.0	21	-0.151	796.5	15	0.150	-0.026
4	n	m	1608.0	1	-0.135	917.8	35	0.147	0.139
5	n	m	---	0	---	743.6	36	0.151	0.152
6	n	m	1493.6	18	-0.137	1488.8	18	0.137	0.000
7	n	m	1774.0	1	-0.134	902.8	35	0.147	0.141
8	n	m	1036.2	24	-0.144	954.6	12	0.146	-0.048
9	n	m	698.4	31	-0.153	796.9	5	0.150	-0.111
10	n	m	1114.3	3	-0.143	819.3	33	0.149	0.125

Table 4.12: Average perceptual distance values, based on responses by Dutch listeners to nasal stimuli produced by Dutch speakers

			Same Responses			Different Responses			
Subj.	1	2	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	ŋ	m	908.0	1	-0.147	812.3	35	0.149	0.142
2	ŋ	m	1025.5	2	-0.144	640.5	34	0.155	0.139
3	ŋ	m	---	0	---	736.3	36	0.151	0.152
4	ŋ	m	---	0	---	616.8	36	0.156	0.156
5	ŋ	m	---	0	---	908.0	36	0.147	0.147
6	ŋ	m	700.0	1	-0.153	775.5	35	0.150	0.143
7	ŋ	m	---	0	---	657.7	36	0.154	0.155
8	ŋ	m	---	0	---	776.1	36	0.150	0.151
9	ŋ	m	---	0	---	884.4	36	0.147	0.148
10	ŋ	m	1957.0	1	-0.132	944.4	35	0.146	0.139
1	ŋ	n	614.0	1	-0.156	895.1	35	0.147	0.139
2	ŋ	n	926.6	4	-0.146	619.5	32	0.156	0.123
3	ŋ	n	1755.0	3	-0.134	1097.2	33	0.143	0.121
4	ŋ	n	614.0	1	-0.156	620.3	35	0.156	0.148
5	ŋ	n	1190.0	1	-0.141	919.8	35	0.147	0.139
6	ŋ	n	---	0	---	823.6	36	0.149	0.151
7	ŋ	n	1025.0	1	-0.144	759.7	35	0.151	0.144
8	ŋ	n	---	0	---	726.9	36	0.152	0.152
9	ŋ	n	---	0	---	817.2	36	0.149	0.150
10	ŋ	n	1396.0	1	-0.138	912.0	35	0.147	0.139
1	n	m	908.7	4	-0.147	817.1	32	0.149	0.117
2	n	m	826.3	30	-0.149	649.7	6	0.154	-0.099
3	n	m	---	0	---	821.3	36	0.149	0.150
4	n	m	726.6	31	-0.152	1119.1	5	0.142	-0.110
5	n	m	---	0	---	905.8	36	0.147	0.147
6	n	m	1214.0	14	-0.141	1053.2	22	0.144	0.034
7	n	m	990.0	1	-0.145	706.5	35	0.152	0.145
8	n	m	793.1	34	-0.150	1114.5	2	0.143	-0.134
9	n	m	---	0	---	795.8	36	0.150	0.150
10	n	m	1010.6	34	-0.145	868.0	2	0.148	-0.129

Table 4.13: Average perceptual distance values, based on responses by English listeners to nasal stimuli produced by Dutch speakers

			Same Responses			Different Responses			
Subj.	1	2	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	ŋ	m	1056.8	3	-0.144	859.4	33	0.148	0.124
2	ŋ	m	1028.0	1	-0.144	1017.1	35	0.144	0.137
3	ŋ	m	961.5	2	-0.146	795.7	34	0.150	0.134
4	ŋ	m	---	0	---	986.2	36	0.145	0.146
5	ŋ	m	---	0	---	818.0	36	0.149	0.150
6	ŋ	m	1297.4	6	-0.140	1325.8	30	0.139	0.093
7	ŋ	m	1046.3	3	-0.144	985.0	33	0.145	0.122
8	ŋ	m	1288.9	7	-0.140	1034.4	29	0.144	0.089
9	ŋ	m	923.6	12	-0.146	1026.3	24	0.144	0.049
10	ŋ	m	1243.5	2	-0.140	1133.5	34	0.142	0.128
1	ŋ	n	1466.4	5	-0.137	832.7	31	0.149	0.109
2	ŋ	n	1510.0	7	-0.137	1102.2	29	0.143	0.089
3	ŋ	n	857.7	10	-0.148	831.6	26	0.149	0.067
4	ŋ	n	1180.2	14	-0.141	1105.4	22	0.143	0.032
5	ŋ	n	1395.0	4	-0.138	988.2	32	0.145	0.115
6	ŋ	n	1463.3	8	-0.137	1389.4	28	0.138	0.077
7	ŋ	n	1271.1	21	-0.140	973.9	15	0.145	-0.022
8	ŋ	n	1107.8	11	-0.143	1171.7	25	0.142	0.056
9	ŋ	n	948.3	19	-0.146	907.8	17	0.147	-0.009
10	ŋ	n	1306.1	10	-0.139	1081.2	26	0.143	0.065
1	n	m	855.5	2	-0.148	819.1	34	0.149	0.133
2	n	m	---	0	---	1017.7	36	0.144	0.145
3	n	m	839.5	2	-0.149	773.1	34	0.150	0.134
4	n	m	1160.0	1	-0.142	1049.0	35	0.144	0.136
5	n	m	---	0	---	831.1	36	0.149	0.149
6	n	m	1389.5	2	-0.138	1293.2	34	0.140	0.125
7	n	m	---	0	---	909.1	36	0.147	0.147
8	n	m	1322.2	9	-0.139	921.4	27	0.147	0.075
9	n	m	817.5	3	-0.149	829.8	33	0.149	0.125
10	n	m	1179.0	2	-0.141	1015.9	34	0.144	0.129

Table 4.14: Average perceptual distance values, based on responses by Dutch listeners to nasal stimuli produced by English speakers

			Same Responses			Different Responses			
Subj.	1	2	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	ŋ	m	1308.0	2	-0.139	850.3	34	0.148	0.133
2	ŋ	m	976.2	14	-0.145	948.6	22	0.146	0.034
3	ŋ	m	1356.3	3	-0.139	1128.4	33	0.142	0.121
4	ŋ	m	1042.5	7	-0.144	782.7	29	0.150	0.093
5	ŋ	m	---	0	---	944.0	36	0.146	0.146
6	ŋ	m	1372.9	5	-0.138	1008.6	31	0.145	0.107
7	ŋ	m	908.0	1	-0.147	755.5	35	0.151	0.143
8	ŋ	m	1660.0	2	-0.135	918.3	34	0.147	0.131
9	ŋ	m	---	0	---	869.2	36	0.148	0.148
10	ŋ	m	1498.3	11	-0.137	1148.8	25	0.142	0.057
1	ŋ	n	1116.3	11	-0.142	1057.5	25	0.144	0.056
2	ŋ	n	972.3	29	-0.145	770.8	7	0.150	-0.088
3	ŋ	n	1914.9	15	-0.132	1611.9	20	0.135	0.023
4	ŋ	n	905.4	14	-0.147	873.9	22	0.148	0.033
5	ŋ	n	1640.5	2	-0.135	1030.1	34	0.144	0.129
6	ŋ	n	---	0	---	870.0	35	0.148	0.149
7	ŋ	n	1146.0	2	-0.142	823.1	34	0.149	0.133
8	ŋ	n	926.0	8	-0.146	984.3	28	0.145	0.081
9	ŋ	n	1147.1	6	-0.142	915.8	30	0.147	0.099
10	ŋ	n	1429.8	11	-0.138	1414.1	25	0.138	0.055
1	n	m	1043.5	2	-0.144	845.5	34	0.148	0.133
2	n	m	986.3	18	-0.145	887.4	18	0.147	0.001
3	n	m	1496.0	1	-0.137	856.6	35	0.148	0.140
4	n	m	884.0	4	-0.147	839.0	32	0.149	0.116
5	n	m	---	0	---	919.2	36	0.147	0.147
6	n	m	---	0	---	894.9	36	0.147	0.149
7	n	m	---	0	---	746.8	36	0.151	0.152
8	n	m	1200.5	3	-0.141	881.3	33	0.147	0.124
9	n	m	---	0	---	850.8	36	0.148	0.149
10	n	m	1265.3	6	-0.140	1116.1	30	0.142	0.096

Table 4.15: Average perceptual distance values, based on responses by English listeners to nasal stimuli produced by English speakers

A repeated measures analysis of variance (ANOVA) was run on the data in Tables 4.4-4.15 in order to determine what effect each of these factors had on the amount of perceptual distance between each different VCCV pair. The results of this ANOVA are given in Table 4.16:

	F	df	Sig.
Manner	21.576	2,17	< **.001**
Manner * Listener	0.591	2,17	0.565
Speaker	0.318	1,18	0.580
Speaker * Listener	0.079	1,18	0.782
Pair	82.604	2,17	< **.001**
Pair * Listener	1.881	2,17	0.183
Manner * Speaker	0.384	2,17	0.687
Manner * Speaker * Listener	0.451	2,17	0.644
Manner * Pair	1.315	4,15	0.309
Manner * Pair * Listener	0.734	4,15	0.583
Speaker * Pair	64.794	2,17	< **.001**
Speaker * Pair * Listener	0.821	2,17	0.457
Manner * Speaker * Pair	1.081	4,15	0.401
Manner * Speaker * Pair * Listener	1.718	4,15	0.198
Listener	0.007	1,18	0.933

Table 4.16: Summary of ANOVA results for perceptual distance data

There are significant main effects in the ANOVA for both manner ($F=21.576$, $df=2,17$, $p < .001$) and place pair ($F = 82.604$, $df= 2,17$, $p < .001$), as well as a significant interaction between the language of speaker and place pair factors ($F=64.794$, $df=2,17$, $p < .001$). All of these factors were significant in the magnitude estimation task, as well.

Post-hoc t-tests—the results of which are given in Table 4.17—also reveal a similar pattern of results to those which were found in the earlier, magnitude estimation task. Perceptual distance is highest, for instance, between stops with bursts, and lowest for nasals and stops without bursts. There was no significant difference between the perceptual distances for nasals and stops without bursts. This pattern of results bears out the predictions for this study, and seems to indicate that the discrepancy between earlier studies which found no significant differences between the perceptual salience of stops and nasals (Ohala 1990, Hura et al. 1992, Winters 2001) and those which did (Mohr & Wang 1968, Pols 1983) may be due to the presence or absence of audible release burst cues for stops in the test stimuli. Chapters 5 and 6 will consider further what relevance stimuli with such acoustic cues have for cross-linguistic patterns in place assimilation.

		T-Tests	
Manner	Mean PD	vs. No Burst	vs. Nasal
Burst	0.081	< **.001**	**0.029**
No Burst	0.068	---	0.539
Nasal	0.070	---	---

		T-Tests	
Pair	Mean PD	vs. Lab-Dor	vs. Cor-Dor
Lab-Cor	0.061	< **.001**	**0.017**
Lab-Dor	0.088	---	< **.001**
Cor-Dor	0.070	---	---

Speaker*Pair			T-Tests	
Speaker	Pair	Mean PD	vs. Lab-Dor	vs. Cor-Dor
English	Lab-Cor	0.081	0.162	< **.001**
	Lab-Dor	0.085	---	< **.001**
	Cor-Dor	0.051	---	---
Dutch	Lab-Cor	0.042	< **.001**	< **.001**
	Lab-Dor	0.091	---	0.091
	Cor-Dor	0.090	---	---

Pair*Speaker			
Pair	Speaker	Mean PD	T-Tests
Lab-Cor	English	0.081	< **.001**
	Dutch	0.042	
Lab-Dor	English	0.085	**0.026**
	Dutch	0.091	
Cor-Dor	English	0.051	< **.001**
	Dutch	0.090	

Table 4.17: Post-hoc t-tests for significant effects in perceptual distance ANOVA

Figure 4.8 provides a graphical representation of the significant speaker*pair interaction in the ANOVA. This figure reveals a familiar pattern: perceptual distances are significantly smaller for coronal-labial contrasts in Dutch and dorsal-coronal contrasts in English. There are no significant differences between the perceptual distances of the other two place contrast pairs in both languages. These results indicate that coronal is the most confusable place of articulation, perceptually, in both languages, but that the place of articulation with which it is most often confused depends on the language in which it is spoken. This interaction presumably reflects the fine-grained phonetic differences between the two languages: the Dutch coronals were produced with a dental place of articulation, and are perceptually closer to bilabials, while English coronals are produced as alveolars, and are perceptually closer to dorsal stops and nasals.

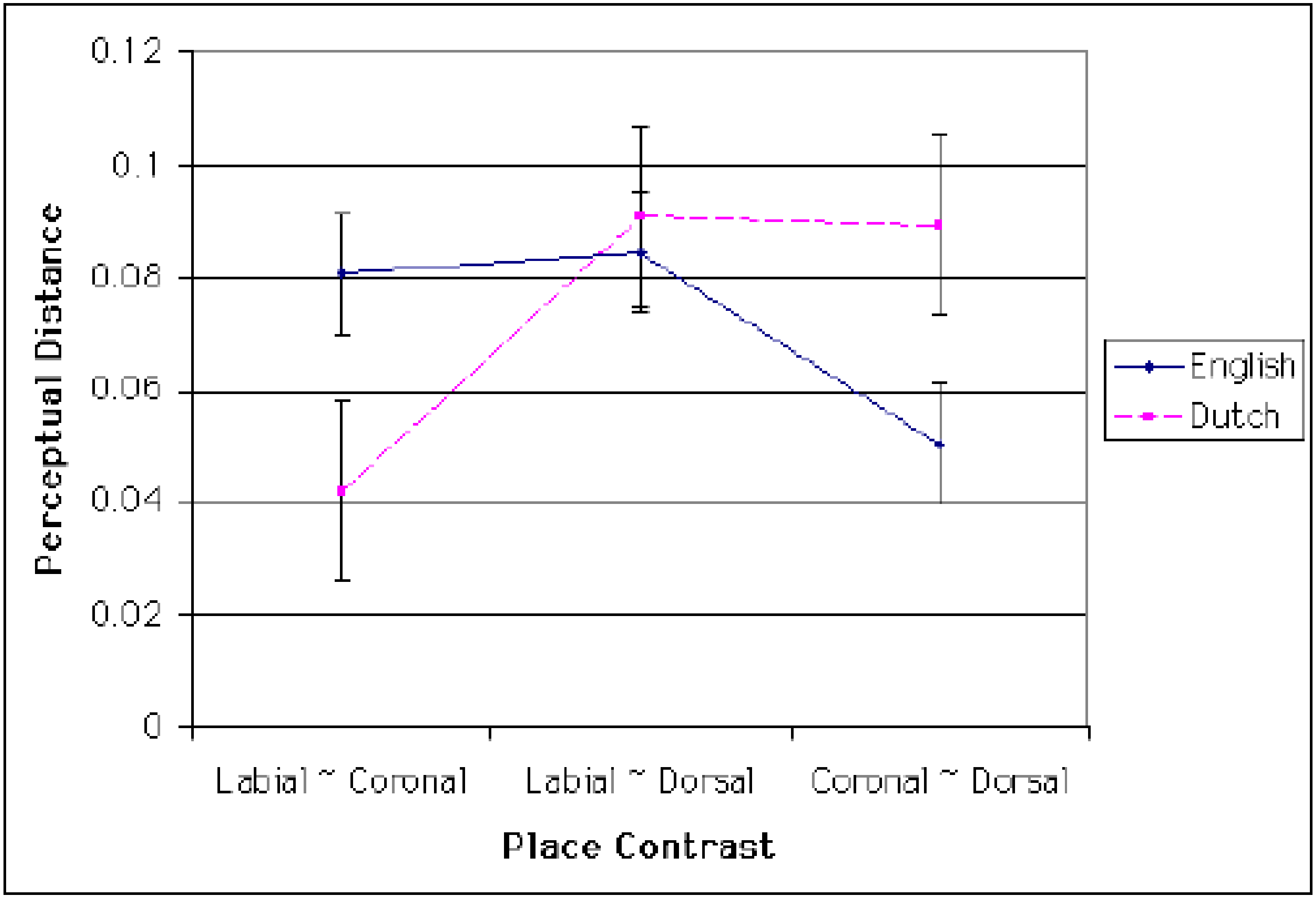


Figure 4.8: Perceptual distance, by place contrast and language of speaker

To a certain extent, this result confirms the analysis of Jun (1995), who claimed that coronals undergo place assimilation in many languages of the world because the perceptual cues to their place of articulation are weaker than the same cues for labials or dorsals. There is not, however, a corresponding effect for manner of articulation in these results—nasal distinctions are not less salient, perceptually, than the same distinctions in stops without bursts. Nor is there a similar effect for the native language of the listener—even though only nasal coronals may undergo place assimilation in Dutch, and the results of Pols (1983) suggested that there may be cross-linguistic differences between Dutch and English listeners' ability to perceive place in nasals. Instead, the results of both this study and the magnitude estimation task suggest that Pols' (1983) findings may be explained in terms of the articulatory differences between the two languages. Pols (1983) only tested the perception of the labial and coronal nasals, /m/ and /n/, in his study. This study has revealed that the small perceptual distance between these two segments in Dutch is not representative of Dutch listeners' general ability to distinguish between the place of articulation of all the nasals in their phonemic inventory.

Also absent from the ANOVA results are significant effects for any language of speaker and language of listener interactions, even though these emerged from the results of the magnitude estimation task. It had been earlier hypothesized that the presence of these effects in estimated magnitudes of difference showed that this task involves higher-level processing beyond simply interpreting the "perceptual distance" between two different sounds. The lack of such effects in this study indicates that AX discrimination and magnitude estimation may, in fact, involve different levels of perceptual processing. Nothing in the results of either study, however, suggests that one type of processing or the other would be more likely to motivate a cross-linguistic tendency for nasals to undergo place assimilation.

This analysis of perceptual distance, however, has only considered the discriminability of place in the C_1 consonant without respect to the place of articulation of the subsequent C_2 consonant. The place of articulation of this C_2 consonant may have a significant effect on the perceptibility of place distinctions in the C_1 consonant. This contextual influence has particular relevance to the study of perception's interaction with place assimilation, since this process involves the C_1 consonant taking on the same place of articulation as the following consonant. Ohala (1990), in fact, suggested that the place cues for C_2 in a heterorganic C_1C_2 sequence can overwhelm the place cues for C_1, thereby deceiving the listener into thinking the entire consonant only has one place of articulation—namely, that of the C_2 consonant. Ohala (1990) claimed, in fact, that simply mis-perceiving consonant clusters in this way is what causes place assimilation processes to emerge in phonology.

Although the AX discrimination experiment was not specifically designed to test Ohala's (1990) theory, it is possible to interpret its results in a way which might be relevant to a theory of "innocent misapprehensions" and their relationship to place assimilation. Place assimilation affects heterorganic clusters of consonants by transforming them into homorganic clusters of consonants. Importantly, it does not change them into other heterorganic clusters (e.g., /pk./ --> [tk]). If misperceptions truly do motivate place assimilation as Ohala (1990) suggested, it should occur as the result of confusing heterorganic clusters with homorganic clusters more often than with other heterorganic clusters.

In terms of the VCCV stimuli pairs in this experiment, this hypothesis would therefore predict that listeners would find the heterorganic-homorganic pairs more confusing than the heterorganic-heterorganic pairs. The perceptual distance between the pairs below on the left, in other words, should be smaller than the perceptual distance between the pairs on the right.

heterorganic ~ homorganic pairs	heterorganic ~ heterorganic pairs
atpa ~ appa	atpa ~ akpa
akpa ~ appa	
apta ~ atta	apta ~ akta
akta ~ atta	
apka ~ akka	apka ~ atka
atka ~ akka	
anpa ~ ampa	anpa ~ aŋpa
aŋpa ~ ampa	
amta ~ anta	amta ~ aŋta
aŋta ~ anta	
amka ~ aŋka	amka ~ anka
anka ~ aŋka	

Table 4.18: Assimilatory (heterorganic-homorganic) and non-assimilatory (heterorganic-heterorganic) contrast pairs

Tables 4.19-4.30 present one attempt to test the hypothesis that heterorganic-homorganic pairs are more perceptually confusable than heterorganic-heterorganic pairs. These tables break down perceptual distance values from the AX discrimination experiment by the type of cluster contrast, where heterorganic-homorganic pairs are labeled as “assimilatory” and heterorganic-heterorganic pairs are “non-assimilatory.”

Each table presents these values for a particular manner of articulation (nasal, stop with burst, stop without burst), broken down further by language of speaker (English, Dutch), and language of listener (English, Dutch).

		Same Responses			Different Responses			
Subj.	Contrast Type	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	Assimilatory	---	0	---	721.0	72	0.152	0.153
2	Assimilatory	1110.2	20	-0.143	1012.8	52	0.144	0.068
3	Assimilatory	856.2	4	-0.148	659.7	68	0.154	0.138
4	Assimilatory	1107.0	16	-0.143	966.6	56	0.145	0.084
5	Assimilatory	1262.2	9	-0.140	851.4	63	0.148	0.112
6	Assimilatory	1833.0	2	-0.133	1260.5	70	0.140	0.133
7	Assimilatory	1024.1	16	-0.144	1069.7	56	0.143	0.086
8	Assimilatory	1644.2	4	-0.135	841.5	67	0.148	0.133
9	Assimilatory	955.0	1	-0.146	667.5	71	0.154	0.150
10	Assimilatory	1126.9	13	-0.142	1008.7	59	0.145	0.093
1	Non-Assim	564.0	1	-0.158	701.6	35	0.153	0.144
2	Non-Assim	1049.4	10	-0.144	1004.1	26	0.145	0.065
3	Non-Assim	785.0	3	-0.150	692.9	33	0.153	0.128
4	Non-Assim	1367.7	9	-0.138	974.7	27	0.145	0.074
5	Non-Assim	1118.7	5	-0.142	865.8	31	0.148	0.108
6	Non-Assim	---	0	---	1084.8	36	0.143	0.144
7	Non-Assim	1135.8	7	-0.142	980.1	29	0.145	0.091
8	Non-Assim	680.5	2	-0.153	840.6	34	0.148	0.133
9	Non-Assim	1215.0	1	-0.141	644.9	35	0.155	0.147
10	Non-Assim	1196.0	2	-0.141	954.5	34	0.146	0.130

Table 4.19: Perceptual Distance, by contrast type, based on responses by Dutch listeners to stop-with-burst stimuli produced by Dutch speakers

		Same Responses			Different Responses			
Subj.	Contrast Type	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	Assimilatory	1323.1	9	-0.139	882.5	62	0.147	0.111
2	Assimilatory	1011.8	5	-0.145	685.5	67	0.153	0.133
3	Assimilatory	1403.0	9	-0.138	983.4	63	0.145	0.112
4	Assimilatory	1278.0	1	-0.140	666.0	71	0.154	0.150
5	Assimilatory	1658.7	4	-0.135	963.5	68	0.146	0.130
6	Assimilatory	1485.0	2	-0.137	828.2	70	0.149	0.143
7	Assimilatory	817.6	12	-0.149	717.1	60	0.152	0.106
8	Assimilatory	1094.4	7	-0.143	892.3	65	0.147	0.120
9	Assimilatory	959.2	8	-0.146	960.5	64	0.146	0.118
10	Assimilatory	1275.0	9	-0.140	1008.9	63	0.145	0.110
1	Non-Assim	1012.5	2	-0.145	882.5	34	0.147	0.132
2	Non-Assim	913.3	3	-0.147	676.2	33	0.153	0.129
3	Non-Assim	1599.7	3	-0.136	841.5	33	0.148	0.125
4	Non-Assim	607.0	1	-0.156	683.4	35	0.153	0.145
5	Non-Assim	---	0	---	940.8	36	0.146	0.147
6	Non-Assim	1139.0	1	-0.142	868.0	35	0.148	0.142
7	Non-Assim	959.5	4	-0.146	782.0	32	0.150	0.118
8	Non-Assim	---	0	---	869.9	36	0.148	0.149
9	Non-Assim	---	0	---	961.3	36	0.146	0.146
10	Non-Assim	1973.5	2	-0.132	978.5	34	0.145	0.130

Table 4.20: Perceptual Distance, by contrast type, based on responses by English listeners to stop-with-burst stimuli produced by Dutch speakers

		Same Responses			Different Responses			
Subj.	Contrast Type	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	Assimilatory	1066.4	15	-0.143	887.1	57	0.147	0.088
2	Assimilatory	1374.4	6	-0.138	950.3	66	0.146	0.123
3	Assimilatory	894.3	15	-0.147	845.3	57	0.148	0.088
4	Assimilatory	1256.0	3	-0.140	995.8	69	0.145	0.134
5	Assimilatory	1679.0	3	-0.135	877.0	69	0.148	0.137
6	Assimilatory	1616.5	17	-0.135	1276.6	55	0.140	0.076
7	Assimilatory	1888.1	6	-0.133	932.8	66	0.146	0.124
8	Assimilatory	1116.2	6	-0.142	911.4	66	0.147	0.124
9	Assimilatory	881.4	9	-0.147	764.6	63	0.151	0.114
10	Assimilatory	1298.5	9	-0.139	1020.6	63	0.144	0.110
1	Non-Assim	---	0	---	756.6	36	0.151	0.152
2	Non-Assim	---	0	---	974.1	36	0.145	0.146
3	Non-Assim	714.5	2	-0.152	733.1	34	0.152	0.135
4	Non-Assim	---	0	---	940.8	36	0.146	0.147
5	Non-Assim	---	0	---	833.4	36	0.149	0.150
6	Non-Assim	1208.5	2	-0.141	1203.3	34	0.141	0.126
7	Non-Assim	1238.7	4	-0.140	861.4	32	0.148	0.117
8	Non-Assim	1071.0	2	-0.143	867.0	34	0.148	0.132
9	Non-Assim	1222.0	1	-0.141	669.4	35	0.154	0.146
10	Non-Assim	1360.0	1	-0.139	931.2	35	0.146	0.139

Table 4.21: Perceptual Distance, by contrast type, based on responses by Dutch listeners to stop-with-burst stimuli produced by English speakers

		Same Responses			Different Responses			
Subj.	Contrast Type	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	Assimilatory	1017.9	7	-0.144	908.4	65	0.147	0.119
2	Assimilatory	944.2	15	-0.146	771.2	57	0.150	0.089
3	Assimilatory	1714.8	12	-0.134	996.8	60	0.145	0.099
4	Assimilatory	880.1	11	-0.147	777.3	61	0.150	0.105
5	Assimilatory	1323.8	3	-0.139	1000.8	69	0.145	0.134
6	Assimilatory	2420.3	3	-0.128	927.6	69	0.146	0.137
7	Assimilatory	1021.5	2	-0.144	748.1	70	0.151	0.144
8	Assimilatory	1075.1	12	-0.143	991.1	60	0.145	0.097
9	Assimilatory	---	0	---	886.8	72	0.147	0.148
10	Assimilatory	1197.3	18	-0.141	1220.4	54	0.141	0.071
1	Non-Assim	934.0	3	-0.146	878.7	33	0.148	0.124
2	Non-Assim	970.5	4	-0.145	756.8	32	0.151	0.118
3	Non-Assim	1093.5	2	-0.143	901.7	34	0.147	0.132
4	Non-Assim	1064.0	2	-0.143	730.9	34	0.152	0.136
5	Non-Assim	---	0	---	986.3	36	0.145	0.146
6	Non-Assim	---	0	---	927.3	36	0.146	0.148
7	Non-Assim	---	0	---	730.4	36	0.152	0.152
8	Non-Assim	1202.0	1	-0.141	867.5	35	0.148	0.140
9	Non-Assim	---	0	---	906.8	36	0.147	0.148
10	Non-Assim	---	0	---	1103.8	36	0.143	0.144

Table 4.22: Perceptual Distance, by contrast type, based on responses by English listeners to stop-with-burst stimuli produced by English speakers

		Same Responses			Different Responses			
Subj.	Contrast Type	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	Assimilatory	901.3	3	-0.147	677.9	69	0.153	0.142
2	Assimilatory	1175.4	21	-0.141	1082.5	51	0.143	0.060
3	Assimilatory	808.4	9	-0.149	669.5	63	0.154	0.116
4	Assimilatory	1086.8	17	-0.143	1026.1	55	0.144	0.077
5	Assimilatory	1058.2	10	-0.144	934.5	62	0.146	0.107
6	Assimilatory	1825.5	8	-0.133	1272.7	64	0.140	0.110
7	Assimilatory	1099.3	25	-0.143	1045.6	47	0.144	0.045
8	Assimilatory	1238.2	6	-0.140	942.0	66	0.146	0.123
9	Assimilatory	987.0	10	-0.145	682.5	62	0.153	0.112
10	Assimilatory	1209.1	21	-0.141	1164.7	51	0.142	0.061
1	Non-Assim	660.0	2	-0.154	746.7	34	0.151	0.136
2	Non-Assim	1124.9	11	-0.142	1133.4	25	0.142	0.056
3	Non-Assim	695.0	2	-0.153	636.3	34	0.155	0.138
4	Non-Assim	979.3	7	-0.145	946.0	29	0.146	0.090
5	Non-Assim	1434.0	3	-0.138	829.8	33	0.149	0.125
6	Non-Assim	1194.5	2	-0.141	1184.6	34	0.141	0.126
7	Non-Assim	1210.9	6	-0.141	1168.0	30	0.142	0.096
8	Non-Assim	641.5	2	-0.155	905.0	34	0.147	0.132
9	Non-Assim	---	0	---	663.5	36	0.154	0.154
10	Non-Assim	1249.7	15	-0.140	1033.2	21	0.144	0.026

Table 4.23: Perceptual Distance, by contrast type, based on responses by Dutch listeners to stop-without-burst stimuli produced by Dutch speakers

		Same Responses			Different Responses			
Subj.	Contrast Type	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	Assimilatory	980.5	22	-0.145	896.5	50	0.147	0.058
2	Assimilatory	1067.6	9	-0.143	668.4	63	0.154	0.117
3	Assimilatory	1223.7	17	-0.141	798.7	55	0.150	0.081
4	Assimilatory	747.5	2	-0.151	666.0	70	0.154	0.146
5	Assimilatory	1608.1	12	-0.135	1018.4	60	0.144	0.099
6	Assimilatory	607.0	1	-0.156	912.9	71	0.147	0.144
7	Assimilatory	864.4	21	-0.148	802.4	51	0.150	0.064
8	Assimilatory	1388.2	6	-0.138	872.9	66	0.148	0.125
9	Assimilatory	1033.0	26	-0.144	997.0	46	0.145	0.041
10	Assimilatory	1150.7	15	-0.142	1018.5	57	0.144	0.086
1	Non-Assim	1059.5	6	-0.144	823.4	30	0.149	0.101
2	Non-Assim	853.8	3	-0.148	661.1	33	0.154	0.129
3	Non-Assim	2070.5	2	-0.131	800.1	34	0.150	0.134
4	Non-Assim	---	0	---	734.3	36	0.152	0.152
5	Non-Assim	1254.0	2	-0.140	1023.0	34	0.144	0.129
6	Non-Assim	---	0	---	846.6	36	0.148	0.150
7	Non-Assim	897.0	10	-0.147	817.1	26	0.149	0.068
8	Non-Assim	791.0	2	-0.150	835.8	33	0.149	0.132
9	Non-Assim	931.1	7	-0.146	1167.6	29	0.142	0.088
10	Non-Assim	1274.5	3	-0.140	948.7	33	0.146	0.123

Table 4.24: Perceptual Distance, by contrast type, based on responses by English listeners to stop-without-burst stimuli produced by Dutch speakers

		Same Responses			Different Responses			
Subj.	Contrast Type	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	Assimilatory	981.1	12	-0.145	879.1	60	0.148	0.100
2	Assimilatory	1202.2	5	-0.141	1080.6	67	0.143	0.125
3	Assimilatory	888.2	19	-0.147	856.7	53	0.148	0.071
4	Assimilatory	1033.0	3	-0.144	966.9	69	0.145	0.134
5	Assimilatory	1261.8	7	-0.140	837.9	65	0.149	0.121
6	Assimilatory	1488.0	20	-0.137	1274.1	52	0.140	0.063
7	Assimilatory	1375.9	8	-0.138	917.0	64	0.147	0.115
8	Assimilatory	1080.5	21	-0.143	1045.2	51	0.144	0.061
9	Assimilatory	795.3	14	-0.150	720.0	58	0.152	0.094
10	Assimilatory	1310.3	16	-0.139	1075.8	56	0.143	0.082
1	Non-Assim	925.0	1	-0.146	785.1	35	0.150	0.142
2	Non-Assim	783.3	3	-0.150	996.9	33	0.145	0.121
3	Non-Assim	794.8	5	-0.150	768.9	31	0.150	0.109
4	Non-Assim	1609.0	1	-0.135	962.6	35	0.146	0.138
5	Non-Assim	1069.0	1	-0.143	838.8	35	0.149	0.141
6	Non-Assim	1603.9	8	-0.135	1186.3	28	0.141	0.080
7	Non-Assim	871.4	8	-0.148	875.9	28	0.148	0.083
8	Non-Assim	1305.0	6	-0.139	924.6	30	0.146	0.099
9	Non-Assim	---	0	---	705.1	36	0.152	0.153
10	Non-Assim	1562.3	7	-0.136	1053.4	29	0.144	0.090

Table 4.25: Perceptual Distance, by contrast type, based on responses by Dutch listeners to stop-without-burst stimuli produced by English speakers

Subj.	Contrast Type	Same Responses Mean	No.	-1/ln(rt)	Different Responses Mean	No.	1/ln(rt)	PD
1	Assimilatory	1098.5	8	-0.143	913.4	64	0.147	0.115
2	Assimilatory	912.8	25	-0.147	768.4	47	0.151	0.047
3	Assimilatory	1863.1	10	-0.133	1220.0	62	0.141	0.105
4	Assimilatory	806.2	33	-0.149	861.5	39	0.148	0.013
5	Assimilatory	1534.0	1	-0.136	978.9	71	0.145	0.142
6	Assimilatory	930.1	6	-0.146	999.8	66	0.145	0.123
7	Assimilatory	1162.0	1	-0.142	756.9	71	0.151	0.148
8	Assimilatory	958.6	23	-0.146	961.1	49	0.146	0.053
9	Assimilatory	900.8	4	-0.147	950.5	68	0.146	0.131
10	Assimilatory	1259.6	28	-0.140	1159.2	44	0.142	0.032
1	Non-Assim	1023.5	2	-0.144	889.9	34	0.147	0.132
2	Non-Assim	868.2	9	-0.148	809.3	27	0.149	0.075
3	Non-Assim	1706.2	5	-0.134	1072.8	31	0.143	0.107
4	Non-Assim	848.7	8	-0.148	762.7	28	0.151	0.084
5	Non-Assim	1238.0	1	-0.140	922.0	35	0.146	0.139
6	Non-Assim	1039.5	2	-0.144	853.7	34	0.148	0.133
7	Non-Assim	1961.0	1	-0.132	719.9	35	0.152	0.145
8	Non-Assim	896.0	1	-0.147	863.6	35	0.148	0.140
9	Non-Assim	930.0	1	-0.146	854.3	35	0.148	0.140
10	Non-Assim	1071.2	6	-0.143	1128.8	30	0.142	0.096

Table 4.26: Perceptual Distance, by contrast type, based on responses by English listeners to stop-without-burst stimuli produced by English speakers

		Same Responses			Different Responses			
Subj.	Contrast Type	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	Assimilatory	877.8	7	-0.148	654.5	65	0.154	0.126
2	Assimilatory	1462.0	1	-0.137	917.1	71	0.147	0.143
3	Assimilatory	721.7	14	-0.152	708.0	58	0.152	0.095
4	Assimilatory	---	0	---	894.9	72	0.147	0.148
5	Assimilatory	1165.0	1	-0.142	785.7	71	0.150	0.147
6	Assimilatory	1468.0	20	-0.137	1305.7	52	0.139	0.064
7	Assimilatory	---	0	---	885.8	72	0.147	0.149
8	Assimilatory	1038.0	18	-0.144	817.7	54	0.149	0.076
9	Assimilatory	724.1	22	-0.152	675.9	50	0.153	0.061
10	Assimilatory	1425.8	5	-0.138	891.1	67	0.147	0.128
1	Non-Assim	---	0	---	630.1	36	0.155	0.156
2	Non-Assim	---	0	---	858.3	36	0.148	0.148
3	Non-Assim	999.4	9	-0.145	640.3	27	0.155	0.079
4	Non-Assim	1300.0	2	-0.139	949.9	34	0.146	0.131
5	Non-Assim	---	0	---	754.2	36	0.151	0.151
6	Non-Assim	1576.0	5	-0.136	1129.3	31	0.142	0.105
7	Non-Assim	1579.5	4	-0.136	876.0	32	0.148	0.117
8	Non-Assim	992.4	9	-0.145	901.8	27	0.147	0.076
9	Non-Assim	725.4	10	-0.152	620.6	26	0.156	0.071
10	Non-Assim	698.0	1	-0.153	848.5	35	0.148	0.141

Table 4.27: Perceptual Distance, by contrast type, based on responses by Dutch listeners to nasal stimuli produced by Dutch speakers

		Same Responses			Different Responses			
Subj.	Contrast Type	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	Assimilatory	908.0	1	-0.147	840.4	71	0.149	0.145
2	Assimilatory	871.2	23	-0.148	655.0	49	0.154	0.058
3	Assimilatory	1885.0	2	-0.133	900.4	70	0.147	0.141
4	Assimilatory	741.6	20	-0.151	717.4	52	0.152	0.070
5	Assimilatory	1190.0	1	-0.141	906.3	71	0.147	0.143
6	Assimilatory	1149.4	12	-0.142	913.0	60	0.147	0.101
7	Assimilatory	990.0	1	-0.145	707.1	71	0.152	0.149
8	Assimilatory	779.3	23	-0.150	771.1	49	0.150	0.055
9	Assimilatory	---	0	---	830.9	72	0.149	0.149
10	Assimilatory	1234.5	23	-0.140	931.7	49	0.146	0.054
1	Non-Assim	810.4	5	-0.149	843.8	31	0.148	0.108
2	Non-Assim	903.2	13	-0.147	610.4	23	0.156	0.047
3	Non-Assim	1625.0	1	-0.135	854.1	35	0.148	0.141
4	Non-Assim	668.9	12	-0.154	659.9	24	0.154	0.053
5	Non-Assim	---	0	---	921.0	36	0.147	0.147
6	Non-Assim	948.7	3	-0.146	826.2	33	0.149	0.126
7	Non-Assim	1025.0	1	-0.144	709.6	35	0.152	0.145
8	Non-Assim	820.5	11	-0.149	861.4	25	0.148	0.059
9	Non-Assim	---	0	---	835.6	36	0.149	0.149
10	Non-Assim	1081.5	13	-0.143	889.3	23	0.147	0.041

Table 4.28: Perceptual Distance, by contrast type, based on responses by English listeners to nasal stimuli produced by Dutch speakers

		Same Responses			Different Responses			
Subj.	Contrast Type	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	Assimilatory	1248.3	8	-0.140	841.9	64	0.148	0.116
2	Assimilatory	1456.0	4	-0.137	1078.5	68	0.143	0.128
3	Assimilatory	876.6	12	-0.148	804.7	60	0.149	0.101
4	Assimilatory	1211.1	11	-0.141	1025.6	61	0.144	0.101
5	Assimilatory	1294.0	3	-0.140	904.1	69	0.147	0.136
6	Assimilatory	1275.9	10	-0.140	1368.2	62	0.138	0.100
7	Assimilatory	1096.9	18	-0.143	976.0	54	0.145	0.074
8	Assimilatory	1130.7	21	-0.142	1068.2	51	0.143	0.061
9	Assimilatory	865.1	29	-0.148	942.7	43	0.146	0.028
10	Assimilatory	1297.0	12	-0.140	1046.9	60	0.144	0.097
1	Non-Assim	908.5	2	-0.147	827.4	34	0.149	0.133
2	Non-Assim	1244.0	4	-0.140	980.1	32	0.145	0.114
3	Non-Assim	825.0	2	-0.149	791.0	34	0.150	0.134
4	Non-Assim	1092.8	4	-0.143	1089.3	32	0.143	0.112
5	Non-Assim	1394.0	1	-0.138	829.0	35	0.149	0.142
6	Non-Assim	1541.6	6	-0.136	1271.9	30	0.140	0.094
7	Non-Assim	1569.0	6	-0.136	916.1	30	0.147	0.100
8	Non-Assim	1460.6	6	-0.137	1021.7	30	0.144	0.098
9	Non-Assim	1108.4	5	-0.143	890.1	31	0.147	0.108
10	Non-Assim	1058.0	2	-0.144	1136.9	34	0.142	0.127

Table 4.29: Perceptual Distance, by contrast type, based on responses by Dutch listeners to nasal stimuli produced by English speakers

		Same Responses			Different Responses			
Subj.	Contrast Type	Mean	No.	-1/ln(rt)	Mean	No.	1/ln(rt)	PD
1	Assimilatory	1186.6	10	-0.141	932.6	62	0.146	0.107
2	Assimilatory	989.9	45	-0.145	913.2	27	0.147	-0.036
3	Assimilatory	1532.8	15	-0.136	1140.8	56	0.142	0.085
4	Assimilatory	985.6	20	-0.145	833.8	52	0.149	0.067
5	Assimilatory	1487.0	1	-0.137	959.5	71	0.146	0.142
6	Assimilatory	1380.5	4	-0.138	936.7	68	0.146	0.132
7	Assimilatory	908.0	1	-0.147	773.0	71	0.150	0.147
8	Assimilatory	1227.6	9	-0.141	937.1	63	0.146	0.110
9	Assimilatory	1119.8	5	-0.142	884.6	67	0.147	0.128
10	Assimilatory	1444.2	21	-0.137	1248.9	51	0.140	0.060
1	Non-Assim	1070.4	5	-0.143	888.1	31	0.147	0.108
2	Non-Assim	955.3	16	-0.146	817.4	20	0.149	0.018
3	Non-Assim	1777.8	4	-0.134	1135.6	32	0.142	0.115
4	Non-Assim	922.1	5	-0.146	828.1	31	0.149	0.108
5	Non-Assim	1794.0	1	-0.133	974.3	35	0.145	0.138
6	Non-Assim	1651.0	1	-0.135	900.1	34	0.147	0.140
7	Non-Assim	1146.0	2	-0.142	779.3	34	0.150	0.135
8	Non-Assim	1146.3	4	-0.142	909.7	32	0.147	0.115
9	Non-Assim	1570.0	1	-0.136	866.6	35	0.148	0.140
10	Non-Assim	1362.1	7	-0.139	1181.3	29	0.141	0.088

Table 4.30: Perceptual Distance, by contrast type, based on responses by English listeners to nasal stimuli produced by English speakers

A repeated measures analysis of variance was run on the data in these tables, taking into account the effects of manner of articulation, contrast type, language of listener and language of speaker had on the individual values of perceptual distances. Table 4.31 lists the results of this ANOVA. Several significant effects emerged from this analysis; only those including the contrast type factor will be addressed here, since all others were identical to those factors which were significant in the previous perceptual distance ANOVA (Table 4.16). Of this subset of factors, the overall contrast type factor was significant ($F = 42.134$, $df = 1,18$, $p < .001$). Post-hoc t-tests (Table 4.32) reveal that the direction of this effect was in the predicted direction: there was significantly more perceptual distance between non-assimilatory pairs than between assimilatory pairs ($p < .001$). This indicates that there is some perceptual motivation for place assimilation to occur in the manner that Ohala (1990) suggested. There are also significant interactions between contrast type and both manner of articulation ($F = 4.775$, $df = 2,17$, $p = .0123$) and language of speaker ($F = 8.341$, $df = 1,18$, $p = .01$). While intriguing, the manner of articulation-contrast type interaction does not provide any new insights into nasal susceptibility to place assimilation. Post-hoc t-tests reveal that this interaction consists primarily of a lack of a significant perceptual difference ($p = .07$) between assimilatory pairs in nasals and stops with bursts—even though non-assimilatory pairs have significantly greater perceptual distances in stops with bursts than in nasals ($p = .005$). The significant language of speaker-contrast type interaction is equally unenlightening; post-hoc t-tests reveal that it consists solely of a significant difference in perceptual distances between non-assimilatory pairs in English and Dutch ($p = .047$). These pairs have significantly greater perceptual distances in English than in Dutch. There are at least two possible explanations for this result--one being that the most confusable contrast in Dutch—labials vs. coronals—is generally more confusable than the most confusable contrast in English (coronals vs. dorsals). The second is that all of the Dutch listeners had some knowledge of English, and this may have enabled them to perceive distinctions in the English stimuli better than the English listeners could perceive the corresponding place distinctions in Dutch.

	F	df	p
Manner	18.445	2,17	< **.001**
Manner * Listener	0.598	2,17	0.561
Speaker	0.025	1,18	0.875
Speaker * Listener	0.149	1,18	0.704
Contrast	42.134	1,18	< **.001**
Contrast * Listener	0.291	1,18	0.596
Manner * Speaker	0.577	2,17	0.572
Manner * Speaker * Listener	0.720	2,17	0.501
Manner * Contrast	4.775	2,17	**0.023**
Manner * Contrast * Listener	2.759	2,17	0.092
Speaker * Contrast	8.341	1,18	**0.010**
Speaker * Contrast * Listener	0.220	1,18	0.645
Manner * Speaker * Contrast	1.400	2,17	0.274
Manner * Speaker * Contrast * Listener	0.670	2,17	0.525
Listener	0.022	1,18	0.884

Table 4.31: Summary of results of analysis of variance for contrast type data

Manner	Mean PD	T-Tests vs. No Burst	vs. Nasal
Burst	0.124	< **<u>.001</u>**	**<u>0.018</u>**
No Burst	0.105	---	0.786
Nasal	0.107	---	---

Contrast	Mean PD	T-Test
Assim	0.104	< **<u>.001</u>**
Non-Assim	0.120	---

Manner * Contrast

Manner	Contrast	Mean PD	T-Tests
Burst	Assim	0.116	< **<u>.001</u>**
	Non-Assim	0.133	
No Burst	Assim	0.095	< **<u>.001</u>**
	Non-Assim	0.116	
Nasal	Assim	0.102	**<u>0.008</u>**
	Non-Assim	0.112	

Contrast * Manner

Contrast	Manner	Mean PD	T-Tests vs. No Burst
Assim	Burst	0.116	< **<u>.001</u>**
	No Burst	0.095	---
	Nasal	0.102	---
Non-Assim	Burst	0.133	**<u>0.004</u>**
	No Burst	0.116	---
	Nasal	0.112	---

Speaker * Contrast

Speaker	Contrast	Mean PD	T-Tests
English	Assim	0.100	< **<u>.001</u>**
	Non-Assim	0.123	---
Dutch	Assim	0.108	**<u>0.006</u>**
	Non-Assim	0.117	---

Contrast * Speaker

Contrast	Speaker	Mean PD	T-Tests
Assim	English	0.100	0.147
	Dutch	0.108	---
Non-Assim	English	0.123	**<u>0.047</u>**
	Dutch	0.117	---

Table 4.32: Summary of post-hoc t-tests for significant effects in contrast type ANOVA

The overall pattern of results in the AX discrimination experiment reflects the results found in the magnitude estimation study: perceptual distinctiveness is greatest in stops with bursts, and significantly less in nasals and stops without bursts. The lack of a significant perceptual difference between these last two classes of sounds confirms earlier non-effects found in Ohala (1990), Hura et al. (1992) and Winters (2001). The fact that audible release burst cues make stops more perceptually salient than nasals may provide an explanation for why stops undergo place assimilation less often, cross-linguistically, than nasals. This hypothesis will be pursued further in the next chapter.

The results of the AX discrimination study also confirmed the unexpected finding of the magnitude estimation study that the confusability of certain place contrasts depended on the language in which they were spoken. Coronal-labial contrasts were highly confusable in Dutch, while coronal-dorsal contrasts were highly confusable in English. This perceptual difference seems to have its origins in the articulatory differences between the two languages. The fact that the most confusable contrasts in both languages involve coronal consonants, however, may account for why only coronals undergo place assimilation in English and Dutch.

Considering the influence of context on the perception of place contrasts also seems to indicate that there is some perceptual motivation for place assimilation to occur because of "innocent misapprehensions." Perceptual distances between heterorganic and homorganic VCCV pairs were considerably lower than the perceptual distances between comparable pairs of heterorganic sequences. Listeners seem, therefore, to be more likely to confuse heterorganic clusters with homorganic clusters in real-world perceptual situations; such confusion might provide the sole, perceptual origin for place assimilation in stops and nasals—as Ohala (1990) suggested.

This data does not, however, reveal significant differences between listener perception of nasals and stops—nor any significant differences between Dutch and English listeners' perception of place in these two groups of sounds. This inquiry into the perception of place of articulation in these two languages has therefore failed to provide any evidence for a strictly perceptual motivation for the cross-linguistic susceptibility of nasals to undergo place assimilation. Chapters 5 and 6 describe efforts to uncover experimental evidence showing that the true motivation for this cross-linguistic tendency has its origins in constraints on articulation.

CHAPTER 5

DO STOPS AND NASALS HAVE RELEASE BURSTS WHEN FOLLOWED BY ANOTHER STOP?

The two previous studies have shown that audible release bursts make stops more perceptually distinct than nasals in English and Dutch. There are no significant perceptual differences, however, between nasals and stops without bursts. This effect holds for stops in C_1 position in VC_1C_2V sequences as well as in VC syllables, indicating that the perceptual advantage provided by release bursts to stops may make them less likely to undergo place assimilation than nasals.

Researchers have traditionally assumed, however, that stops (and nasals) generally lack release bursts when they are produced before other stops in a consonant cluster. Jun (1995), for instance, claims that, "In the consonant cluster C_1C_2, if C_1 is a stop, it is unreleased due to its overlap with C_2 in many languages (most languages surveyed in the present study)." Jun's sentiment echoes the received wisdom of the phonetics world; authoritative works such as Ladefoged (1975), MacKay (1978), Abercombie (1967), Catford (1977) and Jones (1956) all make the same claim. Jun (1995) therefore based his rankings of perceptually-based preservation constraints for C_1 consonants in C_1C_2 clusters on the assumption that consonants in this position do not audible release burst cues. Ohala (1990) made the same assumption when constructing stimuli for his study on perceptual influences on place assimilation, and Winters (2001,

2002) did the same. All of these studies found no significant differences between listeners' ability to perceive place in nasals and stops in a likely context for place assimilation to occur. The AX discrimination experiment described in the previous chapter yielded an identical non-result. This collection of experimental evidence suggests that there can be no perceptual motivation for nasals to undergo place assimilation more often than stops, so long as stops lack release bursts in environments which are likely to induce place assimilation.

Henderson & Repp (1982), however, produced evidence which showed that stops do not always lack release bursts when they are produced before another stop. After analyzing consonant cluster productions in previous, unrelated studies (e.g., Repp 1982), Henderson & Repp (1982) had become skeptical about claims that initial stops in C_1C_2 clusters were always "unreleased." Henderson & Repp realized that it would, in fact, be impossible for all such stops to be articulatorily unreleased. Further, they had also noticed that many such stops actually were produced with audible release bursts.

Spurred on by this evidence—and impeccable reasoning—Henderson & Repp designed a study to determine just how often stops in such clusters had these release bursts, and if these bursts were at all perceptible. Henderson & Repp's (1982) study simply involved having speakers read sentences which had consonant clusters embedded in them. These consonant clusters were generally situated across the morpheme boundary in a compound noun; for instance, the study included the compound "rib cage," which contained the target /bk/ cluster, in the following sentence:

(5.1) Considering he wasn't wearing a seatbelt, the driver was lucky to escape with only bruising to his ribcage and abdomen.

Henderson & Repp made an oscillographic analysis of each consonant cluster production in these sentences in order to determine whether or not the initial consonant had an audible release burst. In many cases, they found that it did; the exact percentages from their results are given below:

	Back-to-front clusters			Front-to-back clusters			
Speaker	/tp/	/kp/	/kt/	/pt/	/tk/	/pk/	mean
NM	25.0	25.0	50.0	62.5	75.0	87.5	45.0
AB	0.0	0.0	50.0	87.5	87.5	87.5	52.1
BR	0.0	0.0	37.5	87.5	87.5	100.0	52.1
CG	12.5	12.5	75.0	75.0	75.0	87.5	56.3
JM	0.0	25.0	87.5	87.5	87.5	75.0	60.4
RK	12.5	87.5	100.0	87.5	100.0	100.0	81.3
mean	8.3	25.0	66.7	81.3	85.4	89.6	57.9

Table 5.1: Percentages of C_1 stops with audible release bursts, by C_1C_2 cluster type, in Henderson & Repp (1982)

As Henderson & Repp noted, the likelihood that any of the C_1 stops has a release burst increases if the place of articulation of the C_2 consonant is further back in the oral tract from the lips. Overall, however, Henderson & Repp had established that stops actually had release bursts before other stops more than 50% of the time.

Since they found that so many stops have release bursts before other stops, Henderson & Repp questioned why many phoneticians (Ladefoged, Abercrombie, Jones, etc.) would assume that these release bursts did not exist. They hypothesized that, even though these bursts appear in many consonant clusters, listeners (and even trained phoneticians) cannot perceive their presence in the acoustic signal. Henderson & Repp tested this hypothesis in a two-part perception experiment. The stimuli for this experiment consisted of the productions of some of the target sequences which had been recorded in Henderson & Repp's production study. These target sequences included "cactus," "ribcage," "Edgar," "bodkin" and "scapegoat." Half of these stimuli had the release bursts following the first stop in the consonant cluster removed; the other stimuli were left as they were in the original recordings. In the first part of the experiment, all of the stimuli were presented to trained phoneticians, who were asked to simply determine whether or not there was a release burst in the consonant cluster. The second part of the experiment had a two-interval forced-choice task, in which the same listeners heard two stimuli and had to determine which one contained the consonant cluster with a release burst in it. Henderson & Repp found that listener performance in these tasks was not much better than chance:

Stimuli	Yes/No	2IFC
cactus	49.4	55.0
ribcage	61.1	58.3
Edgar	62.2	63.9
bodkin	64.4	62.2
scapegoat	67.8	80.5
total	61.0	64.0

Table 5.2: Percent correct release burst detection in Henderson & Repp's (1982) two perceptual tasks

Henderson & Repp (1982) therefore concluded that, even though stops in consonant clusters may have audible release bursts, these bursts are not generally perceptible.

Despite the inability of Henderson & Repp's listeners to consistently perceive the presence of these release bursts in the acoustic signal, the data from the AX discrimination experiment in chapter 4 had indicated that the presence of such release bursts does increase the perceptual salience of place contrasts in stops. It therefore seems plausible that the proportion of stops with release bursts in the actual production of stop-stop clusters may provide these sounds with enough perceptual strength to resist assimilatory processes more often than nasals. Shin (2000), in fact, suggested that release bursts play a key role in the failure of stops to undergo place assimilation in languages like Hindi. "In Hindi...stops maintain their release in the coda position (Rhee 1999). Whereas unreleased stops in the coda undergo assimilation due to the weak perceptual cues in English, coronal oral stops in the coda are released and thus do not undergo assimilation in Hindi."

The following replication of Henderson & Repp's (1982) original study was undertaken in order to test the hypothesis that the proportion of stops in consonant clusters with audible release bursts provides them with a perceptual advantage over nasals that makes them more resistant to place assimilation. This study also tests Shin's (2000) conjecture that stops always have release bursts in languages like Hindi—and Dutch—in which these segments do not undergo place assimilation. This study therefore expands on Henderson & Repp's original approach by investigating the production of both stop-stop and nasal-stop clusters, embedded in compound nouns in Dutch and English sentences. The proportions of nasals and stops in these clusters with audible release bursts will later be used to re-interpret the results of the two previous perception experiments, in order to determine the overall effect these cues might have on the perception of nasals and stops in real-world consonant cluster productions.

5.1 Replication of Henderson & Repp (1982): Materials

The sentences that speakers read in the original Henderson & Repp (1982) study contained consonant clusters which appeared both word-finally (e.g., 'apt') and within words. Intra-word clusters also appeared both across morpheme boundaries (e.g., 'ribcage') and within morphemes (e.g., 'bodkin'). In order to eliminate potential confounds in comparing productions of consonant clusters in such different environments, this study aimed for more uniformity in the structure of the words containing the target consonant cluster sequences. All of these words—in both English and Dutch—were therefore bisyllabic compound nouns (á la 'ribcage'), with the target consonant cluster straddling the morpheme boundary in the compound. Stress always fell on the first syllable of the compound. Compounds were selected for all nine place of articulation combinations in these clusters. The first consonant could therefore have any of three places of articulation (labial, coronal, dorsal), and it could be followed by a stop with any of the same three places of articulation. The first consonant in the cluster could either be a nasal or a voiceless stop; the stimuli therefore included eighteen basic consonant cluster types in each language. Two target compounds were selected for each of these consonant cluster types; all of the target compounds (for both languages) are listed below in Tables 5.3-5.4.

pp		mp	
dro**p** **p**ass	(5)	crea**m** **p**uff	(10)
slo**p** **p**ail	(10)	ti**mep**iece	(25)
pt		mt	
pu**p** **t**ent	(17)	thu**mb** **t**acks	(1)
la**p** **t**op	(22)	Drea**m** **T**eam	(12)
pk		mk	
cu**pc**akes	(16)	ti**me** **c**ard	(1)
po**pc**orn	(20)	Ga**mec**ocks	(2)
tp		np	
po**t** **p**ies	(19)	pe**n** **p**al	(4)
no**te** **p**ad	(24)	bea**n** **p**ole	(13)
tt		nt	
ou**tt**akes	(7)	pi**ne** **t**ar	(3)
ca**tt**ails	(8)	su**n** **t**an	(11)
tk		nk	
stree**tc**ar	(18)	pho**ne** **c**all	(4)
boo**t** **c**amp	(19)	gree**n** **c**ard	(11)
kp		ŋp	
cro**ckp**ot	(9)	Ki**ngp**in	(14)
ba**ckp**ack	(17)	pi**ng** **p**ong	(15)
kt		ŋt	
bla**ckt**op	(21)	ri**ng** **t**oss	(9)
du**ckt**ail	(23)	ha**ng** **t**ime	(12)
kk		ŋk	
sto**ck** **c**ar	(6)	spri**ng** **c**atch	(6)
du**ck** **c**all	(8)	lo**ng** **c**oat	(7)

Table 5.3: Target English compounds for the consonant cluster production study

pp		mp	
knoo**pp**unt - junction	(3)	zwe**mp**ak – swimsuit	(8)
slaa**pp**il – sleeping pill	(16)	bloe**mp**ot – flowerpot	(12)
pt		mt	
loo**pt**ijd – duration	(4)	boo**mt**ak – tree branch	(11)
slee**pt**ouw – tow-rope	(11)	kle**mt**oon – stress, emphasis	(20)
pk		mk	
zee**pk**ist – soap box	(6)	do**mk**op - blockhead	(6)
soe**pk**om – soup bowl	(12)	bloe**mk**ool - cauliflower	(13)
tp		np	
zou**tp**an – saltpan	(7)	loo**np**eil – level of wages	(15)
voe**tp**ad – footpath	(17)	sei**np**aal – signal-post	(19)
tt		nt	
boo**tt**ocht –boat excursion	(1)	tie**nt**al - decade	(4)
flui**tt**oon – flute tone	(2)	stee**nt**ijd – stone age	(5)
tk		nk	
straa**tk**ant –side of a street	(9)	klei**nk**ind – grandchild	(1)
scha**tk**ist – pushcart	(18)	wij**nk**aart – wine list	(14)
kp		ŋp	
koo**kp**unt – boiling point	(2)	spri**ngp**aard – jumping horse	(10)
dru**kp**ers – printing press	(10)	ga**ngp**ad – path	(14)
kt		ŋt	
hoe**kt**and – canine tooth	(5)	le**ngt**e - length	(3)
tre**kt**ocht – hike	(7)	spri**ngt**ouw – jump rope	(17)
kk		ŋk	
lij**kk**ist - coffin	(9)	ha**ngk**ast – hanging wardrobe	(8)
koo**kk**unst – art of cooking	(13)	za**ngk**oor – choir	(15)

Table 5.4: Target Dutch compounds for the consonant cluster production study

The number in parentheses after each compound in Tables 5.3-5.4 represents the number of the sentence in which they appeared in the experiment. All of these sentences, for both languages, are listed below in Tables 5.5-5.6. The author wrote the English sentences in Table 5.5; in general, these sentences included either one or two compounds, with one compound usually appearing at the end of the sentence. The target compounds in each sentence appear in boldface. A native speaker of Dutch wrote corresponding sentences for the Dutch compounds, along the same guidelines. All of these sentences are also listed in Table 5.6, with the author's English translations of each.

1. Josh dropped the box of **thumb tacks** when he realized that he had forgotten his **time card**.

2. This is the second year in a row that the Buckeyes have lost to the South Carolina **Gamecocks** in a bowl game.

3. Few baseball fans will ever forget the day that too much **pine tar** cost George Brett a home run.

4. Anna was delighted to receive a **phone call** from her **pen pal**.

5. The coach almost dropped his whistle when he first saw Gretzky's legendary **drop pass**.

6. Dale raised a few eyebrows when he installed a **spring catch** on the door of his **stock car**.

7. In one of the **outtakes** to the original Star Wars, you can see Han decked out in a **long coat** with a patch over one eye.

8. Hidden among the **cattails**, Sean let out two blasts on his **duck call**.

9. Jim claimed to have won his **crockpot** while playing **ring toss** at the county fair.

10. My Grandma threw the **cream puff** directly into the **slop pail**.

11. Though she still lacked a **green card**, Octavia had developed an incredible **sun tan**.

12. Michael Jordan wasn't the only member of the **Dream Team** with an unusually large **hang time**.

13. Randy assured us that he'd rather be called "**bean pole**" than "chrome dome."

14. Recent laboratory tests show that you can significantly reduce your brainpower by repeatedly watching the movie "**Kingpin**."

15. Once again, Jeff began to tell us the story of how he beat his prom date at **ping pong**.

16. Nancy got a kick out of the recipe for sugar-free **cupcakes** in her husband's new cookbook.

17. Dan was somehow able to stuff an entire **pup tent** into his **backpack**.

Table 5.5: English sentences used in the consonant cluster production study

(Table 5.5 continued)

18. Blanche could hear the **streetcar** from the porch behind her back door.

19. The new recruits ate only oatmeal and **pot pies** at **boot camp**.

20. We decided to enjoy the dog days of summer by holing up in our apartment and watching movies while we ate **popcorn**.

21. When we drove through New Jersey, we saw a surprising amount of roadkill on the **blacktop**.

22. There is actually a section in the lab guide on how to run experiments with the department **lap top**.

23. Joey looked at his reflection in one of the hubcaps of his car as he combed his hair into a **ducktail**.

24. Yesterday I wasted half an hour just trying to balance a **note pad** on top of a bed post.

25. Robert had inherited the antique **timepiece** from his grandfather.

Table 5.5: English sentences used in the consonant cluster production study

1. Mieke besloot haar **kleinkind** te verrassen met een **boottocht** over de Rijn
Mieke decided to surprise her grandchild with a boat tour over the Rhine.

2. Een schrille **fluittoon** uit de ketel geeft aan dat water het **kookpunt** heeft bereikt.
A shrill flute tone tells us that water has reached its boiling point.

3. Een kettingbotsing ter hoogte van het **knooppunt** Eemnes veroorzaakte een file met een **lengte** van meer dan 20 kilometer.
A chain collision off Eemnes caused a traffic jam with a length of over 20 kilometers.

4. Een **tiental** jaren was begroot voor de **looptijd** van een onderzoek naar buitenaards leven op de planeet mars.
A duration of a decade was estimated for research about alien life on the planet Mars.

5. De leraar vertelde tijdens de geschiedenisles dat het in de **steentijd** heel normaal was om zonder verdoving een **hoektand** te trekken.
The teacher explained during the history class that it was very normal in the Stone Age to pull a canine tooth without an anaesthetic.

6. Wat een **domkop** was Peter om over de **zeepkist** uit glijden en zijn arm te breken.
What a blockhead was Peter to slip over the soap box and break his arm.

7. Jan en Marieke zijn van plan om tijdens hun vakantie in de Verenigde Staten een **trektocht** door een **zoutpan** te maken.
Jan and Marieke are planning to take a hike across a saltpan during their vacation in the US.

8. Voordat Bianca naar het zwembad kon gaan, moest ze eerst op zoek naar haar **zwempak**, dat ergens in haar **hangkast** verborgen lag.
Before Bianca could go to the swimming pool, she first needed to find her swimsuit, which was hidden somewhere in her hanging wardrobe.

9. Een enorme menigte had zich verzameld langs de **straatkant** om de **lijkkist** van de koningin te zien passeren en een laatste afscheid te kunnen nemen.
An enormous crowd has gathered along the side of the street to see the coffin of the queen passing by, and to say a last goodbye.

10. Het **springpaard** schrok van het onverwachte geluid van een **drukpers** en sloeg op hol.
The jumping horse was frightened by the unexpected sound of a printing press and bolted.

Table 5.6: Dutch sentences used in the consonant cluster production study

(Table 5.6 continued)

11. Thijs had een **sleeptouw** nodig om de **boomtak**, die de weg blokkeerde, in de berm te trekken.
Thijs needed a tow rope to pull the bough, that was blocking the road, in the verge.

12. Bij gebrek aan een **bloempot** besloot Anne een **soepkom** te gebruiken voor haar nieuwe plant.
Because Anne didn't have a flowerpot she decided to use a soup bowl for her new plant.

13. Petra probeerde haar echtgenoot ervan te overtuigen dat hij best een keer **bloemkool** zou kunnen klaar maken, aangezien er maar weinig **kookkunst** voor nodig is.
Petra tried to convince her husband to make her cauliflower at least once, since it doesn't involve much "art of cooking".

14. Omdat er niet direct een tafel beschikbaar was in het restaurant, besloten Anne en Peter in het **gangpad** te wachten, waar ze wel alvast een blik konden werpen op de **wijnkaart** en het menu.
Because there wasn't a table available in the restaurant right away, Anne and Peter decided to wait in the path, where they could have a first look on the winelist and the menu.

15. Om een verhoging van het **loonpeil** af te dwingen besloot het **zangkoor** te gaan staken.
To force a raise of the level of wages the choir decided to go on a strike.

16. Omdat Hans moeite had om de slaap te vatten, besloot hij om een **slaappil** te nemen.
Because Hans had trouble sleeping, he decided to take a sleeping pill.

17. De kinderen waren druk in de weer op het **voetpad** met hun **springtouw**.
The children were busy playing with their jump rope on the curb.

18. De **schatkist** werd pas vele jaren later ontdekt in een diepe grot in het zuiden van Limburg.
The treasury was only discovered many years later in a deep cave in the southern part of Limburg.

19. Na een zware nachtdienst viel Wouter, op weg naar huis, in slaap en reed met zijn auto tegen een **seinpaal** aan.
On his way home, after an exhausting night shift, Wouter fell asleep and drove his car into a signal-post.

(Table 5.6 continued)

20. Er zijn regionale verschillen in Nederland met betrekking tot de **klemtoon** van een woord.
There are regional differences in the Netherlands with regard to the particular emphasis on a word.

Table 5.6: Dutch sentences used in the consonant cluster production study

5.2 Methods

The participants in this study read each of these sentences (in their native language only) as they were presented to them by a customized SuperCard (version 2.0) program running on a Macintosh PowerPC 6500/275 in a sound-attenuated booth. Figure 5.1 shows the format of the prompt this program gave the participants for each sentence. The participants read each sentence through a Sennheiser HMD 25-L head-mounted microphone, which was connected to the computer via a Symetrix SX 30L Dual microphone pre-amplifier, with the gain set at approximately 20 dB. Through this setup, each reading of each sentence was recorded directly onto the Macintosh's hard drive by the customized SuperCard program.

After each presentation of each sentence, the SuperCard program recorded the input from the head-mounted microphone for a total of ten seconds. The participants were instructed to produce the entire sentence within this ten-second window. Once the program had finished recording, it played back the recording to the listener through the headphones on the head-mounted microphone/headphone set. The program then gave the participants the option of recording the sentence again, in case they had made any mistakes while reading the sentence. The participants had the option of re-recording each sentence as many times as necessary.

The participants in this study were the same Dutch and English speakers who participated in the magnitude estimation study which was described in detail in chapter 3. They worked through this consonant cluster production of the study immediately before the magnitude estimation task. The Dutch speakers were recruited from the community of Dutch speakers at the Ohio State University and the Columbus, Ohio, Dutch club. These speakers were paid $8 each for their participation in both this task and the magnitude estimation study. The English-speaking participants were recruited from introductory linguistics and introductory psycholinguistics courses at the Ohio State University; they received course credit for their participation.

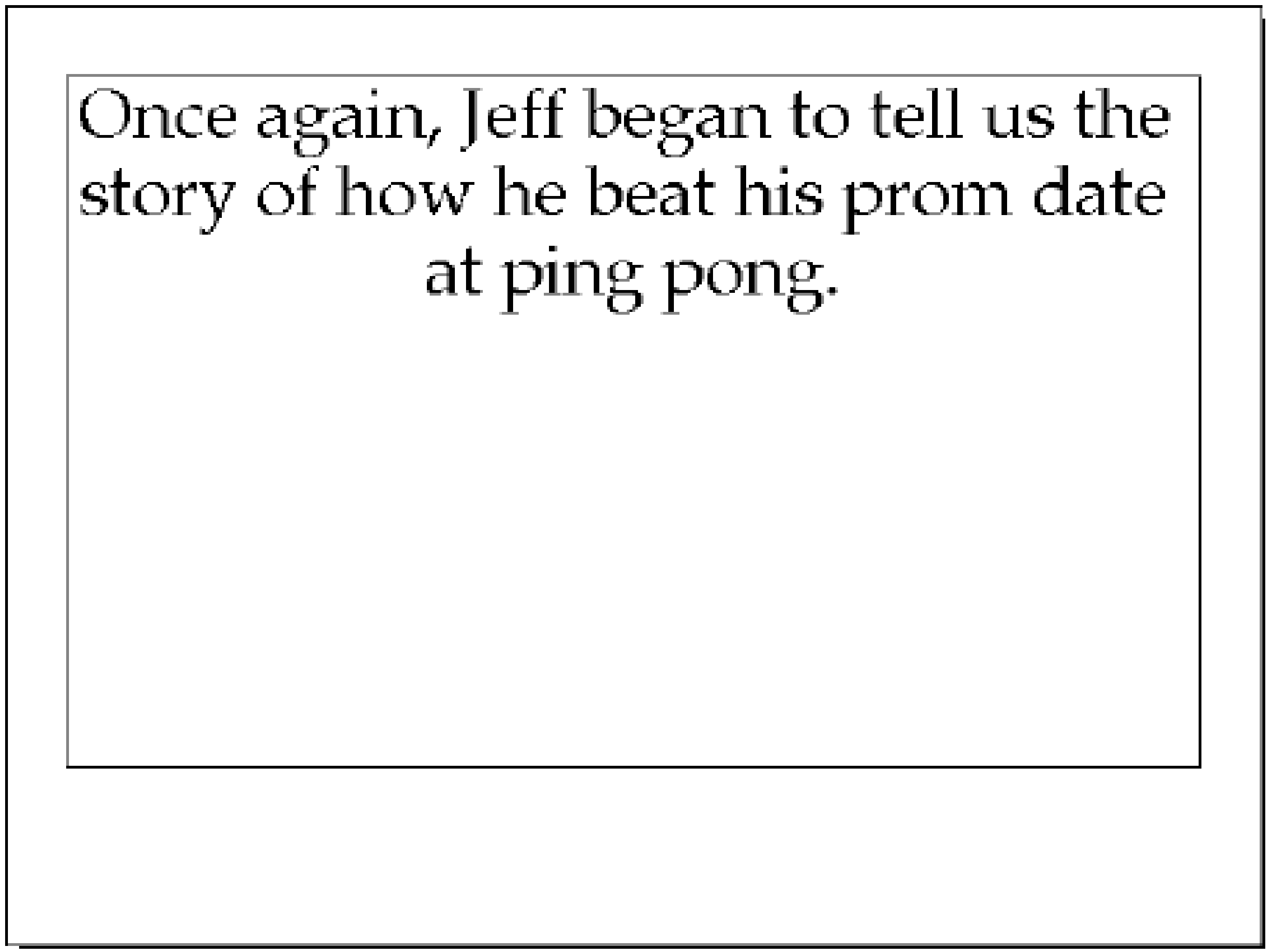


Figure 5.1: Example sentence prompt for consonant cluster production study

5.3 Analysis

The Praat (version 4.0) sound editor was used to determine whether or not the initial stops and nasals in each target consonant cluster had an audible release burst. Figures 5.2-5.4 present examples of the waveform and spectrogram representations in Praat that were used to help guide this analysis process. Figure 5.2 shows an example production of the Dutch word 'trektocht' which has an audible release burst, while Figure 5.3 shows a production of the same word which lacks an audible release burst. Similarly, Figure 5.4 shows a production of the English compound "green card" with an audible release burst after the final /n/ in "green". Table 5.7 shows the results of this analysis, listing the tallies of audible release bursts for each consonant cluster type and manner of articulation, broken down by the two groups of language speakers. Table 5.8 shows this same information in terms of the percentage of all consonant cluster productions which had audible release bursts after the first consonant.

Only <u>audible</u> release bursts were included in these tallies, since it was presumed that only these bursts might have an effect on perception. Several stop and nasal productions seemed to have release bursts which were inaudible; Figure 5.5 presents an example of one of these "hidden" release bursts in the Dutch word "voetpad". Such hidden release bursts were occasionally found after consonants which were followed by a stop whose place of articulation was further forward in the oral tract—in a /t/ or a /k/ before a /p/, for instance. It seems reasonable to assume that these "hidden" release bursts may indicate articulatory overlap in these sequences—where the second consonantal closure has been made before the release of the first. Hence, the release burst of the first consonant is damped by the closed oral cavity into which it is released; such a damped burst may be what appears as a low frequency spike in the spectrogram in Figure 5.5. Such "hidden" release bursts were not counted as audible release bursts, nor included in the tallies in Table 5.7. However, it is worth nothing that such gestural overlap is more likely to occur in consonant clusters which proceed from a further back place of articulation to a further forward place of articulation. This fact may explain why the results of both this study and Henderson & Repp (1982) show that stops preceding dorsals and alveolars are more likely to have audible release bursts.

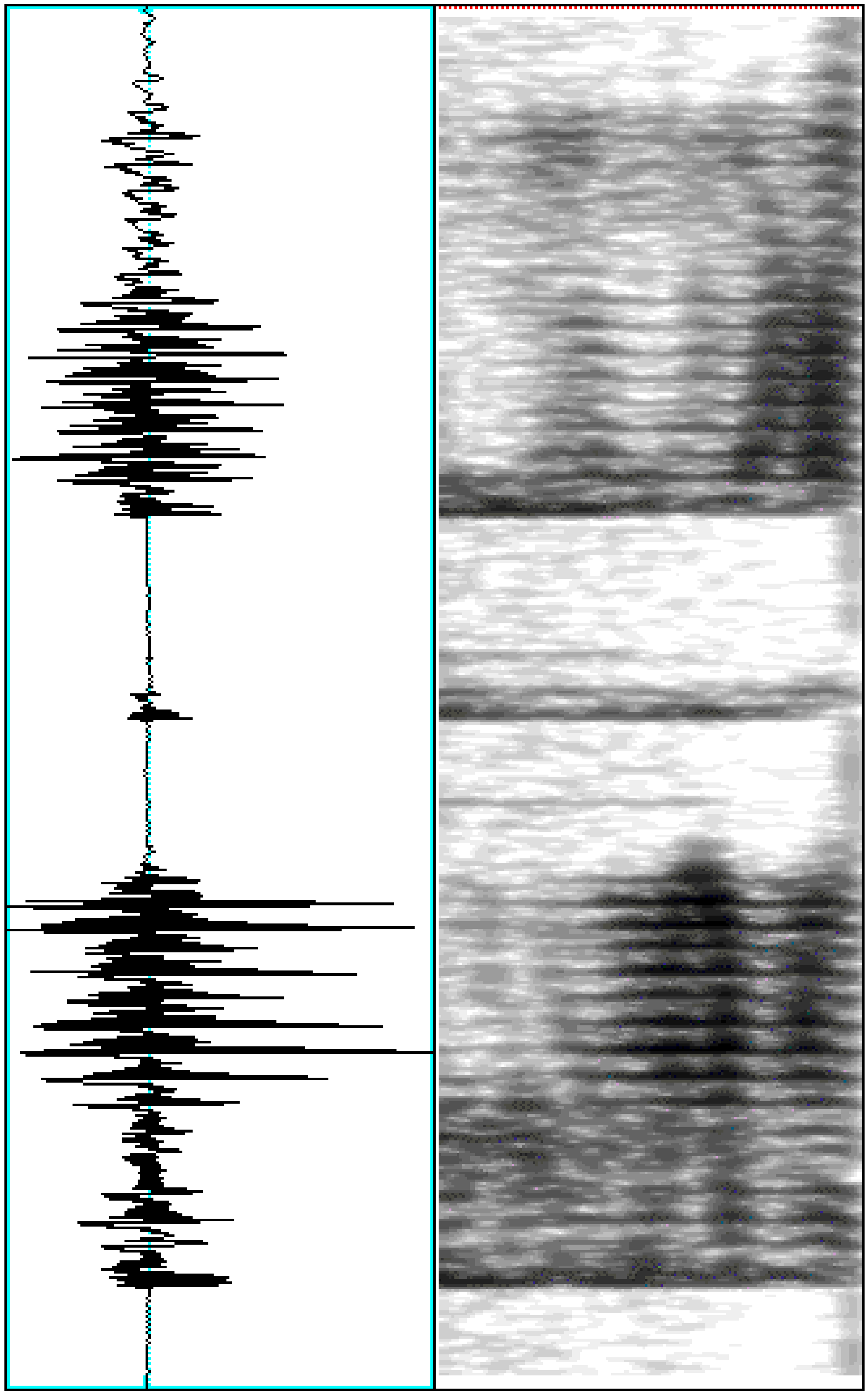


Figure 5.2: Example waveform and spectrogram of “trektocht” production, with burst

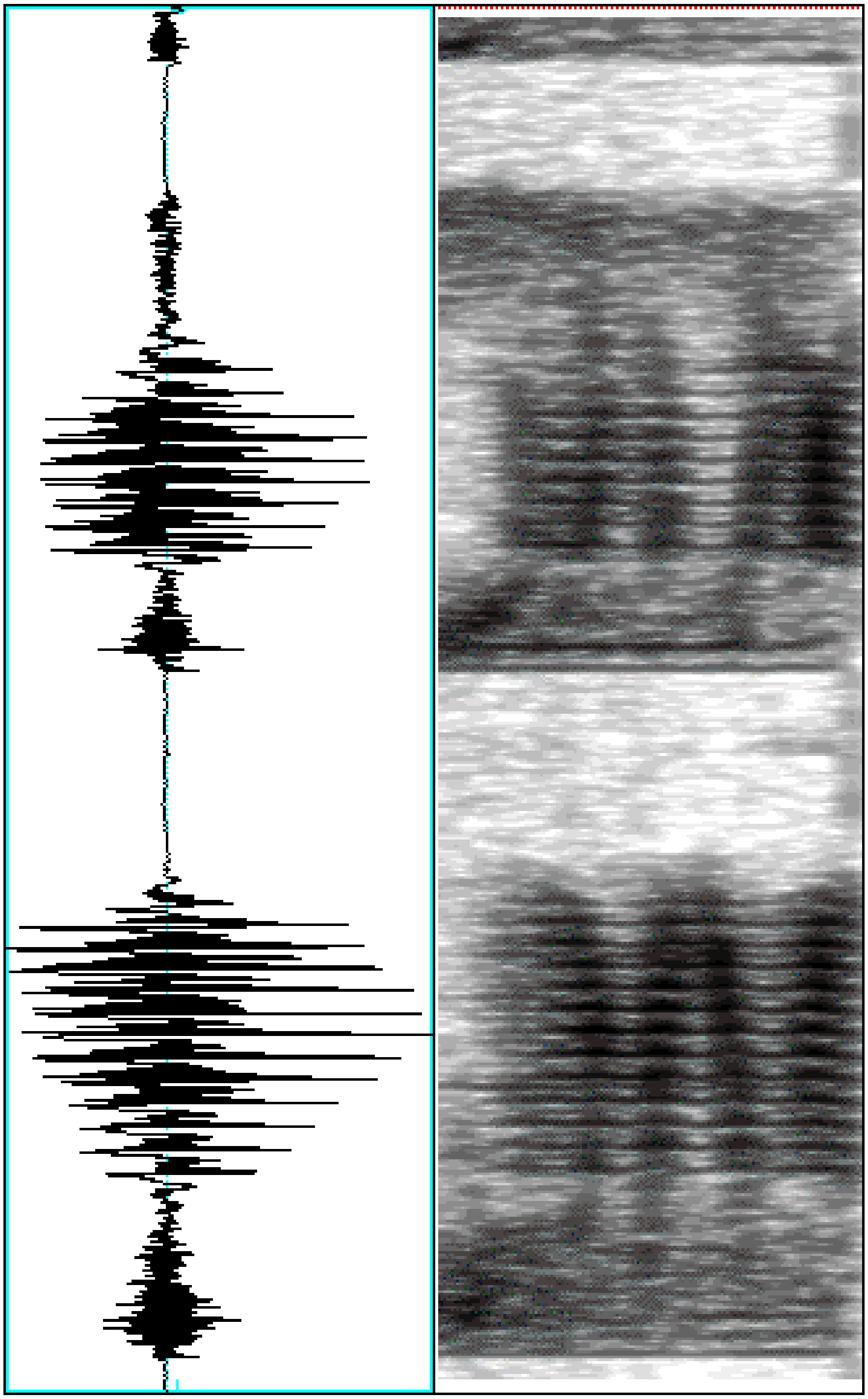


Figure 5.3: Example waveform and spectrogram of “trektocht” production, without burst

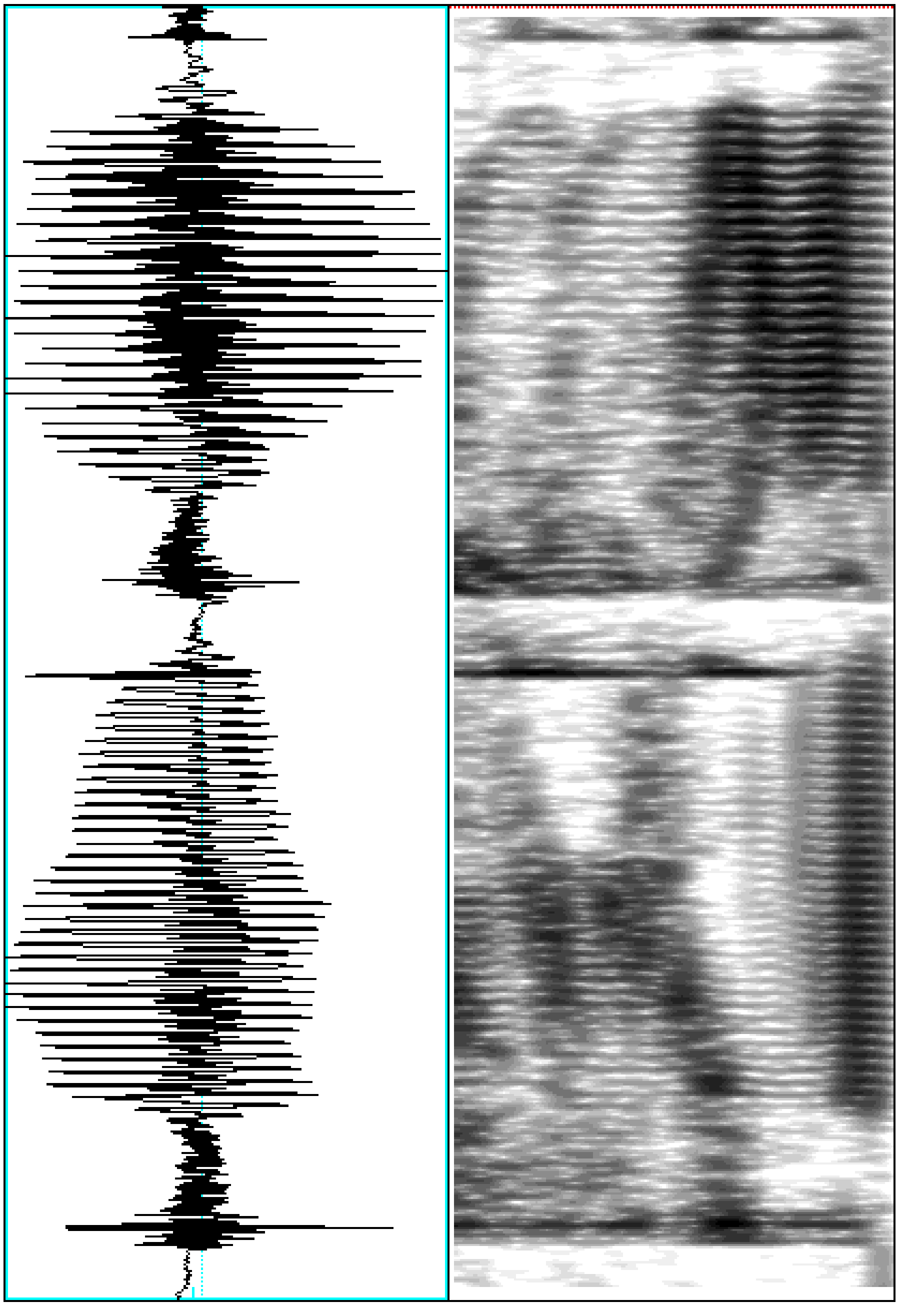


Figure 5.4: Example waveform and spectrogram of “green card” production, with burst

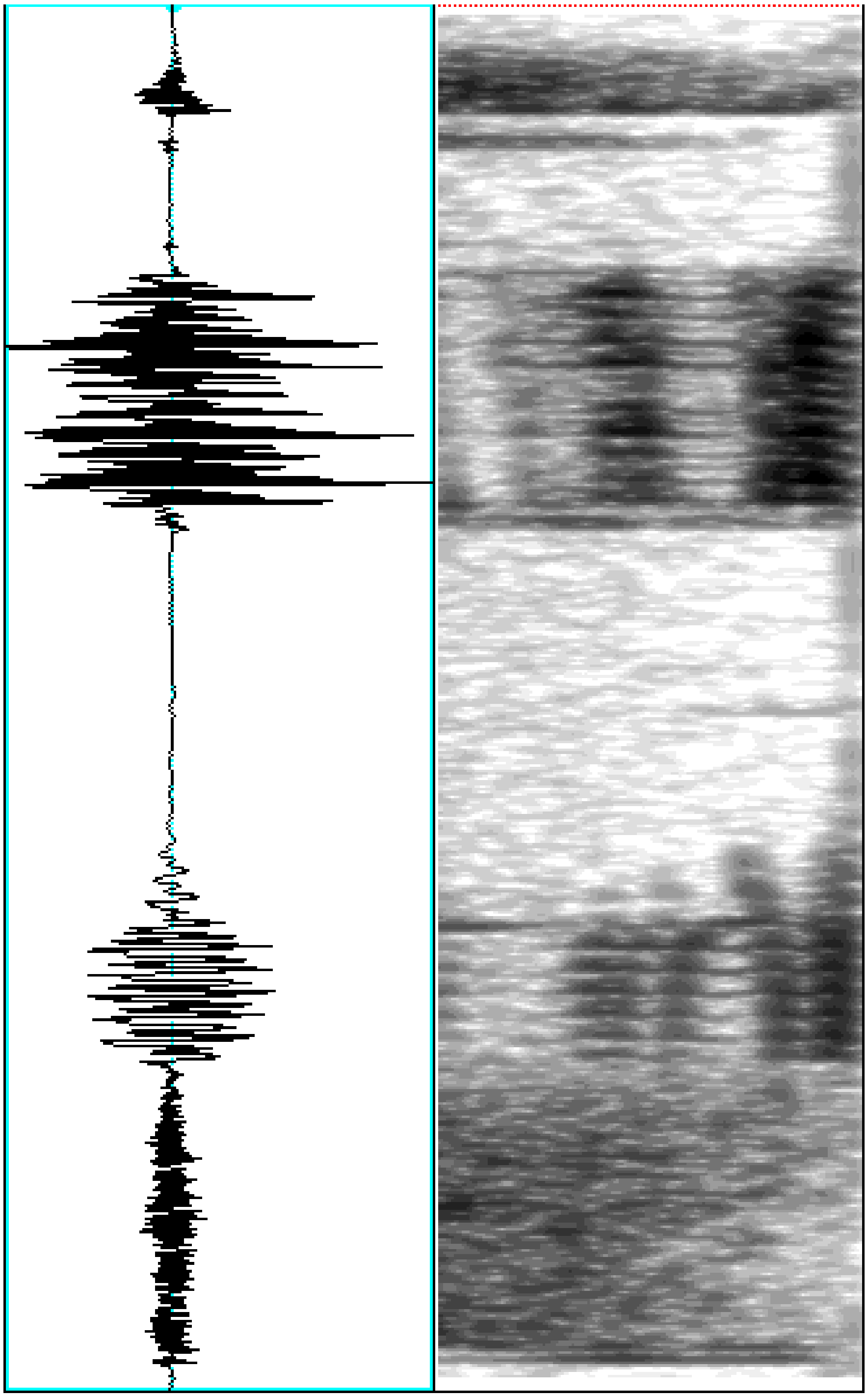


Figure 5.5: Example waveform and spectrogram of "voetpad" production, with "hidden" burst

English Speakers

		C2 =			
	Stops	p	t	k	Total
	p	3	15	14	32
C1 =	t	0	7	10	17
	k	8	11	5	24
	Total	11	33	29	73

	Nasals	p	t	k	Total
	m	0	8	9	17
C1 =	n	0	3	7	10
	ŋ	2	3	3	8
	Total	2	14	19	35

Dutch Speakers

		C2 =			
	Stops	p	t	k	Total
	p	2	15	11	28
C1 =	t	11	4	14	29
	k	8	14	2	24
	Total	21	33	27	81

	Nasals	p	t	k	Total
	m	0	7	8	15
C1 =	n	0	2	6	8
	ŋ	2	4	0	6
	Total	2	13	14	29

Table 5.7: Tallies of audible release bursts in consonant clusters, by cluster type and language of speaker

English Speakers

		C2 =			
	Stops	p	t	k	Mean
	p	18.8%	93.8%	87.5%	66.7%
C1 =	t	0.0%	43.8%	62.5%	35.4%
	k	50.0%	68.8%	31.3%	50.0%
	Mean	22.9%	68.8%	60.4%	50.7%

	Nasals	p	t	k	Mean
	m	0.0%	50.0%	56.3%	35.4%
C1 =	n	0.0%	18.8%	43.8%	20.8%
	ŋ	12.5%	18.8%	18.8%	16.7%
	Mean	4.2%	29.2%	39.6%	24.3%

Dutch Speakers

		C2 =			
	Stops	p	t	k	Mean
	p	12.5%	93.8%	68.8%	58.3%
C1 =	t	68.8%	25.0%	87.5%	60.4%
	k	50.0%	87.5%	12.5%	50.0%
	Mean	43.8%	68.8%	56.3%	56.3%

	Nasals	p	t	k	Mean
	m	0.0%	43.8%	50.0%	31.3%
C1 =	n	0.0%	12.5%	37.5%	16.7%
	ŋ	12.5%	25.0%	0.0%	12.5%
	Mean	4.2%	27.1%	29.2%	20.1%

Table 5.8: Percentages of audible release bursts in consonant clusters, by cluster type and language of speaker

The Praat (version 4.0) sound editor was also used to measure the durations of each consonant cluster in the target compounds. Consonant cluster durations were measured from the offset of the vowel directly preceding the cluster to the onset of the release burst for the second stop in the cluster. This interval included both the murmur portion of nasal consonants and any audible release bursts. The durations of both these murmurs and release bursts were also made in order to guide the construction of the VC-CV sequences for the AX discrimination experiment. Tables 5.9 through 5.12 list the results of these duration measurements.

Dutch Speakers

	Stops	C2 = p	t	k	Mean
C1 =	p	120.2	119.9	136.6	125.6
	t	135.7	108.2	128.8	124.2
	k	152.2	140.2	97.7	130.0
	Mean	136.0	122.8	121.0	126.6

	Nasals	C2 = p	t	k	Mean
C1 =	m	126.6	134.3	140.9	133.9
	n	132.4	109.3	100.5	114.1
	ŋ	136.0	116.9	120.6	124.5
	Mean	131.7	120.2	120.7	124.2

English Speakers

	Stops	C2 = p	t	k	Mean
C1 =	p	135.6	135.3	116.7	129.2
	t	87.3	81.4	85.9	84.8
	k	129.4	125.0	117.1	123.8
	Mean	117.4	113.9	106.5	112.6

	Nasals	C2 = p	t	k	Mean
C1 =	m	103.3	116.7	109.9	110.0
	n	131.7	112.4	101.3	115.1
	ŋ	140.9	137.4	121.3	133.2
	Mean	125.3	122.2	110.8	119.4

Table 5.9: Mean consonant cluster durations, by place and manner cluster

Dutch Speakers

		C2 =			
		p	t	k	Mean
	m	71.7	87.3	82.0	80.5
C1 =	n	63.9	68.8	54.8	62.5
	ŋ	70.7	68.1	73.1	70.6
	Mean	68.7	74.7	69.7	71.1

English Speakers

		C2 =			
		p	t	k	Mean
	m	69.1	89.1	76.3	78.1
C1 =	n	101.3	83.1	72.9	85.7
	ŋ	99.8	101.2	91.7	97.7
	Mean	89.8	91.1	80.2	87.1

Table 5.10: Mean nasal murmur durations, by place cluster

Dutch Speakers

		C2 =			
		p	t	k	Mean
	p	76.0	70.1	79.8	74.0
C1 =	t	69.8	72.0	57.4	62.8
	k	65.3	59.7	60.7	61.6
	Mean	68.4	65.1	66.2	66.3

English Speakers

		C2 =			
		p	t	k	Mean
	p	107.0	90.2	65.9	80.3
C1 =	t	---	52.5	52.6	52.6
	k	59.8	58.5	66.0	60.8
	Mean	67.7	72.2	61.6	67.1

Table 5.11: Mean onset of first consonant release bursts, by place cluster

Dutch Speakers

		C2 =			
		p	t	k	Mean
	p	31.5	19.1	17.5	19.6
C1 =	t	26.4	61.0	30.4	30.3
	k	41.0	29.6	48.5	35.2
	Mean	33.4	25.5	27.2	28.2

English Speakers

		C2 =			
		p	t	k	Mean
	p	20.5	16.3	16.6	16.7
C1 =	t	---	12.2	11.1	11.5
	k	17.7	13.2	6.7	13.8
	Mean	18.3	14.6	13.6	14.7

Table 5.12: Mean release burst durations, by place cluster

Tables 5.13-5.16 present the results of tests for significant differences between the proportions of stops and nasals with audible release bursts in English and Dutch. Each of these tests involves calculating confidence intervals for the differences between proportions of different classes of sounds with audible release bursts. Table 5.13, for instance, presents confidence intervals for the difference in proportions of Dutch and English stops with audible release bursts. Since both confidence intervals in Table 5.13 encompass zero, it is impossible to say that there is a significant difference in the proportions of stops with audible release bursts between the two languages. Table 5.14 shows that the same holds true for Dutch and English nasals--neither language group is more likely to produce these sounds with audible release bursts than the other. Dutch speakers do not, therefore, always produce stops with release bursts—as Shin (2000) had hypothesized--even though stops do have audible release bursts in both Dutch and English more than 50% of the time. This result suggests that something besides the presence of audible release bursts after stops must motivate the resistance of these sounds to place assimilation in Dutch.

Stops	pp	pt	pk	tp	tt	tk	kp	kt	kk	Total	Pct.	Diff.		95%
English	3	15	14	0	7	10	8	11	5	73	0.51	-0.06	+/-	0.12
Dutch	2	15	11	11	4	14	8	14	2	81	0.56			
N =	16	16	16	16	16	16	16	16	16	144				

Table 5.13: Difference in proportions of stops with release bursts as produced in consonant clusters by speakers of English and Dutch

Nasals	mp	mt	mk	np	nt	nk	ŋp	ŋt	ŋk	Total	Pct.	Diff.		95%
English	0	8	9	0	3	7	2	3	3	35	0.24	0.04	+/-	0.1
Dutch	0	7	8	0	2	6	2	4	0	29	0.20			
N =	16	16	16	16	16	16	16	16	16	144				

Table 5.14: Difference in proportions of nasals with release bursts as produced in consonant clusters by speakers of English and Dutch

	pp	pt	pk	Total	Pct.	Diff.		95%
English	3	15	14	32	0.67	0.08	+/-	0.19
Dutch	2	15	11	28	0.58			
N =	16	16	16	48				

	tp	tt	tk	Total	Pct.	Diff.		95%
English	0	7	10	17	0.35	**-0.25**	**+/-**	**0.19**
Dutch	11	4	14	29	0.60			
N =	16	16	16	48				

	kp	kt	kk	Total	Pct.	Diff.		95%
English	8	11	5	24	0.50	0.00	+/-	0.20
Dutch	8	14	2	24	0.50			
N =	16	16	16	48				

	pp	tp	kp	Total	Pct.	Diff.		95%
English	3	0	8	11	0.23	**-0.21**	**+/-**	**0.18**
Dutch	2	11	8	21	0.44			
N =	16	16	16	48				

	pt	tt	kt	Total	Pct.	Diff.		95%
English	15	7	11	33	0.69	0.00	+/-	0.19
Dutch	15	4	14	33	0.69			
N =	16	16	16	48				

	pk	tk	kk	Total	Pct.	Diff.		95%
English	14	10	5	29	0.60	0.04	+/-	0.20
Dutch	11	14	2	27	0.56			
N =	16	16	16	48				

Table 5.15: Difference in proportions of stop-stop clusters with release bursts as produced by speakers of English and Dutch

English	LL	LC	LD	CL	CC	CD	DL	DC	DD	Total	Pct.	Diff.		95%
Stops	3	15	14	1	7	10	8	11	5	73	0.51	**0.26**	+/-	**0.11**
Nasals	0	8	9	1	3	7	2	3	3	35	0.24			
N =	16	16	16	16	16	16	16	16	16	144				

Dutch	LL	LC	LD	CL	CC	CD	DL	DC	DD	Total	Pct.	Diff.		95%
Stops	2	15	11	11	4	14	8	14	2	81	0.56	**0.36**	+/-	**0.10**
Nasals	0	7	8	0	2	6	2	4	0	29	0.20			
N =	16	16	16	16	16	16	16	16	16	144				

Table 5.16: Difference in proportions of nasals and stops with release bursts, as produced by speakers of English and Dutch

The data in Table 5.15, however, reveals one apparent difference between the production of stops in Dutch and English. This table lists burst proportions for the individual places of articulation in both the first second consonants in the cluster. It breaks down these proportions by the language of the speaker and the place of articulation of either the first or second consonant, and tests whether or not there are any significant differences between the burst proportions by language. In most cases, there are none. The production of coronal-labial stop-stop clusters in both languages, however, is an exception. Table 5.15 reveals that Dutch coronal stops have significantly more release bursts when produced before another stop than English coronal stops do. Likewise, Dutch stops have more release bursts when produced before a labial stop than English stops do. Both of these effects seem to be the result of the complete absence of release bursts in English speakers' productions of coronal-labial stop-stop clusters. The English speakers read these clusters in the English compounds "notepad" and "pot pie"; their counterparts in Dutch--"zoutpan" and "voetpad"--had release bursts in 11 out of 16 readings. The tendency for coronal stops in these clusters to lack release bursts may therefore provide some motivation for their tendency to undergo place assimilation in English, while the fact that the same stops in Dutch often have release bursts may explain why they do not undergo the same process.

The data in Table 5.16 also indicates that there may be a similar explanation for nasals' tendency to undergo place assimilation more often than stops, cross-linguistically. The tests in this table confirm that stops have release bursts more often in stop-stop clusters than nasals in nasal-stop clusters do. This effect holds for both Dutch and English speakers, and Table 5.17 shows that it is true for all C_1 and C_2 places of articulation, as well. The robustness of this effect suggests that release bursts may play a role in enabling stops to resist assimilatory processes more often than nasals do, due to the boost in perceptual salience that these particular cues to place of articulation provide.

	LL	LC	LD	Total	Pct.	Diff.		95%
Stops	5	30	25	60	0.625	**0.292**	+/-	**0.135**
Nasals	0	15	17	32	0.333			
N =	32	32	32	96				

	CL	CC	CD	Total	Pct.	Diff.		95%
Stops	11	11	24	46	0.479	**0.292**	+/-	**0.127**
Nasals	0	5	13	18	0.188			
N =	32	32	32	96				

	DL	DC	DD	Total	Pct.	Diff.		95%
Stops	16	25	7	48	0.500	**0.354**	+/-	**0.122**
Nasals	4	7	3	14	0.146			
N =	32	32	32	96				

Table 5.17: Difference in proportions of nasals and stops with release bursts, by place of articulation

One way to determine whether this hypothetical interaction between perceptual salience and release burst production exists is to re-interpret the data of the previous perceptual studies in terms of the appropriate percentages of stops and nasals with release bursts, as determined by the results of this study. Tables 5.18-5.21 present one attempt at such a re-interpretation of the perceptual data. Tables 5.18-5.19 combine the magnitude estimation data (from chapter 3) for place contrasts in the three different manners of articulation--stops with bursts, stops without bursts, and nasals--into the corresponding numbers for just two manners of articulation: stops and nasals. The data for the individual contrasts in this combined manner table is determined by a weighted average of the appropriate contrast estimations for that particular sound both with and without bursts. The weighting of this equation is determined by the percentage of the sounds in the contrast which had release bursts in the production study.

For example: the combined estimated magnitude value for the English stop coronal-dorsal contrast is determined by first calculating the percentage of English coronal and dorsal stops which had release bursts in consonant clusters in the production study. These numbers are 35.4% (17/48) and 50% (24/48), respectively. These two percentages are averaged to provide the relevant overall percentage of stops with bursts for the coronal-dorsal contrast in English: (17 + 24)/96 = 42.7%. The combined estimate magnitude score for the English coronal-dorsal contrast in stops, then, is weighted to include 42.7% of the z-score for this contrast with release bursts, and (1 – 42.7% =) 57.3% of the z-score for this contrast without release bursts. The estimated magnitudes given by each individual listener (Dutch and English) for the coronal-dorsal stop contrasts both with and without bursts are combined according to this percentage of release bursts. For instance, Dutch listener 1 had an average z-score of .373 for English coronal-dorsal stop contrasts with bursts, and an average z-score of -.884 for the same contrasts without bursts. This listener's combined z-score for this contrast--weighted by the percentage of release bursts—is, therefore, .427 x .373 + (1-.427) x (-.884) = -.3472. This calculation can be expressed mathematically as:

$$(5.2) \quad c_{ij} = \frac{(b_i + b_j)}{(n_i + n_j)} \cdot z(b_{ij}) + \left(1 - \frac{(b_i + b_j)}{(n_i + n_j)}\right) \cdot z(nb_{ij})$$

where c_{ij} represents the combined score for the place contrast between sounds i and j, b_{ij} represents the z-score for this place contrast only in stops with bursts, and nb_{ij} represents the same score for this score in stops without bursts. b_i represents the number of i sounds which had release bursts in the consonant cluster production study, while n_i represents the total number of i productions in this study (with or without release bursts).

This same equation was also used to calculate combined z-scores for place contrasts in nasal consonants, even though release bursts were significantly less frequent in these sounds than they were in stop consonants. The calculations for these combined nasal z-scores incorporated the z-scores for place contrasts between stops with release bursts, for the appropriate percentage of nasals which had such bursts in the production study. The decision to incorporate these numbers in the calculations for nasal contrasts presumes that place distinctions in nasals with bursts have the same perceptual distinctiveness as they do in stops with bursts. While it was impossible to obtain empirical data to confirm this presumption, it seems not unreasonable, given that place distinctions in stops and nasals without release bursts have roughly comparable levels of perceptual salience.

Tables 5.18-5.19 present data which has been combined in this way for the z-scores from the magnitude estimation experiment, broken down by manner of articulation (stops, nasals), place contrast (labial-coronal, labial-dorsal, coronal-dorsal), language of speaker (English, Dutch) and language of listener (English, Dutch). Tables 5.20-5.21 present the perceptual distance data from the AX discrimination experiment, which has been combined in the same way.

Manner	Stops			Nasals		
Contrast	Dor-Cor	Dor-Lab	Cor-Lab	Dor-Cor	Dor-Lab	Cor-Lab
English	-1.016	0.049	0.102	-0.809	-0.173	0.652
Listeners	-0.342	0.197	-1.592	-0.342	-0.006	-1.474
	-0.385	-0.292	0.116	0.847	0.154	-1.671
	0.322	-0.186	-0.747	-0.587	-0.349	-1.249
	-0.817	-0.737	-0.884	-0.113	-0.142	-1.417
	-1.002	0.347	-0.538	0.627	0.866	-2.567
	-0.210	0.028	-1.213	-1.032	-0.941	-1.348
	-0.725	-0.447	-0.327	-0.548	-1.133	-1.189
Dutch	-0.534	0.035	-0.845	-0.766	-0.394	-0.916
Listeners	0.065	-0.194	-0.781	-0.500	-0.699	-0.860
	-0.094	-0.421	-0.235	-1.008	-0.433	-1.363
	0.204	-0.293	-0.036	-0.218	0.090	-1.955
	-0.720	0.050	-0.516	-1.011	-0.262	-1.265
	-0.770	-0.329	-0.093	-0.456	0.544	-1.079
	0.063	0.444	-2.004	-0.752	0.019	-1.984
	-0.503	-0.170	-1.072	-0.608	-0.396	-1.317

Table 5.18: Burst-adjusted magnitude estimation data, for Dutch stimuli only

Manner	Stops			Nasals		
Contrast	Dor-Cor	Dor-Lab	Cor-Lab	Dor-Cor	Dor-Lab	Cor-Lab
English	-0.953	-0.476	-0.017	-0.674	-0.290	-0.648
Listeners	-0.239	0.146	-0.336	-0.502	-0.507	0.516
	-0.411	-0.015	0.816	-0.578	-0.835	-0.091
	-0.353	0.131	-0.416	-0.637	-0.410	-0.606
	-0.300	-0.531	0.324	-1.400	-1.091	-0.430
	-1.439	0.118	-0.426	-0.366	0.175	-0.118
	-0.657	0.014	0.575	-1.157	-0.535	-0.515
	-0.188	0.091	0.397	-1.428	-1.095	-0.640
Dutch	-0.347	-0.421	0.047	-0.648	-0.595	-0.557
Listeners	-0.372	0.457	-0.236	-0.611	-0.287	0.161
	-0.078	0.443	-0.469	-0.945	-0.606	-1.047
	-0.696	-0.439	-0.072	-0.789	-0.968	0.017
	-0.782	-0.485	-0.025	-0.986	-0.471	-0.457
	-0.169	-0.244	-0.249	0.133	-0.756	-0.560
	-0.647	0.300	0.039	-1.330	-0.804	-0.981
	-0.567	0.058	-0.248	-1.031	-0.567	-0.154

Table 5.19: Burst-adjusted magnitude estimation data, for English stimuli only

Manner	Stops			Nasals		
Contrast	Lab-Cor	Lab-Dor	Cor-Dor	Lab-Cor	Lab-Dor	Cor-Dor
English	0.047	0.087	0.063	0.075	0.095	0.091
Listeners	0.097	0.069	0.089	-0.025	0.088	0.082
	0.018	0.101	0.099	0.084	0.101	0.083
	0.103	0.093	0.100	-0.031	0.102	0.098
	0.058	0.099	0.093	0.093	0.098	0.094
	0.097	0.095	0.097	0.041	0.094	0.100
	-0.008	0.094	0.096	0.077	0.103	0.097
	0.093	0.065	0.100	-0.044	0.092	0.101
	0.025	0.081	0.088	0.090	0.099	0.099
	0.080	0.069	0.068	-0.046	0.090	0.090
Dutch	0.100	0.093	0.098	0.071	0.104	0.104
Listeners	-0.041	0.091	0.076	0.059	0.099	0.096
	0.081	0.087	0.092	0.009	0.095	0.097
	-0.028	0.098	0.093	0.064	0.099	0.092
	0.030	0.100	0.092	0.082	0.102	0.095
	0.079	0.090	0.087	0.023	0.092	0.067
	-0.020	0.093	0.078	0.067	0.094	0.090
	0.088	0.085	0.088	-0.001	0.090	0.092
	0.097	0.100	0.082	-0.031	0.099	0.102
	0.012	0.068	0.082	0.071	0.088	0.093

Table 5.20: Burst-adjusted perceptual distance values from AX discrimination experiment, Dutch stimuli only

Manner	Stops			Nasals		
Contrast	Lab-Cor	Lab-Dor	Cor-Dor	Lab-Cor	Lab-Dor	Cor-Dor
English	0.090	0.095	0.057	0.087	0.091	0.042
Listeners	0.059	0.056	0.041	0.021	0.034	-0.037
	0.087	0.096	0.032	0.093	0.084	0.019
	0.049	0.072	0.034	0.081	0.069	0.027
	0.098	0.098	0.084	0.098	0.098	0.085
	0.091	0.096	0.079	0.099	0.077	0.097
	0.096	0.100	0.099	0.100	0.097	0.090
	0.055	0.091	0.049	0.082	0.089	0.053
	0.098	0.099	0.083	0.099	0.099	0.072
	0.045	0.076	0.033	0.065	0.053	0.035
Dutch	0.067	0.095	0.064	0.084	0.087	0.067
Listeners	0.092	0.096	0.066	0.095	0.093	0.061
	0.078	0.080	0.031	0.089	0.089	0.042
	0.086	0.093	0.094	0.089	0.096	0.036
	0.100	0.098	0.070	0.100	0.100	0.078
	0.056	0.073	0.033	0.085	0.064	0.047
	0.093	0.094	0.034	0.095	0.083	0.000
	0.067	0.090	0.042	0.064	0.068	0.042
	0.094	0.099	0.045	0.087	0.051	0.005
	0.092	0.094	0.015	0.088	0.088	0.045

Table 5.21: Burst-adjusted perceptual distance values from AX discrimination experiment, English stimuli only

Repeated measures analyses of variance (ANOVAs) were run on the data in these tables, in order to gauge the influence the four factors (manner of articulation, place contrast, language of speaker, and language of listener) had on the combined perceptual distance scores for stops and nasals. Tables 5.22 and 5.23 list the results of these ANOVAs. The analysis of these results only addresses the significant effects of manner of articulation in these ANOVAs, since only the organization of manner of articulation data changed between these ANOVAs and those run for the same data sets in chapters 3 and 4.

	F	df	Sig.
Speaker	7.069	1,14	**0.020**
Speaker * Listener	0.000	1,14	0.989
Manner	18.97	1,14	**< .001**
Manner * Listener	0.488	1,14	0.496
Contrast	7.981	2,13	**0.010**
Contrast * Listener	0.240	2,13	0.790
Speaker * Manner	1.874	1,14	0.193
Speaker * Manner * Listener	1.106	1,14	0.311
Speaker * Contrast	11.96	2,13	**< .001**
Speaker * Contrast * Listener	0.405	2,13	0.675
Manner * Contrast	1.619	2,13	0.236
Manner * Contrast * Listener	1.539	2,13	0.251
Speaker * Manner * Contrast	2.231	2,13	0.147
Speaker * Manner * Contrast * Listener	2.595	2,13	0.113
Listener	0.417	1,14	0.529

Table 5.22: Burst-adjusted Magnitude Estimation ANOVA results

	F	df	Sig.
Speaker	0.816	1,18	0.378
Speaker * Listener	0.073	1,18	0.790
Manner	0.316	1,18	0.581
Manner * Listener	0.423	1,18	0.523
Contrast	80.64	2,17	**< .001**
Contrast * Listener	1.470	2,17	0.258
Speaker * Manner	0.297	1,18	0.592
Speaker * Manner * Listener	0.340	1,18	0.567
Speaker * Contrast	65.68	2,17	**< .001**
Speaker * Contrast * Listener	0.798	2,17	0.467
Manner * Contrast	0.195	2,17	0.824
Manner * Contrast * Listener	0.834	2,17	0.451
Speaker * Manner * Contrast	1.213	2,17	0.322
Speaker * Manner * Contrast * Listener	0.299	2,17	0.745
Listener	0.014	1,18	0.907

Table 5.23: Burst-adjusted AX Discrimination ANOVA results

The only such significant effect in the ANOVA for the adjusted magnitude estimation data is a main effect for manner of articulation (F = 18.97, df = 1,14, p = .001). Post-hoc t-tests (in Table 5.24) show that the direction of this effect is, in fact, in favor of stops, which hold a significant perceptual advantage over nasals (p = .022). In other words, factoring in a realistic percentage of audible release burst cues into measures of the perceptibility of place contrasts does provide stops with a significant perceptual advantage over nasals. This result implies that perception may influence the phonological tendency of nasals to undergo place assimilation more often than stops, in the way that Kohler (1990), Hura et al. (1992) and Jun (1995) envisioned.

However, the ANOVA in Table 5.23 does not reveal a main effect for manner of articulation in the burst-adjusted AX discrimination data (F = .316; df = 1,18; p = .581; n.s.) This non-result indicates that audible release bursts--even though they appear more often after stops than after nasals--do not provide stops with a clear perceptual advantage over nasals in the VCCV sequences that were used in the AX discrimination experiment. In a sense, the failure of release bursts to make a perceptual difference between stops and nasals in these stimuli matters more than the significant effect found in the magnitude estimation data. It indicates that the combination of release bursts and perceptual salience does not have relevance for the process of place assimilation in the context where it is supposed to happen.

Manner	Mean Z	T-test
Stops	-0.293	**0.022**
Nasals	-0.575	---

		T-tests	
Contrast	Mean Z	vs. Dor-Lab	vs. Cor-Lab
Dor-Cor	-0.528	**0.019**	0.972
Dor-Lab	-0.249	---	0.096
Cor-Lab	-0.524	---	---

Speaker	Mean Z	T-test
Dutch	-0.510	0.057
English	-0.357	---

Speaker * Contrast			T-Tests	
Speaker	Contrast	Mean Z	vs. Dor-Lab	vs. Cor-Lab
Dutch	Dor-Cor	-0.394	0.513	**0.019**
	Dor-Lab	-0.349	---	**0.004**
	Cor-Lab	-0.620	---	---
English	Dor-Cor	-0.663	**0.006**	**< .001**
	Dor-Lab	-0.298	---	0.185
	Cor-Lab	-0.110	---	---

Contrast * Speaker			
Contrast	Speaker	Mean Z	T-tests
Dor-Cor	Dutch	-0.394	**0.011**
	English	-0.663	---
Dor-Lab	Dutch	-0.349	0.694
	English	-0.298	---
Cor-Lab	Dutch	-0.620	**0.004**
	English	-0.110	---

Table 5.24: T-tests for significant effects in burst-adjusted magnitude estimation ANOVA

The nature of producing audible release bursts in this context may, in fact, account for the conflicting results between the re-analyses of the two perceptual studies. In order to create the VCCV stimuli, the release bursts of the original VC stimuli had to be cut in order to match the durations of the release bursts that were produced in the readings of the Dutch and English sentences. The perceptual strength of these release burst cues may have diminished as they became shorter in duration. The comparative shortness of release bursts in consonant clusters might also account for Henderson & Repp's earlier finding that even trained phoneticians had difficulty perceiving the presence of these cues in real-world productions of consonant clusters. Henderson & Repp (1982) suggested that the difficulties of perceiving these release bursts could account for why phoneticians have traditionally assumed that they did not exist in consonant clusters. This evidence therefore suggests that release bursts in consonant clusters are markedly different, perceptually, from the same bursts which appear phrase-finally (or, more specifically, at the end of VC syllables). The potential ability of these release bursts to provide stops with the perceptual advantage they need to resist the process of place assimilation thus seems to have been overestimated.

One interesting pattern of release burst production that Henderson & Repp (1982) noticed—i.e., that stops are more likely to have release bursts, the further back the following stop is in the oral tract--provides another justification for this argument. The proportion of release bursts for stops whose place of articulation is further forward than the stops which follow them is significantly greater than the proportion of release bursts of stops whose place of articulation is further back in the oral tract than that of the stops which follow them (see Table 5.25). This pattern is strong, and it has been replicated in both this study and Henderson & Repp (1982). One might, therefore, suspect that it could have an influence on processes of place assimilation. However, there are no known assimilatory processes which affect only stops and nasals with further back places of articulation, whenever they precede stops with further forward places of articulation. Korean, in fact, has an assimilatory process which exhibits just the opposite pattern: further forward stops such as /p/ and /t/ assimilate to a further back place of articulation whenever they precede /k/.

English

Plosives	L-C	L-D	C-D	Total	Pct.	Diff.		95%
Front	15	14	10	39	0.81	**0.42**	+/-	**0.18**
Back	0	8	11	19	0.40			
N =	16	16	16	48				

Nasals	L-C	L-D	C-D	Total	Pct.	Diff.		95%
Front	8	9	7	24	0.50	**0.40**	+/-	**0.17**
Back	0	2	3	5	0.10			
N =	16	16	16	48				

Dutch

Plosives	L-C	L-D	C-D	Total	Pct.	Diff.		95%
Front	15	11	14	40	0.83	0.15	+/-	0.17
Back	11	8	14	33	0.69			
N =	16	16	16	48				

Nasals	L-C	L-D	C-D	Total	Pct.	Diff.		95%
Front	7	8	6	21	0.44	**0.31**	+/-	**0.17**
Back	0	2	4	6	0.13			
N =	16	16	16	48				

Table 5.25: Difference in proportions of nasals and stops with release bursts, depending on whether their place of articulation was in front of or behind the place of articulation of the following consonant, as produced by speakers of English and Dutch

5.4 Conclusion

This study therefore fails to provide any evidence that the presence of release bursts may indirectly affect the ability of stops to resist place assimilation by increasing the perceptual distinctiveness of their place of articulation. Moreover, the results of this production study indicate that perceptual salience of release bursts produced in consonant clusters differs significantly from bursts produced phrase-finally, or at the end of stops in VC syllables. The results of studies which have investigated the peception of place in stops and nasals in VC syllables alone (such as Pols 1983) therefore seem to have little relevance for possible perceptual influences on place assimilation, regardless of whether or not their stop stimuli contained audible release bursts.

The results of this study therefore pound one more nail into the coffin of the hypothesis that perception influences nasals into undergoing place assimilation more often, cross-linguistically, than stops. This hypothesis has not held up under direct testing of VCCV sequences either in clear listening conditions (this study, Winters 2001) or in noise (Winters 2002). Apparent cross-linguistic differences in the perception of nasals between Dutch (Pols 1983) and English have also been shown to be the result of language-specific differences in the production of the coronal place of articulation between these two languages. This study has also shown that the perceptual differences between stops and nasals which emerged in studies such as Mohr & Wang (1968) only exist because of acoustic cues in VC and CV syllables which are not found in the appropriate contexts and conditions in which place assimilation occurs.

All of the relevant evidence thus indicates that nasals' susceptibility to place assimilation is <u>not</u> motivated by their (hypothetically) comparatively weak cues for place of articulation. A much more likely phonetic culprit for this cross-linguistic phonological pattern would therefore seem to be constraints on the <u>articulation</u> of nasals and stops in heterorganic consonant cluster sequences. Chapter 6 examines this possibility in further experimental detail.

CHAPTER 6

OBSERVING ARTICULATORY DIFFICULTIES IN THE PRODUCTION OF NASALS AND STOPS: THE REPETITION/IMITATION TASK

Researchers such as Lindblom (1983) and Kirchner (1998) (and de Courtenay 1895 and Zipf 1949 before them) have all maintained that articulatory forces influence phonological structures through a general principle of least effort. Such researchers theorize, for instance, that assimilatory process occur in phonology in order to reduce the amount of physical effort a speaker has to expend in producing certain consonant clusters. Place assimilation processes, that is, would reduce articulatory effort by changing heterorganic consonant clusters into homorganic consonant clusters. Such transformations eliminate the need to make one of the place of articulation gestures in the consonant cluster, and thereby enable speakers to save that much more biomechanical energy in their articulatory lives.

While this line of reasoning seems intuitively plausible, there have been but few studies which have provided empirical evidence showing that demands for "ease of articulation" or articulatory "laziness" might actually influence phonology in this way. In many ways, the lack of such evidence is the result of assumptions phonologists have traditionally made about the nature of articulatory effort. Both Kirchner (1998) and Lindblom (1983), for instance, assumed that articulatory effort is entirely determined by the amount of physical force or work required to make an articulation. Physical quantities such as “force” or “work” are, in turn, determined entirely by physical laws such as Hooke's Law or Newton's second law. Quantifying the articulatory difficulty of a particular gesture, then, only requires knowing the values of the appropriate variables to feed into the equations defined by these laws: mass, displacement, resistance, acceleration, etc. Since it is very difficult to determine the values of these variables through empirical observation, however, few researchers have ever tried to do so. Instead, researchers such as Kirchner and Lindblom have generally decided to implement their models of articulatory difficulty computationally, using intuitively reasonable values for the variables that determine the amount of physical work required to make an articulatory gesture. The quality of such models is generally tested by comparing predicted articulatory ease effects on phonological systems with the structures of actual phonological systems, as they are found in the languages of the world.

Either Kirchner or Lindblom’s model, for instance, would likely predict that nasals should undergo place assimilation more often, cross-linguistically, than stops. They would likely make this prediction, since heterorganic nasal-stop sequences have an extra manner of articulation gesture (the opening and closing of the velo-pharyngeal port) which heterorganic stop-stop clusters do not have. The extra work needed to lower and raise the mass of the velum to create a nasal consonant therefore ensures that nasal-stop sequences use more physical energy than stop-stop clusters, regardless of the specifics of just how much energy is involved in the production of either. A general articulatory principle of least effort should, therefore, put more pressure on a speaker’s articulatory system to reduce the amount of effort in heterorganic nasal-stop sequences than heterorganic stop-stop sequences, since they come at a higher biomechanical energy cost.

One way to reduce the amount of articulatory effort in producing such consonant sequences is to allow the nasal-stop clusters to assimilate into homorganic clusters, and thereby eliminate the articulatory effort involved in the production of one of the place of articulation gestures. The fact that such processes do, in fact, target nasals more often than stops seems to provide some justification for this theory.

On a smaller scale, however, the case of nasal/stop place assimilation may provide an interesting way to test the basic assumption behind these models of articulatory effort—i.e., that speakers simply avoid articulations which require a lot of physical work or energy. If a desire to reduce articulatory effort motivates optional procesess of place assimilation (as in, e.g., English and Dutch), as well, it should target nasals more often (optionally) than it does stops. Hume & Johnson (2001), in developing a model of the interaction of "external" forces and phonology, conceived of articulatory difficulty as "filtering" phonological patterns in this way. Hume & Johnson draw an analogy between this articulatory filtering force and Ohala's (1981) proposal that "innocent misapprehensions" in perception can be a source of sound change: "The filtering action imposed by production takes a similar form. The cognitive symbolic representation requires that the speaker make a sound that is hard to say. In some instances the speaker will fail to produce the sound and say something else and in this way contribute to a change in p." The difficulty of making certain gestures, in other words, should lead speakers to use more substitutions, reductions, assimilations, deletions, etc., in order to simplify their articulatory task.

There may, however, be other motivations behind casual speech reduction phenomena than just gross considerations of articulatory effort or energy. Boersma

(1998), for instance, suggested that articulatory difficulty—and its influences on phonology—might consist of more than just a desire on the part of speakers to be as lazy as is humanly possible. For this reason, Boersma (1998) criticized previous conceptions of articulatory effort as being overly simplistic:

> "Previous attempts to formalize articulatory effort run short of several generalizations, because they try to express articulatory effort into one variable...Articulatory effort depends on at least six primitives: energy, the presence of articulatory gestures, synchronization of gestures, precision, systemic effort, and coordination, and languages seem to be able to rank these separate measures individually to a certain extent."

Boersma's basic assumption about the overall effect of articulatory effort on the structure of speech is the same as in Kirchner (1998) or Lindblom (1983): that speakers will avoid producing articulations which require a lot of effort. However, the amount of articulatory effort required to make a particular sound, or gesture, in Boersma's model may be determined by any of six independent "primitives"--and not just by physical energy or work alone.

One of Boersma's primitives was the articulatory difficulty of coordinating the production of two or more gestures at the same time. In Boersma's optimality theoretic model, this primitive was formalized with a general family of constraints against the coordination of more than one gesture in the same segment—

(6.1) *Coord ($gesture_1$, $gesture_2$): "The two gestures $gesture_1$ and $gesture_2$ are not coordinated."

The hypothetical difficulty of coordinating gestures in this way might be another factor which makes the production of heterorganic nasal-stop clusters more difficult than the production of heterorganic stop-stop clusters. Heterorganic stop-stop clusters involve the production of two distinct gestures: one for the first stop closure and the second for the second stop closure. Heterorganic nasal-stop clusters add to this combination of gestures a velic gesture, which must be produced in coordination with the first stop gesture. Combining three gestures in this way requires more coordination than just producing two stop gestures in sequence. Coordination should, therefore, make heterorganic nasal-stop gestures more effortful for speakers than heterorganic stop-stop sequences.

An analogy might help clarify the distinction between this kind of articulatory effort and the amount of effort which is determined solely by calculating physical work or energy. Many people have difficulty rubbing their stomach and patting their head at the same time. Executing <u>both</u> of these gestures simultaneously obviously requires more physical energy or work than performing either of the gestures independently. However, the amount of difficulty people have in making these gestures simultaneously is (intuitively, at least) disproportionately greater than the sum of physical work required to make each of the individual gestures independently. It is the effort required to <u>coordinate</u> the gestures which accounts for the disproportionate amount of difficulty. It is not unreasonable to assume that similarly disproportionate amounts of effort may be required to coordinate articulatory gestures--as in, for example, heterorganic nasal-stop sequences.

Establishing that it is difficult to coordinate a particular bundle of gestures in a consonant sequence, however, requires more than just showing that speakers tend to avoid such coordinate feats in favor of easier-to-articulate sequences. Coordinative difficulties—if they exist—should manifest themselves in the hypothetically difficult forms that speakers often try to avoid saying. One possible way of investigating these articulatory difficulties empirically is to examine the acoustic variability of such difficult-to-articulate sequences, whenever speakers succeed in producing them. Measuring the token-to-token variability of particular speech gestures has emerged as a reliable metric of “fluency” in studies of speech development in children. Several of these studies (e.g., Smith 1978, Smith 1992) have noted that children's speech productions are more variable than adults. The duration of children's productions of consonants, for instance, tend to show more variability than adult productions of the same consonants. Researchers have thus concluded that decreasing variability in speech productions reflects the increased skill of older speakers to produce speech consistently. Munson (2000), for instance, treated variability in the duration of consonant cluster productions as a measure of speaker "fluency" in producing those clusters. Speakers may, in the same sense, be less “fluent” at producing heterorganic nasal-stop sequences than they are at producing heterorganic stop-stop sequences, since nasal-stop sequences require the coordination of a greater number of articulatory gestures. If so, this comparative lack of “fluency” should manifest itself in greater token-to-token variability in speakers’ productions of nasal-stop sequences.

There are, therefore, at least two different interpretations of “articulatory difficulty” which seem to have relevance to the question of why nasals tend to undergo place assimilation more often, cross-linguistically, than stops. One of these interpretations (advocated by, among others, Lindblom 1983 and Hume & Johnson 2001) maintains that articulatory difficulty is actively avoided by speakers and therefore leads directly to phonological processes such as assimilation, deletion, etc., which simplify articulatory tasks. Another school of thought maintains that the skill—or fluency—

speakers have at producing certain sound sequences manifests itself in low-level acoustic variability across individual productions of those sound sequences. The greater the variability, the less skilled speakers are at producing those sequences. There are intuitively appealing reasons to suspect that speakers might have more articulatory difficulty—in both of the above senses—with the production of nasal-stop clusters than they do with the production of stop-stop clusters. The following production study therefore attempted to find empirical evidence to support this hypothesis and show that nasal-stop sequences are, in fact, more difficult for speakers to produce both accurately and consistently than stop-stop sequences.

6.1 Methods

The participants' basic task in this study was (following Munson 2000) a repetition, or imitation, task. The participants heard individual VCCV sequences and were instructed to repeat each of these stimuli three consecutive times, while imitating the original stimulus as closely as possible. VCCV sequences were presented to listeners because they included both stops and nasals in a likely context for place assimilation to occur—i.e., before another stop in an intervocalic consonant cluster. These VCCV stimuli therefore consisted of all of the VCCV stimuli from the AX discrimination experiment, along with similarly constructed stimuli from the original speaker productions (see chapter 3) with /u/ and /i/ as the vowels in the VC and CV syllables. For each VCCV stimulus, then, the vowels in both syllables were identical, but could be any of the corner vowels /a/, /i/ or /u/. The stimuli varied along all of the other factors in the AX discrimination experiment, as well: there were three places of articulation (labial, coronal, dorsal) for the first consonant, along with three manners of articulation (nasal, stop with burst, stop without burst) for the same consonant and three places of articulation (labial, coronal, dorsal) for all of the stops in the second consonant position. Participants heard each stimulus type as spoken by both Dutch and English speakers. As in the earlier perception experiments, these speakers were matched across genders and language groups, so that there were effectively two distinct sets of stimuli in the experiment: one with stimuli produced by the Dutch male and English female speakers, and the other with stimuli produced by the English male and Dutch female speakers.

Ten native speakers of Dutch and ten native speakers of English participated in this experiment, with five speakers of each group hearing the two distinct sets of VCCV stimuli. All of the native Dutch speakers had some experience with and knowledge of the English language, while none of the native English speakers had any knowledge of Dutch. The native Dutch speakers were recruited from the student population at the Rijksuniversiteit Groningen, in the Netherlands. The native English speakers were recruited from the student population at the Ohio State University in Columbus, Ohio. All Dutch speakers received 25 Euros for participating in the experiment, while all English speakers received $20; more details on the method of compensating participants are given below.

The listening and recording conditions differed slightly between the two groups of participants, due to extenuating circumstances. For both groups of listeners, the method of stimuli presentation was the same: they heard the stimuli as presented to them by a customized E-prime (version 1.0) program, running on a PC, over Sennheiser HMD

25-L headphones. After playing each stimulus, the customized E-prime program would present the listeners with a screen instructing them to "repeat what they heard" three times.

The English-speaking participants worked through this process inside a sound-attenuated booth. The computer screen, with which they interacted, sat inside the booth, while the PC running the customized E-prime program sat outside the booth. The participants' repetitions of the VCCV stimuli were recorded through the Sennheiser HMD 25-L head-mounted microphones, connected through a Symetrix SX 302 dual microphone pre-amplifier (gain ≈ 70 dB) to a Sony DTC-790 digital audio tape deck, sitting outside the booth. All of the recordings for the native English-speaking participants were made onto Sony PDP 65-C DAT cassettes with this digital audio tape deck.

The recording and listening conditions were thus slightly more optimal for the English-speaking participants than they were for the Dutch-speaking participants. The Dutch participants heard all stimuli as played to them by the same, customized E-Prime program, running on a PC laptop, over the same pair of Sennheiser HMD-25L headphones, in a quiet room. It proved necessary to run this portion of the study in a quiet room, rather than a sound-attenuated booth, since the PC laptop had a tendency to overheat and crash after running for any extended period of time in the small, sound-attenuated booth available at the Rijksuniversiteit Groningen. The Dutch participants' repetitions of the VCCV stimuli were also recorded in the same quiet room. The participants spoke through the Sennheiser HMD-25L head-mounted microphones, connected through a Symetrix SX202 (gain ≈ 50 dB) dual microphone pre-amplifier to a

Sony PCM-M1 portable digital audio recorder, with which the repetitions were recorded onto Quantegy R-04 certified master digital audio tapes. These recording conditions ensured a greater amount of background noise in the recordings of the Dutch participants' repetitions; example waveforms and spectrograms of recordings for both groups of participants are given below in Figures 6.1 and 6.2.

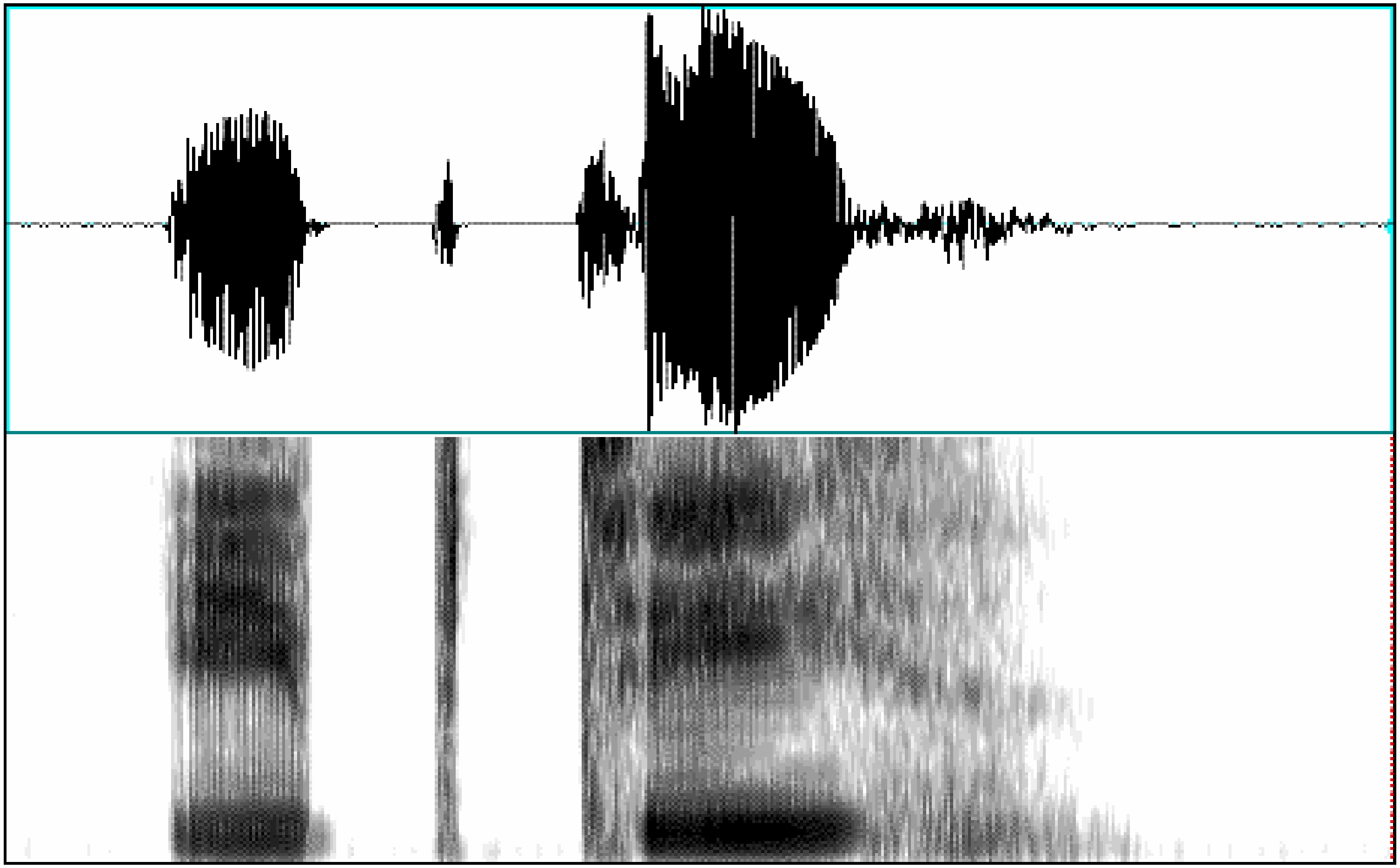


Figure 6.1: Example waveform and spectrogram for English speaker production of /ipti/ (dynamic range = 60 dB)

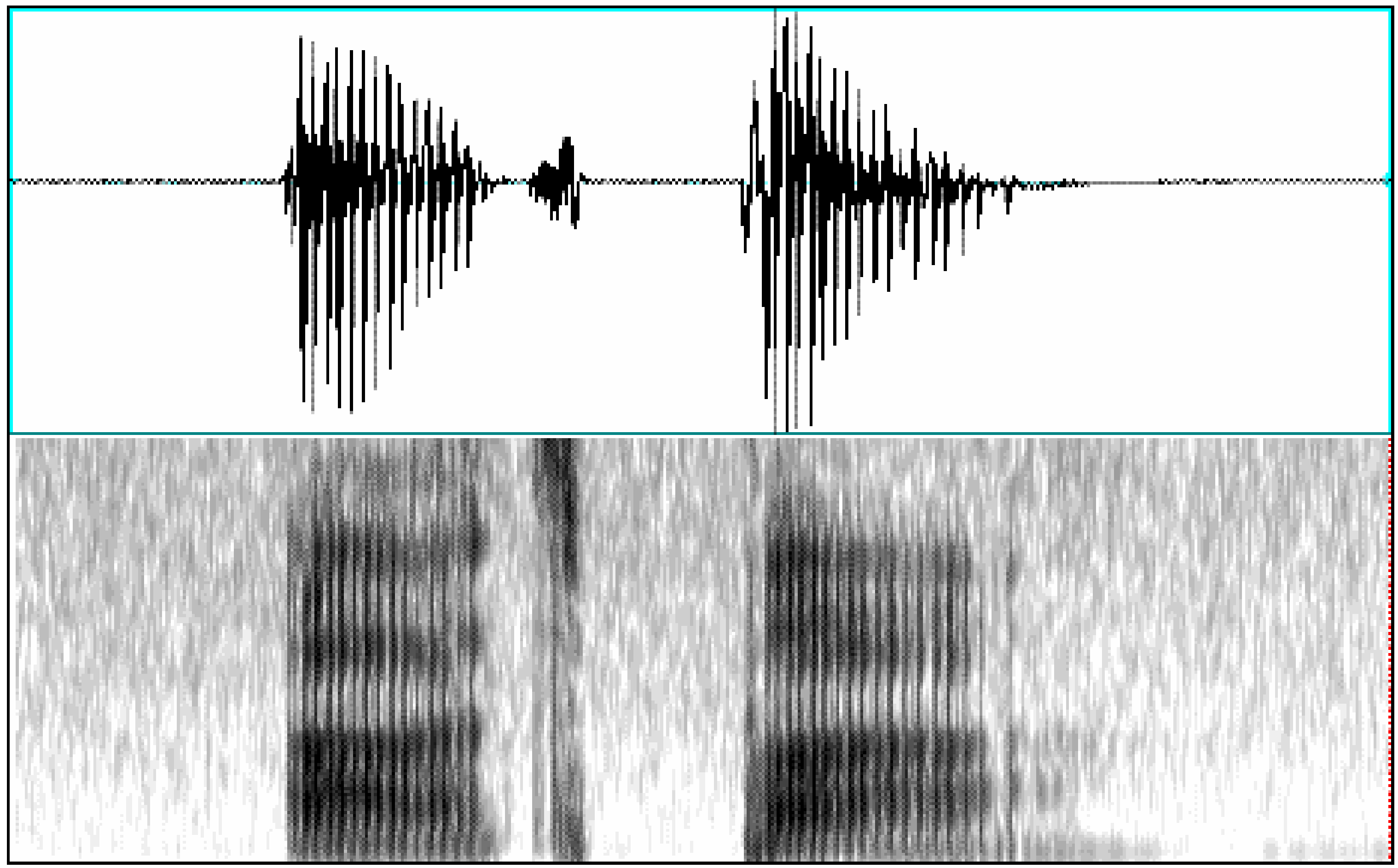


Figure 6.2: Example waveform and spectrogram for Dutch speaker production of /atpa/ (dynamic range = 60 dB)

While there was clearly more background noise in the recordings of the Dutch speakers' repetitions, it was still possible to transcribe these utterances phonetically and analyze them acoustically, in order to determine how consistent and accurate the Dutch speakers had been in making them. The results of this analysis do not provide any reason to believe that the recording and listening conditions had any significant effect on the accuracy and consistency of stimulus repetition between the two groups of participants.

The E-Prime program presented all participants with 162 basic VCCV sequences (162 = 3 vowels x 3 C_1 places of articulation x 3 C_1 manners of articulation x 3 C_2 places of articulation x 2 speakers), and instructed them to repeat each one three times. The participants heard each VCCV sequence twice in the experiment, and therefore made a total of six repetitions of each. In all, each speaker thus made a total of 972 stimulus repetitions.

The instructions given to all of the participants before the experiment informed them that they had an economic incentive to produce repetitions which mimicked the original utterances as closely as possible. All English-speaking participants were nominally guaranteed $8 for participating in the experiment, but were told that they could earn four extra cents for each set of three repetitions which matched the original utterance within a pre-set acoustic threshold. As just one example of such an acoustic threshold, the participants were told that the durations of the segments in their repetitions had to match the segment durations in the original utterance by within five milliseconds.

The English-speaking listeners were therefore told that, with 324 sets of repetitions to make and the opportunity to earn four extra cents with each one, they could earn up to $20 for their participation in the one hour-long experiment. The Dutch participants were given a similar incentive, except with a 10 Euro baseline fee and a reward of five extra cents per set of repetitions, with a possible maximum of 25 Euros.

Listeners were given this economic incentive in order to motivate them to produce all of the consonant clusters in the VCCV sequences precisely as they heard them. Fowler & Dekle (1991) have shown that such incentives can greatly improve subject performance in tasks with which they would otherwise be completely unfamiliar, or unskilled. Economic incentives were thus used in this experiment in the hopes that they would encourage the participants to accurately reproduce consonant clusters which they would normally allow to undergo place assimilation in casual speech in English and Dutch.

In order to prevent participants from being able to perform the repetition/imitation task too well, however, they were also given a cognitive load task to perform while repeating the stimuli they heard. This cognitive load task was adapted from Harnsberger & Pisoni (1999), who determined that requiring speakers to remember a series of digits while they performed a production task was an effective and reliable means of inducing reduction phenomena in the participants' speech. Harnsberger & Pisoni (1999) also found that the length of the digit series that induced the most repetitions varied by speaker. This optimal length could, for example, be as low as five or six digits for some speakers, but as high as eight for others. Mielke (to appear, p.c.), however, found that a string of eight digits provided a large enough cognitive load for most speakers to induce a significant number of reductions in their speech.

All speakers in this study therefore had to perform the repetition/imitation task while keeping a string of eight digits in memory. The customized E-Prime program presented the listeners with one such string before presenting any of the VCCV stimuli, along with the instructions that they should memorize the numbers and be able to recall them after a certain number of VCCV repetitions. After 27 sets of repetitions, the E-Prime program would then prompt the participants to say the last string of memorized digits out loud. After each recall prompt, the E-Prime program would then present the participants with a new string of digits, to be recalled later. This memorization process repeated itself after every 27 VCCV stimuli, for a total of 12 times in the entire experiment. The participants therefore memorized 12 different strings throughout the experiment. The same 12 strings of eight digits were used for all participants; they were presented in the following order:

(6.2) The 12 eight-digit strings

91536028
63597142
08369725
29073645
72058431
37156924
58246317
83524076
42789153
12834607
38654027
84597613

All participants were informed that all the extra cash earned by repeating VCCV sequences accurately within each set of 27 stimuli was also contingent upon remembering and recalling the string of digits for that set of stimuli correctly. The participants therefore had an economic incentive to perform both the repetition and the cognitive load task well. However, all participants were given the advertised maximum amount of money ($20 for English-speaking subjects, and 25 Euros for Dutch-speaking subjects) at the end of the experiment, regardless of performance.

The exact instructions given to all participants in this experiment is given below:

(6.3) Repetition/Imitation task instructions

THE PATTERNS IN PLACE PRODUCTION EXPERIMENT

The basic premise of this experiment is quite simple--the computer will play a "nonsense word" to you over the headphones, and then ask you to repeat what you have heard, to the best of your ability, three times in a row.

Since this task is so simple, I have included a number of complicating factors to make it interesting. First of all, I am providing you with an incentive for repeating what you hear as accurately and precisely as possible. The computer will automatically record each one of your repetitions and then calculate how well they match the original utterance.

For each set of three repetitions which matches the original utterance within a pre-determined, acoustic threshold, I will add 5 cents to the base fee of 10 euros I am paying you to participate in this experiment.

This may not sound like much, but with some 300 original utterances to repeat, you stand to make as much as 25 euros for one hour of your time if you repeat them all perfectly well.

Press any key on the keyboard to continue.

In making its calculations, the computer will (among other things) take into consideration how well the length of each sound in your repetitions matches the length of the same sounds in the original. So, if the original utterance has been spoken very quickly, you will need to repeat what you hear very quickly as well. Likewise, if the original sounds slow, you should repeat it at the same, slow speed.

The other complicating factor in this experiment is that you will have to remember a series of numbers while you do the task. Before you begin repeating anything, the computer will present you with a series of numbers; you should memorize these numbers and keep them in mind. Eventually, the computer will stop presenting you with words to repeat and ask you what the original numbers were.

If you can remember these numbers perfectly, the computer will give you credit for the accurate repetitions you've made since you had to memorize the numbers. If not, the *computer will not credit you for any of the repetitions you've made since you had to memorize the numbers.*

Press any key on the keyboard, and the computer will run you through a short demonstration of this process.

Once the computer has asked you to recall one series of numbers, it will present you with a new series, which it will ask you to recall at a later time. The experiment will continue in this way--alternating between repetitions and the memorization of new numbers--until it is finished.

To sum up, then, the extra five cents you can earn for each series of repetitions depends on you being able to: 1) make a series of accurate repetitions and 2) remember the series of numbers correctly.

If you have any questions about this experiment--or if any of the instructions are unclear--please let the experimenter know now.

You should also be aware that, since you will have to make a lot of repetitions during this experiment, you should feel free to take a breaks during the experiment to get a drink or a breath of fresh air.

Once you are ready to begin, please press any key on the keyboard to begin the testing.

Prior to the experiment, all participants ran through a short demonstration of both the number memorization recall task and the stimulus repetition task. All of the stimuli for both sets of speakers were randomized prior to the experiment and then broken down into twelve separate groups of 27 stimuli each for presentation to the participants. Each of these twelve groups of 27 was then randomized once more, at run time, by the customized E-Prime program, prior to presentation to the participants. The entire experiment lasted approximately one hour for all subjects.

6.2 Analysis

All digital recordings of the participants' repetitions were optically transferred to PC via a HDSA model ODL-276 Optical Data Link connected to a Creative Software SoundBlaster Live! Drive on the PC. The digital transfer was made using CoolEdit 2000 software at a 44100 Hz sampling rate, on a mono channel, with 16-bit resolution. CoolEdit 2000 software was also used to save each individual repetition from all production sessions into separate .wav files. In general, speakers made—as instructed--there were three repetitions for each of the 324 stimuli in the experiment. In some cases, however, speakers produced less than three repetitions, and, in other cases, some speakers produced more. In these latter cases, only the first three repetitions were saved for further analysis.

The repetitions in each .wav file were analyzed in three different ways. First, a trained phonetician made a broad phonemic transcription of all repetitions, with the help of the sound editing tool in the Praat software package. These transcriptions were made by listening to each repetition--focusing, in particular, on the first, VC, syllable in each--and occasionally looking at formant tracks to guide in the identification of the place of articulation of any of the stop or nasal consonants.

Secondly, measurements were made of the durations of each segment in all of the recorded repetitions. These measurements were made with the help of a customized Praat script (Welby, 2002), given below in full in Appendix 6.1. This script automatically saved the value--in the time dimension--of vertical markers placed in a labeling window in the Praat sound editor. Such vertical markers were placed at the following segmental markers throughout each repetition of the VCCV stimuli:

1. The beginning of the first vowel
2. The end of the first vowel (beginning of the first consonant)

3. The end of the first consonant (beginning of the first stop release burst)
4. The end of the first release burst
5. The end of the second consonant (beginning of the second stop release burst)
6. The end of the aspiration for the second consonant.
7. The end of the second vowel.

Figure 6.3 shows an example of how these markers were placed in the spectrogram window.

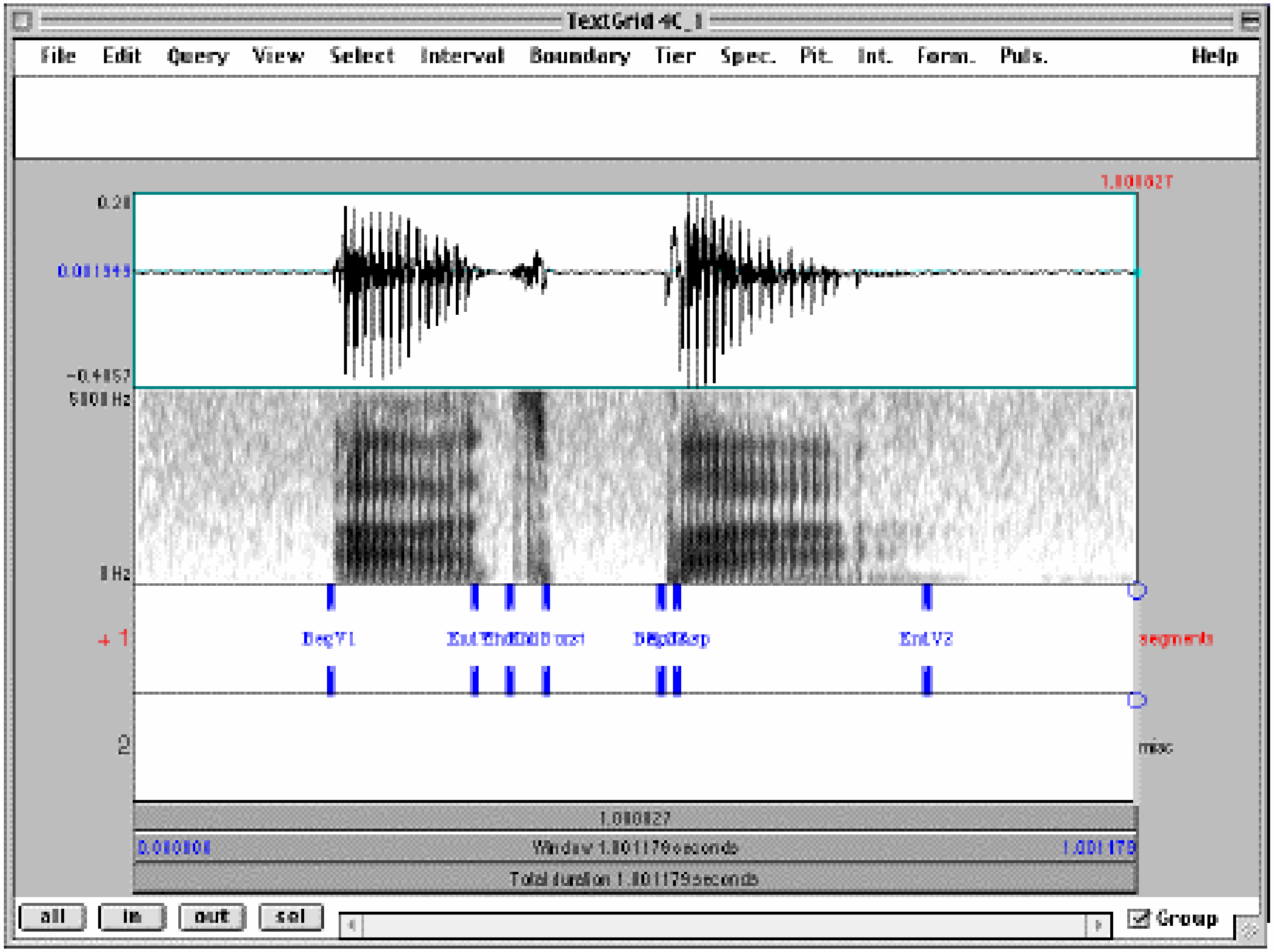


Figure 6.3: Analysis window in Praat for marking acoustic landmarks in spoken VCCV repetitions (Dutch-speaker)

Not all of these acoustic landmarks were present in all recorded repetitions; some repetitions, for instance, lacked a release burst for the first consonant. In these cases, the markers for these acoustic landmarks were simply deleted, and the Praat script only saved the timing information for all of the other markers.

Many of the repetitions that were recorded in the production experiment did not maintain the same vowel-stop-stop-vowel or vowel-nasal-stop-vowel structure of the original stimuli. Some had epenthetic third--or even fourth--consonants in the consonant cluster, while others deleted nasals, voiced consonants, changed stops into fricatives, etc. Most of these changes had a potential effect on the placement of markers for the consonant duration cues (markers 2, 3 and 4). However, most of these markers were placed without regard to extraneous changes in the structure of the sounds themselves--marker three was therefore placed at the end of the first consonant, whether that was cued by the beginning of a stop release burst, the end of a nasal murmur, or the end of consonant frication. Any epenthetic third or fourth consonants between the first and last consonants in the cluster were simply disregarded.

Marker #2--for the offset of the first vowel and onset of the first consonant--was placed prior to any voicing bars, in the case of stops, and prior to any nasal murmur, for nasals. Guidelines from Munson (2000) were also used to determine the onset of nasal murmur for these markers--primarily looking for nasal zeroes, increased formant bandwidth and decreased amplitude in higher formants. Marked off segments were also played and listened to in order to perceptually confirm the acoustic discontinuities between each.

The last acoustic analysis done on each repetition was the measurement of the first four formant values throughout the entire repetition. Such measurements were made with Praat's automatic formant tracking algorithm (Burg method, time step = 0.01 seconds, maximum number of formants = 4, maximum formant = 5500 Hz, window length = .025 seconds, pre-emphasis = 50 Hz). Such measurements were made for each formant individually, and saved by a customized Praat script to an external file. All four formant measurement files for each repetition were subsequently combined by a customized Perl script, and saved as a single text file for later analysis.

6.3 Results

The repetition/imitation task had both perceptual and articulatory components. The participants were asked to first listen to a VCCV stimulus and then attempt to imitate what they had heard, in three consecutive repetitions, as closely as possible. The successful replication of a VCCV stimulus by any of the participants therefore not only required the correct perception of the sounds in the stimulus string, but also the subsequent, accurate production of what had been heard. The perceptual aspects of this task therefore involved, to a certain extent, a process of perceptual identification--as in, e.g., Winters (2001) or Pols (1983). The identification task in this study, however, involved an open response set and also required the listeners to articulate their responses, instead of merely registering them with the push of a button or the use of a pencil and paper. This articulatory aspect of the repetition/imitation task may, therefore, have induced patterns in the responses which would not exist in the results of tasks which tested perception alone. The ability of the participants in this study to accurately repeat the VCCV stimuli--as they perceived them--may therefore be inferred by comparing the patterns of results in this study to those of the previous AX discrimination experiment. While perceptual difficulties may induce similar effects in the results of both studies, other significant effects may emerge in the reptition/imitation task alone, which are presumably the result of articulatory constraints on the spoken repetition process.

In order to make this broad comparison between the two experiments, then, the transcribed repetitions in this experiment were treated as if they were responses to the original VCCV stimuli in a perceptual identification task. Determining the accuracy of these responses therefore required not only calculating the percentage of correct responses (i.e., "hits") to the stimuli, but also factoring out the leveling effects of bias towards certain kinds of responses on the part of the participants. Such response bias manifests itself in "false alarms"--inappropriate responses of one category, or sound, when a stimulus containing a different category or sound has been presented to the listener. A response of /apka/ to an /atka/ stimulus, for instance, includes a "false alarm" for the /p/ response option. One method of factoring out the effects of bias in identification responses is to calculate the sensitivity of listeners to distinctions between particular sound categories which have been presented in the stimuli. One such measure of sensitivity is A', which determines sensitivity without assuming particular internal distributions on which listener responses are based (Grier 1971). A' is calculated as follows:

(6.4) A-prime = $(.5 + (H - FA) * (1 + H - FA)) / (4 * H * (1 - FA))$

where FA = percentage of false alarms, and H = percentage of hits for a particular response category.

A' was used to determine the accuracy of the participants' spoken responses to the auditorily presented VCCV stimuli in this study. Such A' values were calculated only with respect to the place of articulation of the first consonant in the VCCV stimuli, since this consonant is in the most likely environment for place assimilation to occur. Correct repetitions of the place of articulation of these consonants counted as "hits" while inappropriate places of articulation for these consonants were considered "false alarms" on the part of the speaker. Only the place of articulation of the first consonant in the listeners' responses was considered in determining "hit" and "false alarm" totals; the manner of articulation of these consonants was not taken into account. Both /apta/ and /afta/ would have counted as labial responses for the C_1 consonant in the intervocalic cluster, for instance.

Place of articulation "hits" and "false alarms" were tabulated only for listener responses which had less than three consonants in the intervocalic cluster. Listener responses with more than two consonants in these clusters--e.g., /aptka/--were simply left out of the analysis, since it was impossible to determine which consonant (if any) the listener may have thought corresponded to the C_1 consonant in the original stimulus. It was also impossible to distinguish between responses with sequences of two homorganic stop consonants and responses with only a single, intervocalic stop consonant. An /appa/ response without a release burst after the initial /p/, for instance, would be indistinguishable from an /apa/ response. When such responses were made to stimuli with a homorganic stop cluster in the VCCV sequence, the analysis gave the listeners the benefit of the doubt. Responses transcribed with single, intervocalic consonants in these cases were considered to be equivalent to a response with a homorganic stop cluster in between the two vowels. An /apa/ response to an /appa/ stimulus, in other words, registered a "hit" for correct identification of the place of articulation for the first consonant in the cluster. Similar assumptions were not made, however, for single consonant responses to stimuli with heterorganic consonant clusters in between the vowels. In these cases, the listeners registered neither a "hit" nor a "false alarm," but only a "miss" for the place of articulation for the first consonant, since they had misperceived the presence of at least one place of articulation in the cluster and not clearly replaced it with another.

A' values were calculated in this way for all listener responses, with respect to only the place of articulation of the first consonant in each consonant cluster. Tables 6.1-6.4 list hit and false alarm counts--and the corresponding A' values--for the places of articulation in these C_1 consonants, for all listeners and stimuli types. Each of these tables lists hit and false alarm counts for a particular language of listener, language of speaker combination, and breaks down the data by the manner and place of articulation of the C_1 consonant. A repeated measures analysis of variance was run on the A-prime values that were calculated from breaking down the response data in this way, considering only language of speaker, language of listener, manner and place of articulation of the first consonant as potential influences the A' values for listener sensitivity to place distinctions in the first consonant of the intervocalic cluster.

	Burst Stimuli					No Burst Stimuli					Nasal Stimuli				
#1	H	FA	P(H)	P(FA	A'	H	FA	P(H)	P(FA	A'	H	FA	P(H)	P(FA	A'

				)					)					)	
D	52	14	0.96	0.13	0.96	33	42	0.61	0.39	0.68	40	50	0.74	0.46	0.72
C	34	4	0.63	0.04	0.89	16	29	0.30	0.27	0.53	14	12	0.26	0.11	0.68
L	40	16	0.74	0.15	0.87	24	18	0.44	0.17	0.74	33	13	0.61	0.12	0.84
#2	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	40	1	0.74	0.01	0.93	42	40	0.78	0.37	0.79	17	27	0.31	0.25	0.57
C	36	25	0.67	0.23	0.80	21	21	0.39	0.19	0.69	21	36	0.39	0.33	0.56
L	42	18	0.78	0.17	0.88	23	15	0.43	0.14	0.75	26	35	0.48	0.32	0.64
#3	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	41	6	0.76	0.06	0.92	42	38	0.78	0.35	0.80	40	52	0.74	0.48	0.71
C	6	13	0.11	0.12	0.48	14	11	0.26	0.10	0.70	19	20	0.35	0.19	0.67
L	36	21	0.67	0.19	0.82	30	19	0.56	0.18	0.79	24	7	0.44	0.06	0.82
#4	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	37	14	0.69	0.13	0.86	32	33	0.59	0.31	0.72	20	35	0.37	0.32	0.55
C	30	15	0.56	0.14	0.81	20	35	0.37	0.32	0.55	24	42	0.44	0.39	0.55
L	46	20	0.85	0.19	0.90	24	18	0.44	0.17	0.74	24	17	0.44	0.16	0.75
#5	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	50	19	0.93	0.18	0.93	24	44	0.44	0.41	0.54	13	34	0.24	0.31	0.40
C	24	7	0.44	0.06	0.82	12	32	0.22	0.30	0.39	17	39	0.31	0.36	0.45
L	36	25	0.67	0.23	0.80	17	30	0.31	0.28	0.54	22	36	0.41	0.33	0.57
#6	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	40	11	0.74	0.10	0.89	31	29	0.57	0.27	0.74	35	14	0.65	0.13	0.85
C	36	14	0.67	0.13	0.86	19	29	0.35	0.27	0.59	8	15	0.15	0.14	0.52
L	40	15	0.74	0.14	0.88	25	27	0.46	0.25	0.69	46	44	0.85	0.41	0.82
#7	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	42	18	0.78	0.17	0.88	31	41	0.57	0.38	0.66	21	24	0.39	0.22	0.66
C	51	22	0.94	0.20	0.93	14	31	0.26	0.29	0.46	25	37	0.46	0.34	0.61
L	28	0	0.52	0.00	0.88	17	29	0.31	0.27	0.55	31	23	0.57	0.21	0.77
#8	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	45	7	0.83	0.06	0.94	23	33	0.43	0.31	0.61	24	12	0.44	0.11	0.78
C	42	13	0.78	0.12	0.90	15	36	0.28	0.33	0.43	14	14	0.26	0.13	0.66
L	49	5	0.91	0.05	0.96	21	34	0.39	0.31	0.57	43	55	0.80	0.51	0.74
#9	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	29	21	0.54	0.19	0.77	18	12	0.33	0.11	0.73	19	10	0.35	0.09	0.76
C	41	20	0.76	0.19	0.87	26	28	0.48	0.26	0.69	13	19	0.24	0.18	0.59
L	27	12	0.50	0.11	0.80	32	38	0.59	0.35	0.69	34	59	0.63	0.55	0.58
#10	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'

D	35	12	0.65	0.11	0.86	28	40	0.52	0.37	0.63	47	37	0.87	0.34	0.85
C	49	21	0.91	0.19	0.92	14	32	0.26	0.30	0.45	17	14	0.31	0.13	0.70
L	28	10	0.52	0.09	0.82	20	28	0.37	0.26	0.61	21	26	0.39	0.24	0.64

Table 6.1: Hit, False Alarm, and A-Prime totals for all Dutch listeners, in response to all Dutch stimuli

	Burst Stimuli					No Burst Stimuli					Nasal Stimuli				
#1	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	24	1	0.44	0.01	0.85	23	31	0.43	0.29	0.63	48	34	0.89	0.31	0.87
C	39	30	0.72	0.28	0.81	20	30	0.37	0.28	0.59	26	11	0.48	0.10	0.80
L	36	31	0.67	0.29	0.78	25	32	0.46	0.30	0.65	30	12	0.56	0.11	0.83
#2	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	25	2	0.46	0.02	0.85	19	36	0.35	0.33	0.52	21	36	0.39	0.33	0.56
C	38	47	0.70	0.44	0.71	25	40	0.46	0.37	0.59	26	42	0.48	0.39	0.59
L	26	23	0.48	0.21	0.72	25	17	0.46	0.16	0.76	19	15	0.35	0.14	0.71
#3	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	31	13	0.57	0.12	0.83	29	20	0.54	0.19	0.77	38	35	0.70	0.32	0.78
C	33	27	0.61	0.25	0.77	38	23	0.70	0.21	0.83	28	27	0.52	0.25	0.72
L	31	18	0.57	0.17	0.80	31	17	0.57	0.16	0.81	24	0	0.44	0.00	0.86
#4	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	35	9	0.65	0.08	0.87	22	30	0.41	0.28	0.62	39	20	0.72	0.19	0.85
C	40	39	0.74	0.36	0.78	32	33	0.59	0.31	0.72	31	33	0.57	0.31	0.71
L	26	10	0.48	0.09	0.81	25	17	0.46	0.16	0.76	22	16	0.41	0.15	0.74
#5	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	43	19	0.80	0.18	0.88	13	39	0.24	0.36	0.33	20	20	0.37	0.19	0.68
C	39	9	0.72	0.08	0.90	12	35	0.22	0.32	0.35	31	45	0.57	0.42	0.64
L	31	21	0.57	0.19	0.78	21	42	0.39	0.39	0.50	26	19	0.48	0.18	0.75
#6	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	28	3	0.52	0.03	0.86	33	33	0.61	0.31	0.74	46	22	0.85	0.20	0.89
C	30	26	0.56	0.24	0.75	26	32	0.48	0.30	0.66	19	8	0.35	0.07	0.77
L	36	36	0.67	0.33	0.75	27	11	0.50	0.10	0.81	50	17	0.93	0.16	0.94
#7	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	37	17	0.69	0.16	0.85	39	50	0.72	0.46	0.71	26	26	0.48	0.24	0.70
C	43	38	0.80	0.35	0.81	22	27	0.41	0.25	0.65	38	39	0.70	0.36	0.76
L	23	2	0.43	0.02	0.84	12	13	0.22	0.12	0.64	30	2	0.56	0.02	0.88
#8	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	41	31	0.76	0.29	0.82	33	43	0.61	0.40	0.68	27	25	0.50	0.23	0.72
C	34	15	0.63	0.14	0.84	21	24	0.39	0.22	0.66	26	23	0.48	0.21	0.72
L	31	9	0.57	0.08	0.85	18	21	0.33	0.19	0.65	41	19	0.76	0.18	0.87
#9	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	26	13	0.48	0.12	0.79	14	13	0.26	0.12	0.67	12	8	0.22	0.07	0.71
C	33	29	0.61	0.27	0.76	34	38	0.63	0.35	0.72	23	20	0.43	0.19	0.72

L	32	13	0.59	0.12	0.83	23	32	0.43	0.30	0.62	43	40	0.80	0.37	0.80
#10	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	30	13	0.56	0.12	0.82	28	44	0.52	0.41	0.60	49	29	0.91	0.27	0.89
C	37	15	0.69	0.14	0.86	19	21	0.35	0.19	0.66	23	6	0.43	0.06	0.82
L	39	21	0.72	0.19	0.85	23	27	0.43	0.25	0.66	43	12	0.80	0.11	0.91

Table 6.2: Hit, False Alarm, and A-Prime totals for all Dutch listeners, in response to all English stimuli

	Burst Stimuli					No Burst Stimuli					Nasal Stimuli				
#1	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	48	9	0.89	0.08	0.95	45	32	0.83	0.30	0.85	14	18	0.26	0.17	0.62
C	20	3	0.37	0.03	0.82	11	12	0.20	0.11	0.64	20	53	0.37	0.49	0.36
L	47	27	0.87	0.25	0.88	38	22	0.70	0.20	0.83	29	28	0.54	0.26	0.72
#2	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	48	21	0.89	0.19	0.91	44	41	0.81	0.38	0.81	45	39	0.83	0.36	0.83
C	18	2	0.33	0.02	0.82	3	0	0.06	0.00	0.76	6	13	0.11	0.12	0.48
L	34	25	0.63	0.23	0.79	33	41	0.61	0.38	0.69	28	30	0.52	0.28	0.70
#3	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	44	19	0.81	0.18	0.89	38	52	0.70	0.48	0.69	30	45	0.56	0.42	0.62
C	12	0	0.22	0.00	0.81	9	12	0.17	0.11	0.60	13	24	0.24	0.22	0.53
L	47	37	0.87	0.34	0.85	26	27	0.48	0.25	0.70	25	24	0.46	0.22	0.71
#4	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	39	19	0.72	0.18	0.85	22	35	0.41	0.32	0.58	13	25	0.24	0.23	0.51
C	24	16	0.44	0.15	0.75	12	32	0.22	0.30	0.39	19	36	0.35	0.33	0.52
L	36	25	0.67	0.23	0.80	28	34	0.52	0.31	0.67	33	36	0.61	0.33	0.72
#5	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	32	6	0.59	0.06	0.87	18	15	0.33	0.14	0.70	19	27	0.35	0.25	0.61
C	19	23	0.35	0.21	0.64	28	44	0.52	0.41	0.60	19	31	0.35	0.29	0.57
L	41	23	0.76	0.21	0.85	24	25	0.44	0.23	0.69	31	23	0.57	0.21	0.77
#6	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	48	6	0.89	0.06	0.95	31	34	0.57	0.31	0.71	38	41	0.70	0.38	0.75
C	53	5	0.98	0.05	0.98	19	30	0.35	0.28	0.58	28	23	0.52	0.21	0.74
L	43	6	0.80	0.06	0.93	22	26	0.41	0.24	0.66	17	15	0.31	0.14	0.69
#7	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	51	12	0.94	0.11	0.95	31	41	0.57	0.38	0.66	37	8	0.69	0.07	0.89
C	37	8	0.69	0.07	0.89	9	29	0.17	0.27	0.31	9	9	0.17	0.08	0.65
L	43	9	0.80	0.08	0.92	20	31	0.37	0.29	0.59	50	49	0.93	0.45	0.84
#8	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	42	2	0.78	0.02	0.94	21	18	0.39	0.17	0.71	18	12	0.33	0.11	0.73
C	53	10	0.98	0.09	0.97	24	50	0.44	0.46	0.48	35	38	0.65	0.35	0.73
L	44	5	0.81	0.05	0.94	19	26	0.35	0.24	0.62	31	28	0.57	0.26	0.74
#9	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	51	7	0.94	0.06	0.97	31	37	0.57	0.34	0.69	34	14	0.63	0.13	0.84
C	49	0	0.91	0.00	0.98	12	27	0.22	0.25	0.46	24	16	0.44	0.15	0.75

L	45	4	0.83	0.04	0.95	27	28	0.50	0.26	0.70	45	29	0.83	0.27	0.86

#10	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	45	9	0.83	0.08	0.93	33	32	0.61	0.30	0.74	21	1	0.39	0.01	0.84
C	46	0	0.85	0.00	0.96	19	21	0.35	0.19	0.66	20	22	0.37	0.20	0.66
L	48	13	0.89	0.12	0.93	27	25	0.50	0.23	0.72	48	46	0.89	0.43	0.83

Table 6.3: Hit, False Alarm, and A-Prime totals for all English listeners, in response to all Dutch stimuli

	Burst Stimuli					No Burst Stimuli					Nasal Stimuli				
#1	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	30	2	0.56	0.02	0.88	28	20	0.52	0.19	0.76	24	7	0.44	0.06	0.82
C	48	14	0.89	0.13	0.93	35	17	0.65	0.16	0.83	44	55	0.81	0.51	0.75
L	46	21	0.85	0.19	0.90	36	22	0.67	0.20	0.82	27	4	0.50	0.04	0.85
#2	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	29	6	0.54	0.06	0.85	30	26	0.56	0.24	0.75	47	19	0.87	0.18	0.91
C	39	8	0.72	0.07	0.90	23	14	0.43	0.13	0.76	34	11	0.63	0.10	0.86
L	41	39	0.76	0.36	0.79	33	33	0.61	0.31	0.74	46	4	0.85	0.04	0.95
#3	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	33	2	0.61	0.02	0.89	18	25	0.33	0.23	0.61	40	30	0.74	0.28	0.82
C	49	7	0.91	0.06	0.96	27	25	0.50	0.23	0.72	30	16	0.56	0.15	0.80
L	52	18	0.96	0.17	0.95	33	32	0.61	0.30	0.74	41	5	0.76	0.05	0.92
#4	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	33	25	0.61	0.23	0.78	12	40	0.22	0.37	0.27	14	21	0.26	0.19	0.58
C	32	21	0.59	0.19	0.79	17	36	0.31	0.33	0.48	31	49	0.57	0.45	0.61
L	36	14	0.67	0.13	0.86	18	39	0.33	0.36	0.47	19	27	0.35	0.25	0.61
#5	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	13	3	0.24	0.03	0.78	14	17	0.26	0.16	0.63	20	7	0.37	0.06	0.79
C	36	36	0.67	0.33	0.75	37	43	0.69	0.40	0.72	41	40	0.76	0.37	0.78
L	32	22	0.59	0.20	0.79	22	26	0.41	0.24	0.66	28	14	0.52	0.13	0.80
#6	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	19	6	0.35	0.06	0.79	15	15	0.28	0.14	0.67	37	19	0.69	0.18	0.84
C	51	29	0.94	0.27	0.91	42	55	0.78	0.51	0.72	44	23	0.81	0.21	0.88
L	32	25	0.59	0.23	0.77	18	17	0.33	0.16	0.68	30	9	0.56	0.08	0.84
#7	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	29	7	0.54	0.06	0.85	19	36	0.35	0.33	0.52	31	15	0.57	0.14	0.82
C	33	19	0.61	0.18	0.81	19	38	0.35	0.35	0.50	24	16	0.44	0.15	0.75
L	44	30	0.81	0.28	0.85	18	31	0.33	0.29	0.55	48	28	0.89	0.26	0.89
#8	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	28	8	0.52	0.07	0.83	17	20	0.31	0.19	0.64	32	6	0.59	0.06	0.87
C	46	28	0.85	0.26	0.87	33	42	0.61	0.39	0.68	49	24	0.91	0.22	0.91
L	42	7	0.78	0.06	0.92	23	24	0.43	0.22	0.69	48	1	0.89	0.01	0.97
#9	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	51	22	0.94	0.20	0.93	23	43	0.43	0.40	0.53	39	4	0.72	0.04	0.92
C	47	7	0.87	0.06	0.95	18	36	0.33	0.33	0.50	44	10	0.81	0.09	0.92

L	32	3	0.59	0.03	0.88	18	23	0.33	0.21	0.63	52	12	0.96	0.11	0.96
#10	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	21	6	0.39	0.06	0.80	10	32	0.19	0.30	0.31	15	4	0.28	0.04	0.78
C	42	6	0.78	0.06	0.92	31	32	0.57	0.30	0.72	54	37	1.00	0.34	0.91
L	49	38	0.91	0.35	0.87	23	30	0.43	0.28	0.64	45	4	0.83	0.04	0.95

Table 6.4: Hit, False Alarm, and A-Prime totals for all English listeners, in response to all English stimuli

main effect for language of speaker (F = 16.52, df = 1,18, p = .005). Post-hoc t-tests (in Table 6.6) reveal that listeners were able to reproduce the English stimuli more accurately than the Dutch stimuli. A similar language of speaker effect emerged in the AX discrimination task, as well, and may be the result of the Dutch participants' familiarity with the English language.

	F	df	Sig.
Speaker	10.52	1,18	**0.010**
Speaker * Listener	1.601	1,18	0.222
Manner	24.03	2,17	**< .001**
Manner * Listener	3.968	2,17	**0.040**
Place	5.614	2,17	**0.010**
Place * Listener	0.772	2,17	0.478
Speaker * Manner	64.73	2,17	**< .001**
Speaker * Manner * Listener	0.111	2,17	0.895
Speaker * Place	20.69	2,17	**< .001**
Speaker * Place * Listener	4.051	2,17	**0.040**
Manner * Place	8.328	4,15	**< .001**
Manner * Place * Listener	0.305	4,15	0.870
Speaker * Manner * Place	0.394	4,15	0.810
Speaker * Manner * Place * Listener	1.405	4,15	0.280
Listener	4.086	1,18	0.058

Table 6.5: Results of ANOVA on A-prime data for all stimuli

Speaker	Mean A'	T-test
Dutch	0.799	**<u>0.005</u>**
English	0.829	---

		T-tests	
Manner	Mean A'	vs. No Burst	vs. Nasal
Burst	0.866	**0.007**	< **.001**
No Burst	0.834	---	< **.001**
Nasal	0.742	---	---

Manner * Listener

Manner	Listener	Mean A'	T-tests
Burst	Dutch	0.844	0.064
	English	0.888	---
No Burst	Dutch	0.837	0.656
	English	0.831	---
Nasal	Dutch	0.713	0.115
	English	0.770	---

Listener * Manner

			T-tests	
Listener	Manner	Mean A'	vs. No Burst	vs. Nasal
Dutch	Burst	0.844	0.675	**0.002**
	No Burst	0.837	---	**0.001**
	Nasal	0.713	---	---
English	Burst	0.888	**0.001**	< **.001**
	No Burst	0.831	---	**0.023**
	Nasal	0.770	---	---

		T-tests	
Place	Mean A'	vs. Coronal	vs. Labial
Dorsal	0.828	**0.003**	0.198
Coronal	0.795	---	**0.031**
Labial	0.818	---	---

Table 6.6: Results of post-hoc t-tests on significant effects in A' ANOVA for all stimuli

(Table 6.6 continued)

Speaker * Manner

Speaker	Manner	Mean A'	T-tests vs. No Burst	T-tests vs. Nasal
Dutch	Burst	0.882	**<u>0.006</u>**	< **<u>.001</u>**
	No Burst	0.836	---	< **<u>.001</u>**
	Nasal	0.678	---	---
English	Burst	0.850	0.146	**<u>0.018</u>**
	No Burst	0.831	---	0.144
	Nasal	0.805	---	---

Manner * Speaker

Manner	Speaker	Mean A'	T-tests
Burst	Dutch	0.882	0.064
	English	0.850	
No Burst	Dutch	0.836	0.587
	English	0.831	
Nasal	Dutch	0.678	< **<u>.001</u>**
	English	0.805	

Speaker * Place

Speaker	Place	Mean A'	T-tests vs. Coronal	T-tests vs. Labial
Dutch	Dorsal	0.830	< **<u>.001</u>**	0.065
	Coronal	0.761	---	**<u>0.023</u>**
	Labial	0.804	---	---
English	Dorsal	0.825	0.705	0.470
	Coronal	0.828	---	0.506
	Labial	0.833	---	---

Place * Speaker

Place	Speaker	Mean A'	T-tests
Dorsal	Dutch	0.830	0.649
	English	0.825	
Coronal	Dutch	0.761	**<u>0.001</u>**
	English	0.828	
Labial	Dutch	0.804	**<u>0.022</u>**
	English	0.833	

Table 6.6: Results of post-hoc t-tests on significant effects in A' ANOVA for all stimuli

(Table 6.6 continued)

Speaker * Listener * Place

Speaker	Listener	Place	Mean A'	T-tests vs. Coronal	T-tests vs. Labial
Dutch	Dutch	Dorsal	0.819	**0.005**	0.152
		Coronal	0.754	---	0.158
		Labial	0.792	---	---
	English	Dorsal	0.842	**0.003**	0.109
		Coronal	0.768	---	**0.024**
		Labial	0.816	---	---
English	Dutch	Dorsal	0.816	0.062	0.805
		Coronal	0.792	---	**0.018**
		Labial	0.813	---	---
	English	Dorsal	0.834	0.063	0.172
		Coronal	0.865	---	0.343
		Labial	0.853	---	---

Speaker * Place * Listener

Speaker	Place	Listener	Mean A'	T-tests
Dutch	Dorsal	Dutch	0.819	0.259
		English	0.842	
	Coronal	Dutch	0.754	0.586
		English	0.768	
	Labial	Dutch	0.792	0.372
		English	0.816	
English	Dorsal	Dutch	0.816	0.457
		English	0.834	
	Coronal	Dutch	0.792	**0.004**
		English	0.865	
	Labial	Dutch	0.813	**0.044**
		English	0.853	

Table 6.6: Results of post-hoc t-tests on significant effects in A’ ANOVA for all stimuli

(Table 6.6 continued)

Listener * Place * Speaker

Listener	Place	Speaker	Mean A'	T-tests
Dutch	Dorsal	Dutch	0.819	0.835
		English	0.816	
	Coronal	Dutch	0.754	**0.045**
		English	0.792	
	Labial	Dutch	0.792	0.246
		English	0.813	
English	Dorsal	Dutch	0.842	0.632
		English	0.834	
	Coronal	Dutch	0.768	**0.001**
		English	0.865	
	Labial	Dutch	0.816	0.080
		English	0.853	

Manner * Place

Manner	Place	Mean A'	T-tests vs. Coronal	vs. Labial
Burst	Dorsal	0.882	0.136	0.188
	Coronal	0.851	---	0.388
	Labial	0.864	---	---
No Burst	Dorsal	0.842	0.521	**0.008**
	Coronal	0.837	---	**0.003**
	Labial	0.822	---	---
Nasal	Dorsal	0.759	< **.001**	0.522
	Coronal	0.696	---	**0.001**
	Labial	0.769	---	---

Table 6.6: Results of post-hoc t-tests on significant effects in A’ ANOVA for all stimuli

(Table 6.6 continued)

Place * Manner			T-tests	
Place	Manner	Mean A'	vs. No Burst	vs. Nasal
Dorsal	Burst	0.882	< **.001**	**0.001**
	No Burst	0.842	---	**0.005**
	Nasal	0.759	---	---
Coronal	Burst	0.851	0.409	< **.001**
	No Burst	0.837	---	< **.001**
	Nasal	0.696	---	---
Labial	Burst	0.864	**0.007**	< **.001**
	No Burst	0.822	---	**0.004**
	Nasal	0.769	---	---

Table 6.6: Results of post-hoc t-tests on significant effects in A’ ANOVA for all stimuli

As in the AX discrimination experiment, the language of speaker effect also interacts with place of articulation. There is a main effect for place of articulation in this analysis of variance ($F = 5.614$, $df = 2,17$, $p = .013$). The post-hoc t-tests reveal that speakers of both languages were able to reproduce dorsals and labials more accurately than coronals. To a certain extent, the direction of this effect reflects the pattern of results found in the earlier perception studies, which revealed coronal to be the most confusable place of articulation in both Dutch and English. These perceptual studies also revealed, however, that the place of articulation which was most often confused with coronals depended on the language of the speaker producing the coronals. A different pattern emerges in this data. While the ANOVA reveals a significant interaction between the language of the speaker and place of articulation ($F = 20.69$, $df = 2,17$, $p < .001$), post-hoc t-tests show that significant differences between A' values for particular places of articulation exists only for the Dutch stimuli. For these stimuli, listeners reproduced the dorsal place of articulation most accurately, followed by the labial and then the coronal place of articulation. For the English stimuli, there were no significant differences between the reproduction accuracy of any of the three places of articulation. This suggests, again, that the unfamiliarity of English participants with the Dutch stimuli may have significantly decreased the overall accuracy of the reproduction of these stimuli.

However, analysis of the significant 3-way interaction between language of speaker, place of articulation, and language of listener ($F = 4.051$, $df = 2,17$, $p = .036$) reveals that the Dutch repeaters had difficulty reproducing some of the Dutch stimuli, as well. Both groups of participants reproduced Dutch coronals less accurately than either

Dutch dorsals or labials. Furthermore, there were no significant effects of the language of the listener within this set of stimuli. Such effects only emerged in the English stimuli, where the Dutch listeners were significantly worse than the English listeners at reproducing the coronal and labial places of articulation. This effect may perhaps not be unexpected, since the English listeners should presumably be better at reproducing stimuli in their own language than Dutch speakers are. The absence of a similar effect for the Dutch stimuli is therefore surprising; this may reflect cross-linguistic difficulties in perceiving the Dutch stimuli, or perhaps even problems the English-speaking transcriber had in transcribing the spoken Dutch responses correctly. Investigating how language of speaker effects interact with place of articulation, however, does indicate that both groups of listeners were worse at reproducing the Dutch coronals than they were at reproducing English coronals. This effect suggests, once again, that the general perceptual confusability of Dutch coronals may account for the inaccuracy of their reproduction in the repetition/imitation task.

Other significant effects in this analysis of variance, however, do not correspond to similar significant effects in the earlier perception studies. These effects therefore seem to have strictly articulatory origins. For instance, the direction of the main effect for manner of articulation ($F = 24.03$, $df = 2,17$, $p < .001$) differs from that found in the magnitude estimation and AX discrimination experiments. Post-hoc t-tests reveal that repetition accuracy was highest for stops with bursts, followed by stops without bursts, and, finally, worst for nasals. This differs from the pattern of results found in the perception-only studies, which failed to yield any significant differences between the discriminability of place in nasals and stops without bursts. The poorer performance of nasals in this study therefore appears to be the result of articulatory difficulties the listeners had in reproducing them accurately. Such articulatory difficulties would provide a likely motivation for the cross-linguistic susceptibility of nasals to place assimilation.

A significant interaction in the ANOVA between the language of listener and manner of articulation ($F = 3.968$, $df = 2,17$, $p = .039$) seems to confirm this hypothesis. The post-hoc t-tests reveal that the accuracy of reproducing nasal place of articulation was significantly lower than the accuracy of reproducing stops without bursts, regardless of the language of the participant producing the repetitions. The cross-linguistic persistence of this effect suggests that it might be able to account for nasals' broader susceptibility to place assimilation in many languages of the world.

Articulatory difficulties with the production of place in nasal-stop clusters did, however, interact to some extent with perception. The analysis of variance in Table 6.5 reveals a significant interaction between the language of speaker and the manner of articulation factors ($F = 64.73$, $df = 2,17$, $p < .001$). Post-hoc t-tests show different directions of significance in the manner of articulation effect, depending on the language in which the stimuli were spoken. For the Dutch stimuli, repetition accuracy for nasals was significantly worse than the repetition accuracy for stops without bursts, which, in turn, was worse than the repetition accuracy for stops <u>with</u> bursts. The distinctions between A' values for different manners of articulation are less clear for the English stimuli, however; while stops with bursts had significantly higher repetition accuracy than nasals, there were no further significant differences between repetition accuracies for the other pairs of manners of articulation. The lack of such a difference between nasals and stops without bursts in the English stimuli suggests that part of the difficulty in reproducing nasals accurately may stem from all listeners' inability to perceive the place of articulation of Dutch nasals correctly.

Manner of articulation also interacted significantly with place of articulation in the ANOVA (F = 8.328, df = 4,25, p = .001). For stimuli which had stops with bursts, repetition accuracy was relatively high for all three places of articulation. For stops without bursts, however, the labial place of articulation emerged as having the lowest repetition accuracy, while in nasals, speakers had the most difficulty reproducing coronals accurately. This last effect corresponds to the pattern of results found in the perception experiments, where coronals in both Dutch and English were highly confusable with other places of articulation. The difficulty of reproducing labials in the stop without burst stimuli has no perceptual correlate in the previous studies, however--even though it seems to have perceptual origins, since it only emerges once burst cues have been removed from the stop stimuli. It is therefore unclear whether the poor repetition accuracy for labials in the stop without burst stimuli is a result of speaker difficulty in making labial stops in consonant clusters, or simply the perceptual weakness of transition cues to labial place of articulation.

Breaking the A-prime response data down into finer detail, reveals further, unexpected differences in repetition accuracy for burst and no burst stimuli. Part of the rationale behind hypothesizing that articulatory difficulty might motivate nasals' cross-linguistic susceptibility to place assimilation was the recognition of the fact that nasal-stop heterorganic clusters require a greater amount of effort and articulatory coordination than stop-stop heterorganic clusters do. Likewise, heterorganic clusters <u>in general</u> should require greater articulatory effort and coordination than homorganic sequences of consonants do. The different levels of articulatory difficulty between the two cluster types supposedly motivates the process of place assimilation, which changes a heterorganic consonant cluster to a homorganic cluster, thereby yielding a consonant sequence that is easier for speakers to produce. Repetition accuracy--as measured in A-prime--should therefore be higher for the initial consonants in homorganic clusters than it is for the initial consonants in heterorganic clusters.

The results of this study, however, show that this is not always the case. Tables 6.7-6.10 present hit and false alarm totals, broken down by place of articulation, for responses to only those stimuli which had homorganic consonant clusters. These tables also list the corresponding A-prime values for each place of articulation in each stimulus type; an analysis of variance was run on these A-prime values in order to determine how four factors--language of speaker, language of listener, manner of articulation and place of articulation of the first consonant--affected repetition accuracy. The results of this ANOVA are given in Table 6.11. Table 6.12 lists the post-hoc t-test for the solitary significant effect (for manner of articulation) in this ANOVA.

	Burst Stimuli					No Burst Stimuli					Nasal Stimuli				
#1	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	17	11	0.94	0.31	0.90	16	0	0.89	0.00	0.97	16	11	0.89	0.31	0.87
C	9	0	0.50	0.00	0.88	18	1	1.00	0.03	0.99	14	2	0.78	0.06	0.92
L	7	9	0.39	0.25	0.64	18	1	1.00	0.03	0.99	9	2	0.50	0.06	0.84
#2	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	4	0	0.22	0.00	0.81	18	1	1.00	0.03	0.99	16	2	0.89	0.06	0.95
C	0	25	0.00	0.69	N/A	18	0	1.00	0.00	1.00	16	1	0.89	0.03	0.96
L	7	18	0.39	0.50	0.37	17	0	0.94	0.00	0.99	18	1	1.00	0.03	0.99
#3	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	5	5	0.28	0.14	0.67	13	8	0.72	0.22	0.83	12	18	0.67	0.50	0.65
C	0	13	0.00	0.36	N/A	12	5	0.67	0.14	0.85	9	6	0.50	0.17	0.77
L	5	12	0.28	0.33	0.43	12	3	0.67	0.08	0.88	9	0	0.50	0.00	0.88
#4	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	4	11	0.22	0.31	0.38	18	3	1.00	0.08	0.98	8	6	0.44	0.17	0.74
C	3	13	0.17	0.36	0.13	18	0	1.00	0.00	1.00	15	16	0.83	0.44	0.79
L	10	13	0.56	0.36	0.66	15	0	0.83	0.00	0.96	9	0	0.50	0.00	0.88
#5	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	15	9	0.83	0.25	0.87	15	3	0.83	0.08	0.93	13	4	0.72	0.11	0.88
C	3	3	0.17	0.08	0.65	18	2	1.00	0.06	0.99	17	5	0.94	0.14	0.95
L	14	10	0.78	0.28	0.83	14	0	0.78	0.00	0.94	15	0	0.83	0.00	0.96
#6	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	4	11	0.22	0.31	0.38	18	0	1.00	0.00	1.00	14	2	0.78	0.06	0.92
C	6	14	0.33	0.39	0.44	18	0	1.00	0.00	1.00	8	2	0.44	0.06	0.82
L	4	15	0.22	0.42	0.20	18	0	1.00	0.00	1.00	17	11	0.94	0.31	0.90
#7	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	8	8	0.44	0.22	0.70	18	0	1.00	0.00	1.00	15	4	0.83	0.11	0.92
C	15	18	0.83	0.50	0.77	18	1	1.00	0.03	0.99	12	7	0.67	0.19	0.82
L	4	0	0.22	0.00	0.81	17	0	0.94	0.00	0.99	10	6	0.56	0.17	0.79
#8	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	9	6	0.50	0.17	0.77	18	0	1.00	0.00	1.00	18	2	1.00	0.06	0.99
C	9	13	0.50	0.36	0.62	18	0	1.00	0.00	1.00	11	0	0.61	0.00	0.90
L	14	3	0.78	0.08	0.91	18	0	1.00	0.00	1.00	17	6	0.94	0.17	0.94
#9	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	0	13	0.00	0.36	N/A	6	2	0.33	0.06	0.78	9	1	0.50	0.03	0.86
C	10	14	0.56	0.39	0.64	12	5	0.67	0.14	0.85	8	7	0.44	0.19	0.72
L	2	10	0.11	0.28	0.07	16	13	0.89	0.36	0.85	10	18	0.56	0.50	0.55

#10	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	2	12	0.11	0.33	-0.08	17	0	0.94	0.00	0.99	17	0	0.94	0.00	0.99
C	13	15	0.72	0.42	0.74	18	0	1.00	0.00	1.00	16	2	0.89	0.06	0.95
L	2	7	0.11	0.19	0.29	18	1	1.00	0.03	0.99	17	2	0.94	0.06	0.97

Table 6.7: Hit, False Alarm, and A-Prime totals for all Dutch listeners, in response to all Dutch stimuli with homorganic clusters

	Burst Stimuli					No Burst Stimuli					Nasal Stimuli				
#1	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	4	0	0.22	0.00	0.81	13	0	0.72	0.00	0.93	16	6	0.89	0.17	0.92
C	9	18	0.50	0.50	0.50	18	1	1.00	0.03	0.99	16	3	0.89	0.08	0.95
L	7	16	0.39	0.44	0.44	17	5	0.94	0.14	0.95	13	0	0.72	0.00	0.93
#2	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	1	0	0.06	0.00	0.76	12	0	0.67	0.00	0.92	14	1	0.78	0.03	0.93
C	3	29	0.17	0.81	-1.28	18	7	1.00	0.19	0.95	17	5	0.94	0.14	0.95
L	6	15	0.33	0.42	0.40	17	0	0.94	0.00	0.99	15	0	0.83	0.00	0.96
#3	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	3	12	0.17	0.33	0.19	8	0	0.44	0.00	0.86	9	9	0.50	0.25	0.71
C	0	17	0.00	0.47	N/A	16	11	0.89	0.31	0.87	12	14	0.67	0.39	0.72
L	4	12	0.22	0.33	0.33	15	3	0.83	0.08	0.93	9	0	0.50	0.00	0.88
#4	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	5	9	0.28	0.25	0.53	9	1	0.50	0.03	0.86	17	0	0.94	0.00	0.99
C	4	29	0.22	0.81	-0.91	18	7	1.00	0.19	0.95	18	10	1.00	0.28	0.93
L	2	5	0.11	0.14	0.43	16	3	0.89	0.08	0.95	9	0	0.50	0.00	0.88
#5	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	18	2	1.00	0.06	0.99	12	0	0.67	0.00	0.92	14	0	0.78	0.00	0.94
C	15	3	0.83	0.08	0.93	18	0	1.00	0.00	1.00	18	4	1.00	0.11	0.97
L	13	3	0.72	0.08	0.90	18	6	1.00	0.17	0.96	17	0	0.94	0.00	0.99
#6	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	1	3	0.06	0.08	0.37	18	4	1.00	0.11	0.97	18	3	1.00	0.08	0.98
C	0	26	0.00	0.72	N/A	18	5	1.00	0.14	0.97	15	0	0.83	0.00	0.96
L	0	24	0.00	0.67	N/A	9	0	0.50	0.00	0.88	18	0	1.00	0.00	1.00
#7	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	5	11	0.28	0.31	0.47	18	7	1.00	0.19	0.95	13	4	0.72	0.11	0.88
C	8	17	0.44	0.47	0.47	17	0	0.94	0.00	0.99	15	9	0.83	0.25	0.87
L	12	0	0.67	0.00	0.92	12	0	0.67	0.00	0.92	13	0	0.72	0.00	0.93
#8	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	12	16	0.67	0.44	0.68	18	3	1.00	0.08	0.98	15	7	0.83	0.19	0.89
C	4	9	0.22	0.25	0.46	18	0	1.00	0.00	1.00	11	6	0.61	0.17	0.82
L	10	3	0.56	0.08	0.84	15	0	0.83	0.00	0.96	15	0	0.83	0.00	0.96
#9	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	0	12	0.00	0.33	N/A	2	7	0.11	0.19	0.29	4	0	0.22	0.00	0.81
C	2	20	0.11	0.56	-0.75	9	13	0.50	0.36	0.62	10	8	0.56	0.22	0.76
L	3	9	0.17	0.25	0.35	5	15	0.28	0.42	0.32	14	14	0.78	0.39	0.78

#10	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	3	13	0.17	0.36	0.13	18	6	1.00	0.17	0.96	18	2	1.00	0.06	0.99
C	1	15	0.06	0.42	-1.28	17	0	0.94	0.00	0.99	16	0	0.89	0.00	0.97
L	3	12	0.17	0.33	0.19	13	0	0.72	0.00	0.93	18	0	1.00	0.00	1.00

Table 6.8: Hit, False Alarm, and A-Prime totals for all Dutch listeners, in response to English stimuli with homorganic clusters

	Burst Stimuli					No Burst Stimuli					Nasal Stimuli				
#1	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	15	7	0.83	0.19	0.89	18	4	1.00	0.11	0.97	2	6	0.11	0.17	0.36
C	3	3	0.17	0.08	0.65	9	0	0.50	0.00	0.88	6	19	0.33	0.53	0.25
L	11	12	0.61	0.33	0.72	15	8	0.83	0.22	0.88	12	9	0.67	0.25	0.80
#2	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	15	20	0.83	0.56	0.74	11	15	0.61	0.42	0.66	14	15	0.78	0.42	0.77
C	0	0	0.00	0.00	N/A	0	0	0.00	0.00	N/A	1	3	0.06	0.08	0.37
L	1	15	0.06	0.42	-1.28	9	19	0.50	0.53	0.47	7	14	0.39	0.39	0.50
#3	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	11	13	0.61	0.36	0.70	11	9	0.61	0.25	0.77	6	16	0.33	0.44	0.37
C	0	0	0.00	0.00	N/A	11	2	0.61	0.06	0.87	7	5	0.39	0.14	0.73
L	11	19	0.61	0.53	0.58	12	9	0.67	0.25	0.80	13	7	0.72	0.19	0.85
#4	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	12	3	0.67	0.08	0.88	18	1	1.00	0.03	0.99	12	1	0.67	0.03	0.90
C	9	6	0.50	0.17	0.77	16	0	0.89	0.00	0.97	17	5	0.94	0.14	0.95
L	15	9	0.83	0.25	0.87	18	1	1.00	0.03	0.99	17	2	0.94	0.06	0.97
#5	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	3	6	0.17	0.17	0.50	9	0	0.50	0.00	0.88	9	6	0.50	0.17	0.77
C	0	21	0.00	0.58	N/A	13	11	0.72	0.31	0.79	9	8	0.50	0.22	0.73
L	8	13	0.44	0.36	0.58	16	5	0.89	0.14	0.93	16	3	0.89	0.08	0.95
#6	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	15	6	0.83	0.17	0.90	18	4	1.00	0.11	0.97	17	8	0.94	0.22	0.92
C	17	0	0.94	0.00	0.99	17	0	0.94	0.00	0.99	18	1	1.00	0.03	0.99
L	12	3	0.67	0.08	0.88	15	0	0.83	0.00	0.96	9	1	0.50	0.03	0.86
#7	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	16	7	0.89	0.19	0.91	18	2	1.00	0.06	0.99	14	0	0.78	0.00	0.94
C	9	8	0.50	0.22	0.73	15	0	0.83	0.00	0.96	9	2	0.50	0.06	0.84
L	8	5	0.44	0.14	0.76	16	3	0.89	0.08	0.95	18	11	1.00	0.31	0.92
#8	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	10	2	0.56	0.06	0.86	12	0	0.67	0.00	0.92	9	0	0.50	0.00	0.88
C	18	10	1.00	0.28	0.93	18	9	1.00	0.25	0.94	18	9	1.00	0.25	0.94
L	11	0	0.61	0.00	0.90	15	0	0.83	0.00	0.96	15	3	0.83	0.08	0.93
#9	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	16	7	0.89	0.19	0.91	18	0	1.00	0.00	1.00	18	4	1.00	0.11	0.97
C	14	0	0.78	0.00	0.94	18	0	1.00	0.00	1.00	14	0	0.78	0.00	0.94
L	9	3	0.50	0.08	0.82	18	0	1.00	0.00	1.00	18	0	1.00	0.00	1.00

#10	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	12	9	0.67	0.25	0.80	14	3	0.78	0.08	0.91	9	0	0.50	0.00	0.88
C	11	0	0.61	0.00	0.90	15	3	0.83	0.08	0.93	13	3	0.72	0.08	0.90
L	12	10	0.67	0.28	0.78	16	1	0.89	0.03	0.96	18	7	1.00	0.19	0.95

Table 6.9: Hit, False Alarm, and A-Prime totals for all English listeners, in response to Dutch stimuli with homorganic clusters

	Burst Stimuli					No Burst Stimuli					Nasal Stimuli				
#1	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	8	1	0.44	0.03	0.84	12	0	0.67	0.00	0.92	1	3	0.06	0.08	0.37
C	12	11	0.67	0.31	0.77	18	6	1.00	0.17	0.96	14	26	0.78	0.72	0.57
L	13	9	0.72	0.25	0.82	18	0	1.00	0.00	1.00	9	1	0.50	0.03	0.86
#2	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	3	6	0.17	0.17	0.50	18	3	1.00	0.08	0.98	12	2	0.67	0.06	0.89
C	3	7	0.17	0.19	0.45	12	0	0.67	0.00	0.92	16	5	0.89	0.14	0.93
L	5	30	0.28	0.83	-0.83	15	6	0.83	0.17	0.90	16	3	0.89	0.08	0.95
#3	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	9	1	0.50	0.03	0.86	9	0	0.50	0.00	0.88	8	5	0.44	0.14	0.76
C	15	3	0.83	0.08	0.93	18	6	1.00	0.17	0.96	13	10	0.72	0.28	0.81
L	17	9	0.94	0.25	0.92	18	3	1.00	0.08	0.98	17	1	0.94	0.03	0.98
#4	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	16	0	0.89	0.00	0.97	12	0	0.67	0.00	0.92	12	1	0.67	0.03	0.90
C	18	2	1.00	0.06	0.99	18	0	1.00	0.00	1.00	12	11	0.67	0.31	0.77
L	18	0	1.00	0.00	1.00	18	6	1.00	0.17	0.96	12	6	0.67	0.17	0.84
#5	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	0	3	0.00	0.08	N/A	9	0	0.50	0.00	0.88	9	3	0.50	0.08	0.82
C	5	29	0.28	0.81	-0.65	12	13	0.67	0.36	0.73	15	12	0.83	0.33	0.84
L	0	13	0.00	0.36	N/A	14	6	0.78	0.17	0.88	10	0	0.56	0.00	0.89
#6	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	0	0	0.00	0.00	N/A	8	0	0.44	0.00	0.86	14	3	0.78	0.08	0.91
C	15	18	0.83	0.50	0.77	18	10	1.00	0.28	0.93	17	6	0.94	0.17	0.94
L	12	9	0.67	0.25	0.80	18	0	1.00	0.00	1.00	12	2	0.67	0.06	0.89
#7	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	5	3	0.28	0.08	0.73	15	0	0.83	0.00	0.96	16	3	0.89	0.08	0.95
C	4	8	0.22	0.22	0.50	18	3	1.00	0.08	0.98	15	2	0.83	0.06	0.94
L	18	16	1.00	0.44	0.89	18	0	1.00	0.00	1.00	18	0	1.00	0.00	1.00
#8	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	7	3	0.39	0.08	0.78	14	0	0.78	0.00	0.94	8	2	0.44	0.06	0.82
C	12	11	0.67	0.31	0.77	18	4	1.00	0.11	0.97	17	10	0.94	0.28	0.91
L	18	3	1.00	0.08	0.98	18	0	1.00	0.00	1.00	15	0	0.83	0.00	0.96
#9	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	18	10	1.00	0.28	0.93	18	0	1.00	0.00	1.00	15	0	0.83	0.00	0.96
C	11	1	0.61	0.03	0.89	18	0	1.00	0.00	1.00	17	1	0.94	0.03	0.98
L	14	0	0.78	0.00	0.94	18	0	1.00	0.00	1.00	16	4	0.89	0.11	0.94

#10	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	2	3	0.11	0.08	0.57	10	0	0.56	0.00	0.89	6	3	0.33	0.08	0.76
C	6	6	0.33	0.17	0.68	18	6	1.00	0.17	0.96	18	12	1.00	0.33	0.92
L	16	21	0.89	0.58	0.77	17	3	0.94	0.08	0.96	15	0	0.83	0.00	0.96

Table 6.10: Hit, False Alarm, and A-Prime totals for all English listeners, in response to English stimuli with homorganic clusters

	F	df	Sig.
Speaker	2.415	1,10	0.151
Speaker * Listener	3.2	1,10	0.104
Manner	14.48	2,9	< **.001**
Manner * Listener	2.619	2,9	0.127
Place	1.789	2,9	0.222
Place * Listener	1.363	2,9	0.304
Speaker * Manner	2.662	2,9	0.124
Speaker * Manner * Listener	2.05	2,9	0.185
Speaker * Place	3.819	2,9	0.063
Speaker * Place * Listener	1.943	2,9	0.199
Manner * Place	1.516	4,7	0.296
Manner * Place * Listener	1.089	4,7	0.431
Speaker * Manner * Place	1.402	4,7	0.326
Speaker * Manner * Place * Listener	1.748	4,7	0.244
Listener	1.924	1,10	0.196

Table 6.11: Results of ANOVA on A-prime data for stimuli with homorganic clusters

		T-tests	
Manner	Mean A'	vs. No Burst	vs. Nasal
Burst	0.569	< **.001**	**0.004**
No Burst	0.925	---	**0.003**
Nasal	0.841	---	---

Table 6.12: Results of post-hoc t-tests on significant effects in ANOVA for stimuli with homorganic clusters

Tables 6.13-6.16 list similar hit and false alarm totals for responses to only the heterorganic stimuli in this experiment. These totals are also broken down by stimulus type and place of articulation, and are presented alongside their corresponding A-prime values. An analysis of variance was also run on these A-prime values, in order to determine the effects language of speaker, language of listener, place and manner of articulation of the C_1 consonant might have had on repetition accuracy for place in the first consonant of the heterorganic clusters alone. Table 6.17 shows the results of this analysis of variance.

	Burst Stimuli					No Burst Stimuli					Nasal Stimuli				
#1	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	35	3	0.97	0.04	0.98	16	5	0.44	0.07	0.81	24	37	0.67	0.51	0.64
C	25	3	0.69	0.04	0.91	4	1	0.11	0.01	0.74	0	10	0.00	0.14	N/A
L	33	7	0.92	0.10	0.95	6	1	0.17	0.01	0.77	24	11	0.67	0.15	0.84
#2	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	36	0	1.00	0.00	1.00	24	1	0.67	0.01	0.91	1	24	0.03	0.33	-2.36
C	36	0	1.00	0.00	1.00	9	0	0.25	0.00	0.81	5	35	0.14	0.49	-0.29
L	35	0	0.97	0.00	0.99	6	1	0.17	0.01	0.77	8	34	0.22	0.47	0.10
#3	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	36	1	1.00	0.01	1.00	29	9	0.81	0.13	0.91	28	34	0.78	0.47	0.74
C	6	0	0.17	0.00	0.79	8	0	0.22	0.00	0.81	10	14	0.28	0.19	0.60
L	30	9	0.83	0.13	0.91	18	0	0.50	0.00	0.88	15	6	0.42	0.08	0.79
#4	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	33	3	0.92	0.04	0.97	14	0	0.39	0.00	0.85	12	24	0.33	0.33	0.50
C	27	0	0.75	0.00	0.94	8	6	0.22	0.08	0.69	9	25	0.25	0.35	0.37
L	36	3	1.00	0.04	0.99	9	1	0.25	0.01	0.80	15	10	0.42	0.14	0.75
#5	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	35	5	0.97	0.07	0.97	9	3	0.25	0.04	0.76	0	30	0.00	0.42	N/A
C	21	3	0.58	0.04	0.87	0	0	0.00	0.00	N/A	0	34	0.00	0.47	N/A
L	22	5	0.61	0.07	0.87	3	0	0.08	0.00	0.77	7	36	0.19	0.50	-0.05
#6	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	36	0	1.00	0.00	1.00	13	0	0.36	0.00	0.84	21	12	0.58	0.17	0.80
C	30	0	0.83	0.00	0.96	7	0	0.19	0.00	0.80	0	13	0.00	0.18	N/A
L	36	0	1.00	0.00	1.00	7	4	0.19	0.06	0.72	29	33	0.81	0.46	0.77
#7	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	34	10	0.94	0.14	0.95	13	5	0.36	0.07	0.78	6	20	0.17	0.28	0.29
C	36	2	1.00	0.03	0.99	1	3	0.03	0.04	0.37	13	22	0.36	0.31	0.56
L	24	0	0.67	0.00	0.92	0	0	0.00	0.00	N/A	21	17	0.58	0.24	0.76
#8	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	36	1	1.00	0.01	1.00	5	0	0.14	0.00	0.78	6	10	0.17	0.14	0.55
C	33	0	0.92	0.00	0.98	3	6	0.08	0.08	0.50	3	14	0.08	0.19	0.13
L	35	2	0.97	0.03	0.99	3	1	0.08	0.01	0.73	26	49	0.72	0.68	0.55
#9	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	29	8	0.81	0.11	0.91	12	0	0.33	0.00	0.83	10	9	0.28	0.13	0.68
C	31	5	0.86	0.07	0.94	17	10	0.47	0.14	0.77	5	12	0.14	0.17	0.44
L	25	2	0.69	0.03	0.91	11	10	0.31	0.14	0.68	24	41	0.67	0.57	0.59

#10	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	33	0	0.92	0.00	0.98	11	1	0.31	0.01	0.81	30	37	0.83	0.51	0.76
C	36	6	1.00	0.08	0.98	2	0	0.06	0.00	0.76	1	12	0.03	0.17	-0.79
L	26	0	0.72	0.00	0.93	2	0	0.06	0.00	0.76	4	24	0.11	0.33	-0.08

Table 6.13: Hit, False Alarm, and A-Prime totals for all Dutch listeners, in response to Dutch stimuli with heterorganic clusters

	Burst Stimuli					No Burst Stimuli					Nasal Stimuli				
#1	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	20	0	0.56	0.00	0.89	10	3	0.28	0.04	0.77	32	27	0.89	0.38	0.85
C	29	9	0.81	0.13	0.91	8	2	0.22	0.03	0.77	10	8	0.28	0.11	0.70
L	29	14	0.81	0.19	0.88	8	3	0.22	0.04	0.75	17	11	0.47	0.15	0.76
#2	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	24	2	0.67	0.03	0.90	7	3	0.19	0.04	0.74	7	33	0.19	0.46	0.04
C	35	16	0.97	0.22	0.93	13	9	0.36	0.13	0.73	9	37	0.25	0.51	0.10
L	18	6	0.50	0.08	0.82	7	3	0.19	0.04	0.74	4	15	0.11	0.21	0.25
#3	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	28	1	0.78	0.01	0.94	21	3	0.58	0.04	0.87	29	26	0.81	0.36	0.81
C	33	10	0.92	0.14	0.94	28	6	0.78	0.08	0.91	16	13	0.44	0.18	0.73
L	27	6	0.75	0.08	0.90	16	6	0.44	0.08	0.80	15	0	0.42	0.00	0.85
#4	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	30	0	0.83	0.00	0.96	13	4	0.36	0.06	0.79	22	19	0.61	0.26	0.76
C	36	10	1.00	0.14	0.97	20	5	0.56	0.07	0.85	13	23	0.36	0.32	0.54
L	24	4	0.67	0.06	0.89	9	0	0.25	0.00	0.81	13	14	0.36	0.19	0.67
#5	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	25	2	0.69	0.03	0.91	1	0	0.03	0.00	0.76	6	19	0.17	0.26	0.32
C	24	0	0.67	0.00	0.92	0	2	0.00	0.03	N/A	13	40	0.36	0.56	0.26
L	18	7	0.50	0.10	0.81	3	1	0.08	0.01	0.73	9	19	0.25	0.26	0.48
#6	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	27	0	0.75	0.00	0.94	15	1	0.42	0.01	0.84	28	19	0.78	0.26	0.84
C	30	0	0.83	0.00	0.96	14	6	0.39	0.08	0.78	4	8	0.11	0.11	0.50
L	36	12	1.00	0.17	0.96	18	0	0.50	0.00	0.88	32	17	0.89	0.24	0.90
#7	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	32	6	0.89	0.08	0.95	21	7	0.58	0.10	0.84	13	20	0.36	0.28	0.59
C	35	17	0.97	0.24	0.93	10	2	0.28	0.03	0.79	23	30	0.64	0.42	0.68
L	11	0	0.31	0.00	0.83	0	0	0.00	0.00	N/A	17	1	0.47	0.01	0.86
#8	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	26	9	0.72	0.13	0.88	15	3	0.42	0.04	0.82	12	17	0.33	0.24	0.60
C	30	0	0.83	0.00	0.96	9	0	0.25	0.00	0.81	15	16	0.42	0.22	0.68
L	21	3	0.58	0.04	0.87	3	0	0.08	0.00	0.77	26	16	0.72	0.22	0.83
#9	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	26	1	0.72	0.01	0.92	12	3	0.33	0.04	0.79	8	8	0.22	0.11	0.66
C	30	6	0.83	0.08	0.93	25	15	0.69	0.21	0.83	13	12	0.36	0.17	0.69
L	29	4	0.81	0.06	0.93	18	9	0.50	0.13	0.79	29	26	0.81	0.36	0.81

#10	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	27	0	0.75	0.00	0.94	10	10	0.28	0.14	0.67	31	26	0.86	0.36	0.84
C	36	0	1.00	0.00	1.00	8	3	0.22	0.04	0.75	7	6	0.19	0.08	0.67
L	36	3	1.00	0.04	0.99	10	6	0.28	0.08	0.73	25	9	0.69	0.13	0.87

Table 6.14: Hit, False Alarm, and A-Prime totals for all Dutch listeners, in response to English stimuli with heterorganic clusters

	Burst Stimuli					No Burst Stimuli					Nasal Stimuli				
#1	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	33	1	0.92	0.01	0.98	27	7	0.75	0.10	0.90	12	12	0.33	0.17	0.68
C	17	0	0.47	0.00	0.87	8	3	0.22	0.04	0.75	14	34	0.39	0.47	0.41
L	36	12	1.00	0.17	0.96	23	11	0.64	0.15	0.83	17	19	0.47	0.26	0.68
#2	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	33	0	0.92	0.00	0.98	33	12	0.92	0.17	0.93	31	24	0.86	0.33	0.85
C	18	0	0.50	0.00	0.88	3	0	0.08	0.00	0.77	5	8	0.14	0.11	0.56
L	33	10	0.92	0.14	0.94	24	17	0.67	0.24	0.80	21	13	0.58	0.18	0.80
#3	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	33	6	0.92	0.08	0.95	27	18	0.75	0.25	0.83	24	27	0.67	0.38	0.73
C	12	0	0.33	0.00	0.83	1	1	0.03	0.01	0.63	6	18	0.17	0.25	0.35
L	36	15	1.00	0.21	0.95	14	9	0.39	0.13	0.75	12	17	0.33	0.24	0.60
#4	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	27	4	0.75	0.06	0.92	4	0	0.11	0.00	0.78	1	16	0.03	0.22	-1.31
C	15	0	0.42	0.00	0.85	1	0	0.03	0.00	0.76	2	22	0.06	0.31	-0.72
L	21	5	0.58	0.07	0.86	8	1	0.22	0.01	0.79	16	31	0.44	0.43	0.51
#5	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	29	0	0.81	0.00	0.95	6	0	0.17	0.00	0.79	10	20	0.28	0.28	0.50
C	19	0	0.53	0.00	0.88	21	7	0.58	0.10	0.84	10	23	0.28	0.32	0.45
L	33	9	0.92	0.13	0.94	8	1	0.22	0.01	0.79	15	20	0.42	0.28	0.63
#6	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	33	0	0.92	0.00	0.98	13	3	0.36	0.04	0.80	21	26	0.58	0.36	0.68
C	36	0	1.00	0.00	1.00	8	3	0.22	0.04	0.75	10	20	0.28	0.28	0.50
L	31	0	0.86	0.00	0.97	7	1	0.19	0.01	0.78	8	13	0.22	0.18	0.56
#7	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	35	4	0.97	0.06	0.98	13	1	0.36	0.01	0.83	23	2	0.64	0.03	0.90
C	28	0	0.78	0.00	0.94	0	3	0.00	0.04	N/A	0	2	0.00	0.03	N/A
L	35	0	0.97	0.00	0.99	4	0	0.11	0.00	0.78	32	36	0.89	0.50	0.80
#8	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	32	0	0.89	0.00	0.97	9	1	0.25	0.01	0.80	9	12	0.25	0.17	0.61
C	32	0	0.89	0.00	0.97	11	11	0.31	0.15	0.67	17	29	0.47	0.40	0.57
L	33	1	0.92	0.01	0.98	4	3	0.11	0.04	0.67	16	24	0.44	0.33	0.60
#9	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	35	0	0.97	0.00	0.99	13	4	0.36	0.06	0.79	16	8	0.44	0.11	0.78
C	35	0	0.97	0.00	0.99	0	0	0.00	0.00	N/A	10	16	0.28	0.22	0.57
L	36	0	1.00	0.00	1.00	9	0	0.25	0.00	0.81	27	29	0.75	0.40	0.76

#10	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	33	0	0.92	0.00	0.98	19	9	0.53	0.13	0.81	12	1	0.33	0.01	0.82
C	35	0	0.97	0.00	0.99	10	1	0.28	0.01	0.80	7	17	0.19	0.24	0.43
L	36	0	1.00	0.00	1.00	11	0	0.31	0.00	0.83	30	38	0.83	0.53	0.75

Table 6.15: Hit, False Alarm, and A-Prime totals for all English listeners, in response to Dutch stimuli with heterorganic clusters

	Burst Stimuli					No Burst Stimuli					Nasal Stimuli				
#1	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	22	1	0.61	0.01	0.90	16	0	0.44	0.00	0.86	23	4	0.64	0.06	0.88
C	36	2	1.00	0.03	0.99	23	3	0.64	0.04	0.89	30	29	0.83	0.40	0.81
L	33	6	0.92	0.08	0.95	18	7	0.50	0.10	0.81	18	3	0.50	0.04	0.85
#2	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	26	0	0.72	0.00	0.93	12	4	0.33	0.06	0.78	35	16	0.97	0.22	0.93
C	36	0	1.00	0.00	1.00	14	3	0.39	0.04	0.81	18	6	0.50	0.08	0.82
L	36	6	1.00	0.08	0.98	18	6	0.50	0.08	0.82	30	0	0.83	0.00	0.96
#3	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	24	1	0.67	0.01	0.91	9	5	0.25	0.07	0.73	32	25	0.89	0.35	0.86
C	34	3	0.94	0.04	0.97	15	0	0.42	0.00	0.85	17	6	0.47	0.08	0.81
L	35	6	0.97	0.08	0.97	15	5	0.42	0.07	0.80	24	4	0.67	0.06	0.89
#4	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	17	7	0.47	0.10	0.80	0	2	0.00	0.03	N/A	2	9	0.06	0.13	0.17
C	14	1	0.39	0.01	0.84	5	0	0.14	0.00	0.78	19	27	0.53	0.38	0.63
L	18	4	0.50	0.06	0.84	0	5	0.00	0.07	N/A	7	12	0.19	0.17	0.54
#5	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	13	0	0.36	0.00	0.84	5	0	0.14	0.00	0.78	11	3	0.31	0.04	0.78
C	31	5	0.86	0.07	0.94	25	8	0.69	0.11	0.87	26	28	0.72	0.39	0.75
L	32	6	0.89	0.08	0.95	8	8	0.22	0.11	0.66	18	14	0.50	0.19	0.75
#6	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	19	4	0.53	0.06	0.85	7	0	0.19	0.00	0.80	23	12	0.64	0.17	0.83
C	36	8	1.00	0.11	0.97	30	14	0.83	0.19	0.89	27	16	0.75	0.22	0.85
L	20	11	0.56	0.15	0.80	0	3	0.00	0.04	N/A	18	4	0.50	0.06	0.84
#7	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	24	0	0.67	0.00	0.92	4	4	0.11	0.06	0.64	15	4	0.42	0.06	0.81
C	29	6	0.81	0.08	0.92	7	1	0.19	0.01	0.78	9	10	0.25	0.14	0.64
L	26	4	0.72	0.06	0.91	0	1	0.00	0.01	N/A	30	17	0.83	0.24	0.87
#8	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	21	0	0.58	0.00	0.90	3	1	0.08	0.01	0.73	24	4	0.67	0.06	0.89
C	33	6	0.92	0.08	0.95	21	8	0.58	0.11	0.84	32	14	0.89	0.19	0.91
L	24	2	0.67	0.03	0.90	5	4	0.14	0.06	0.67	33	1	0.92	0.01	0.98
#9	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	33	6	0.92	0.08	0.95	5	3	0.14	0.04	0.70	24	4	0.67	0.06	0.89
C	36	2	1.00	0.03	0.99	6	3	0.17	0.04	0.72	27	7	0.75	0.10	0.90
L	18	0	0.50	0.00	0.88	0	1	0.00	0.01	N/A	36	5	1.00	0.07	0.98

#10	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'	H	FA	P(H)	P(FA)	A'
D	19	3	0.53	0.04	0.86	0	3	0.00	0.04	N/A	9	1	0.25	0.01	0.80
C	36	0	1.00	0.00	1.00	19	3	0.53	0.04	0.86	36	25	1.00	0.35	0.91
L	33	12	0.92	0.17	0.93	6	6	0.17	0.08	0.65	30	4	0.83	0.06	0.94

Table 6.16: Hit, False Alarm, and A-Prime totals for all English listeners, in response to English stimuli with heterorganic clusters

	F	df	Sig.
Speaker	20.975	1,9	**0.001**
Speaker * Listener	0.648	1,9	0.441
Manner	39.073	2,8	< **.001**
Manner * Listener	1.192	2,8	0.352
Place	10.045	2,8	**0.007**
Place * Listener	1.5	2,8	0.280
Speaker * Manner	6.499	2,8	**0.021**
Speaker * Manner * Listener	0.669	2,8	0.539
Speaker * Place	15.891	2,8	**0.002**
Speaker * Place * Listener	0.16	2,8	0.855
Manner * Place	18.034	4,6	**0.002**
Manner * Place * Listener	1.959	4,6	0.220
Speaker * Manner * Place	2.3	4,6	0.173
Speaker * Manner * Place * Listener	0.508	4,6	0.734
Listener	2.484	1,9	0.149

Table 6.17: Results of ANOVA on A-prime data for stimuli with heterorganic clusters

These two ANOVAs yield strikingly different results. The analysis of variance for the heterorganic cluster data yields significant effects which are almost identical to the significant effects in the ANOVA for all of the stimuli. There are significant main effects for language of speaker, manner of articulation and place of articulation, as well as significant interactions between speaker and manner, speaker and place, and manner and place. Two of the significant effects in the first ANOVA, however, disappear in the second: the interaction between manner and language of listener, as well as the speaker by place by listener interaction; the language-specific listener effects, in other words, cease to have an effect on A-prime values in the heterorganic cluster stimuli. Post-hoc t-tests (see Table 6.18) show that the direction of these effects is, in most cases, identical to the same effects in the general ANOVA. Crucially, the direction of the manner of articulation effect is the same: repetition accuracy is highest for stops with bursts, followed by stops without bursts, and worst for nasals. Figure 6.4 provides a graphic comparison of the manner of articulation effect for both the heterorganic stimuli alone (labeled XY in the Figure) and all stimuli, taken together. This figure shows that repetition accuracy is slightly lower for stops without bursts in heterorganic clusters –and is significantly lower for nasals in these clusters. Such an effect is to be expected, if heterorganic clusters are truly more difficult, articulatorily, to reproduce than homorganic clusters. This pattern of results therefore suggests that there is an articulatory motivation for both stops (without bursts) and nasals to undergo place assimilation. Moreover, this motivation is stronger in the case of nasals. Figure 6.4 also shows, however, that repetition accuracy changes in the opposite direction for stops with bursts in heterorganic clusters. For this class of sounds, repetition accuracy actually increases in heterorganic clusters. It therefore seems to be easier for speakers to produce stops with bursts in heterorganic clusters than in homorganic clusters.

Speaker	Mean A'	T-test
English	0.811	**0.002**
Dutch	0.729	---

Manner	Mean A'	T-tests vs. No Burst	T-tests vs. Nasal
Burst	0.934	< **.001**	< **.001**
No Burst	0.779	---	**0.014**
Nasal	0.587	---	---

Place	Mean A'	T-tests vs. Coronal	T-tests vs. Labial
Dorsal	0.768	0.336	0.353
Coronal	0.744	---	**0.011**
Labial	0.798	---	---

Speaker * Manner

Speaker	Manner	Mean A'	T-tests vs. No Burst	T-tests vs. Nasal
English	Burst	0.918	< **.001**	**0.001**
	No Burst	0.785	---	0.201
	Nasal	0.727	---	---
Dutch	Burst	0.951	< **.001**	< **.001**
	No Burst	0.774	---	**0.006**
	Nasal	0.440	---	---

Manner * Speaker

Manner	Speaker	Mean A'	T-tests
Burst	English	0.918	**0.008**
	Dutch	0.951	---
No Burst	English	0.785	0.526
	Dutch	0.774	---
Nasal	English	0.727	**0.001**
	Dutch	0.440	---

Table 6.18: Results of post-hoc t-tests on significant effects in ANOVA for stimuli with heterorganic clusters

(Table 6.18 continued)

Speaker * Place

Speaker	Place	Mean A'	T-tests vs. Coronal	T-tests vs. Labial
English	Dorsal	0.795	0.067	**<u>0.010</u>**
	Coronal	0.815	---	0.348
	Labial	0.825	---	---
Dutch	Dorsal	0.743	0.147	0.623
	Coronal	0.670	---	**<u>0.014</u>**
	Labial	0.772	---	---

Place * Speaker

Place	Speaker	Mean A'	T-tests
Dorsal	English	0.795	0.250
	Dutch	0.743	---
Coronal	English	0.815	**<u>0.003</u>**
	Dutch	0.670	---
Labial	English	0.825	**<u>0.011</u>**
	Dutch	0.772	---

Manner * Place

Manner	Place	Mean A'	T-tests vs. Coronal	T-tests vs. Labial
Burst	Dorsal	0.938	0.838	0.051
	Coronal	0.940	---	0.156
	Labial	0.926	---	---
No Burst	Dorsal	0.801	0.154	**<u>0.002</u>**
	Coronal	0.772	---	0.896
	Labial	0.770	---	---
Nasal	Dorsal	0.563	0.444	0.183
	Coronal	0.502	---	**<u>0.001</u>**
	Labial	0.691	---	---

Table 6.18: Results of post-hoc t-tests on significant effects in ANOVA for stimuli with heterorganic clusters

(Table 6.18 continued)

Place * Manner			T-tests	
Place	Manner	Mean A'	vs. No Burst	vs. Nasal
Dorsal	Burst	0.938	< **.001**	**0.006**
	No Burst	0.801	---	0.059
	Nasal	0.563	---	---
Coronal	Burst	0.940	< **.001**	< **.001**
	No Burst	0.772	---	**0.001**
	Nasal	0.502	---	---
Labial	Burst	0.926	< **.001**	< **.001**
	No Burst	0.770	---	0.071
	Nasal	0.691	---	---

Table 6.18: Results of post-hoc t-tests on significant effects in ANOVA for stimuli with heterorganic clusters

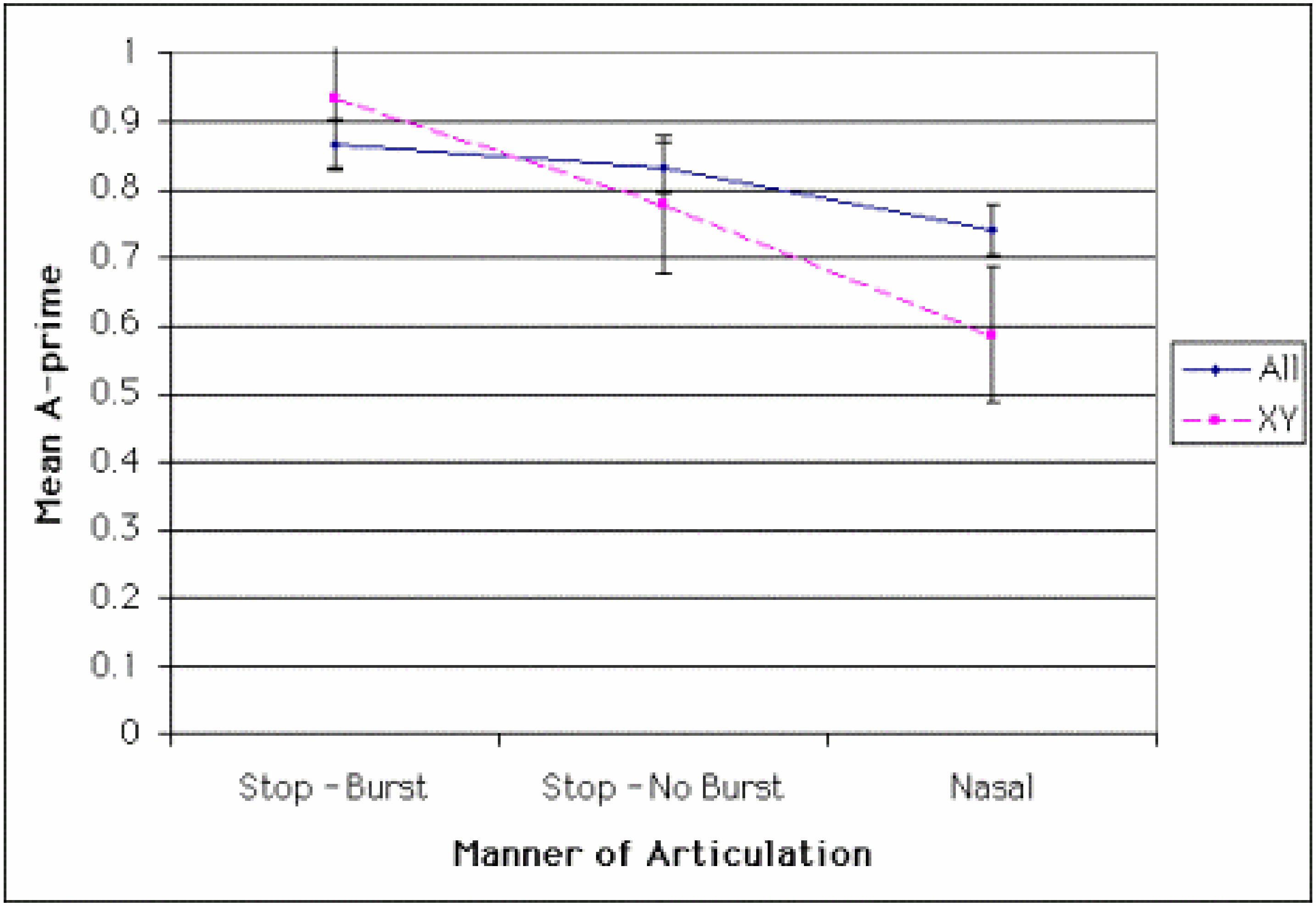


Figure 6.4: A-prime values, by manner and stimulus type

This curious effect of manner of articulation is the only significant effect (F = 14.48, df = 2,9, p = .002) found in the ANOVA for the homorganic cluster stimuli. Post-hoc t-tests show that repetition accuracy is best in this class of stimuli for stops without bursts, followed by nasals, and worst for stops with bursts. Since stops with bursts had the greatest perceptual distinctiveness in both the AX discrimination and magnitude estimation experiments, there seems to be no perceptual explanation for their dramatic decline in repetition accuracy in homorganic clusters (see Figure 6.5). Instead, the articulatory difficulty of repeating two homorganic stops with bursts in rapid succession seems to be at fault here; rather than attempt such an articulatory feat, speakers apparently favor the option of producing the first stop in the sequence with a different place of articulation.

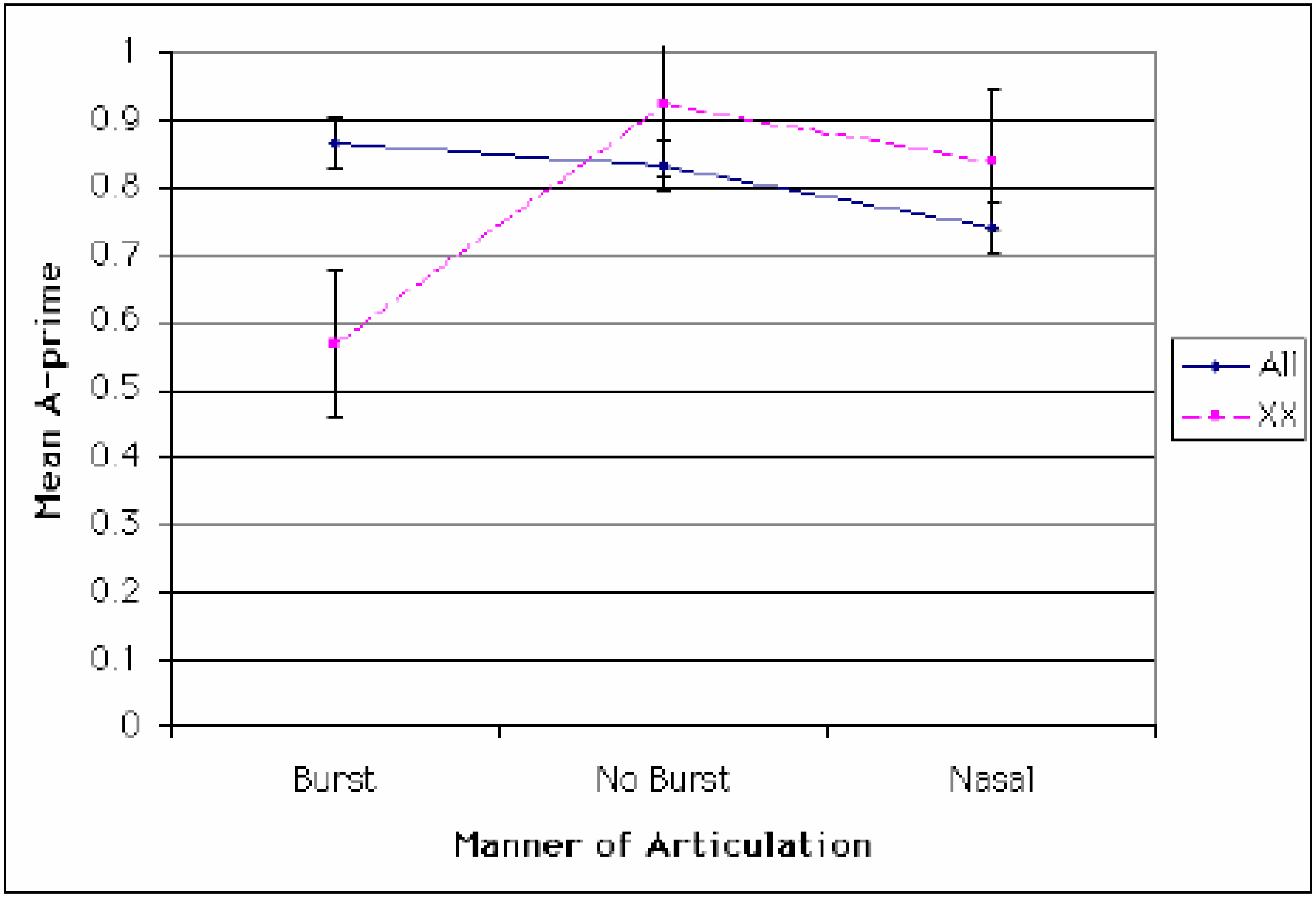


Figure 6.5: A-prime values, by manner and stimulus type

Analyzing the types of mistakes participants made in attempting to reproduce the original, VCCV stimuli helps make such articulatory tendencies clear. Table 6.19 presents the kinds and number of responses listeners gave to all homorganic stimuli. These tables categorize each response type by the relationship its place(s) of articulation has with respect to the places of articulation in the consonant cluster in the original stimulus. All such stimuli contained two consonants with identical places of articulation (X). Tables 6.19 therefore represents the correct responses to these stimuli as XX, for the two consecutive consonants which share the same place of articulation (X) as was in both consonants in the original stimulus cluster. Other possible response types include YX, where the place of articulation of the first consonant in the response differs from that in the original cluster. XY responses are also possible--though considerably rarer. Other response types in Table 6.19 include Y and YY responses--in which there is only one place of articulation in either a singleton, intervocalic consonant, or a homorganic consonant cluster--but this place of articulation differs from the place of articulation in the stimulus cluster. The last possible response type is YZ, but these responses were also quite rare, as is shown in Table 6.19 below.

Bursts	XX	XY	YX	YY	YZ	X	Y	Total
DD	79	21	382	0	0	213	0	695
CC	113	11	407	0	0	172	0	703
LL	46	9	333	2	0	303	0	693
Total	238	41	1122	2	0	688	0	2091

No Burst	XX	XY	YX	YY	YZ	X	Y	Total
DD	24	12	102	2	30	521	22	713
CC	11	0	85	0	0	624	0	720
LL	10	1	91	0	0	611	5	718
Total	45	13	278	2	30	1756	27	2151

Nasal	XX	XY	YX	YY	YZ	X	Y	Total
DD	428	0	220	0	0	59	0	707
CC	522	3	169	4	1	15	0	714
LL	544	1	160	0	0	10	0	715
Total	1494	4	549	4	1	84	0	2136

Table 6.19: Raw tallies of listener response types to stimuli with homorganic (XX) clusters

Bursts	XX	XY	YX	YY	YZ	X	Y	Total
DD	11.0%	2.9%	53.1%	0.0%	0.0%	29.6%	0.0%	96.5%
CC	15.7%	1.5%	56.5%	0.0%	0.0%	23.9%	0.0%	97.6%
LL	6.4%	1.3%	46.3%	0.3%	0.0%	42.1%	0.0%	96.3%
Total	11.0%	1.9%	51.9%	0.1%	0.0%	31.9%	0.0%	96.8%

No Burst	XX	XY	YX	YY	YZ	X	Y	Total
DD	3.3%	1.7%	14.2%	0.3%	4.2%	72.4%	3.1%	99.0%
CC	1.5%	0.0%	11.8%	0.0%	0.0%	86.7%	0.0%	100.0%
LL	1.4%	0.1%	12.6%	0.0%	0.0%	84.9%	0.7%	99.7%
Total	2.1%	0.6%	12.9%	0.1%	1.4%	81.3%	1.3%	99.6%

Nasal	XX	XY	YX	YY	YZ	X	Y	Total
DD	59.4%	0.0%	30.6%	0.0%	0.0%	8.2%	0.0%	98.2%
CC	72.5%	0.4%	23.5%	0.6%	0.1%	2.1%	0.0%	99.2%
LL	75.6%	0.1%	22.2%	0.0%	0.0%	1.4%	0.0%	99.3%
Total	69.2%	0.2%	25.4%	0.2%	0.0%	3.9%	0.0%	98.9%

Table 6.20: Percentages of listener response types to stimuli with homorganic (XX) clusters

Figure 6.6 represents the percentages of the three most common response types to homorganic stimuli—XX, YX and X—in graphical form. This figure shows that YX responses accounted for just over 50% of all responses to homorganic stimuli which had stops with bursts in the first consonant of the intervocalic cluster. The two groups of speakers, in other words, failed to correctly reproduce the place of articulation of these consonants more than half the time. The corresponding rates of YX responses to nasal and no-burst stimuli was 25% and 13%, respectively. Much more common responses to these groups of stimuli were X responses (81.3% for stops without bursts) and XX responses (69.2% for nasals), both of which were considered correct responses to homorganic stimuli. The participants in this study therefore had more success producing the homorganic stop-stop stimuli which lacked a release burst after the first stop, because they (the listeners) seemed to be biased towards producing the stimuli which included these release bursts as heterorganic sequences. The results of the earlier, consonant cluster production study provide an interesting, independent confirmation of this finding. Table 6.21 below shows that release bursts were far more prevalent after the first stop in heterorganic clusters than they were in homorganic clusters.

Plosives	LC	LD	CL	CD	DL	DC	Total	Pct.	Diff.		95%
Heterorganic	30	25	11	24	16	25	131	0.68	0.44	+/-	0.11
	LL		CC		DD						
Homorganic	5		11		7		23	0.24			

Nasals	LC	LD	CL	CD	DL	DC	Total	Pct.	Diff.		95%
Heterorganic	15	17	0	13	4	7	56	0.24	0.17	+/-	0.08
	LL		CC		DD						
Homorganic	0		5		3		8	0.07			

Table 6.21: Tallies and difference in proportions of heterorganic and homorganic clusters with stop bursts, from consonant cluster production study (chapter 5).

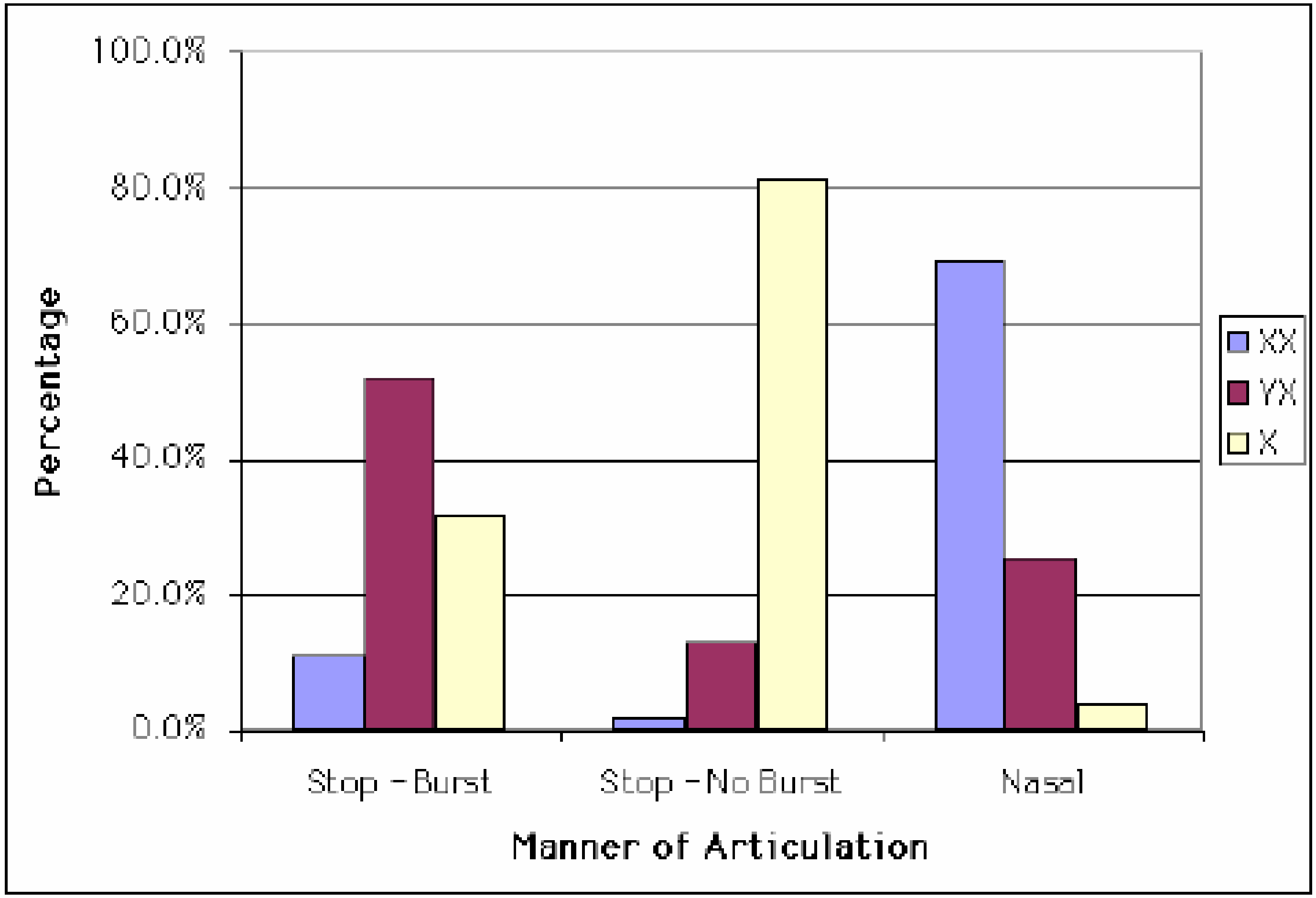


Figure 6.6: Percentages of response types to homorganic (XX) stimuli

Analyzing the response types to heterorganic stimuli in a similar way—as in Tables 6.22-6.23—helps support the argument that articulatory difficulty motivates the process of place assimilation. The earlier, A-prime analysis established that repetition accuracy for nasals in heterorganic sequences was particularly low—perhaps because of the articulatory difficulties listeners had in coordinating the gestures required to produce such sequences. The A-prime analysis also confirmed that repetition accuracy increased for both nasals and stops without bursts in homorganic sequences. A phonological process which changes heterorganic sequences into homorganic sequences--as in, for example, place assimilation--appears, therefore, to simplify a speaker's articulatory task by allowing them to say a consonant sequence which is easier for them to reproduce accurately.

This A-prime analysis cannot, however, establish or confirm that listeners are actually biased towards simplifying their articulatory tasks in this way whenever they fail to reproduce a certain place of articulation in a heterorganic consonant cluster correctly. “Missed" responses to such stimuli could, logically, involve heterorganic clusters just as often as they do homorganic clusters; a mistaken response to an /atpa/ stimulus could, for instance, be either /akpa/ or /appa/. Both of these responses would decrease the A-prime value for /t/ in such heterorganic sequences, but only an assimilated, /appa/ response would also decrease the difficulty of the speaker's articulatory task.

Tables 6.22-6.23 therefore presents the tallies and percentages of particular response types to heterorganic stimuli in order to determine whether or not speakers were biased towards producing one of these mistake types or the other. The categorization of these mistakes follows the same pattern as in Tables 6.19-6.20 above. This means that

YY or Y responses, for instance, represent responses where the first consonant has undergone place assimilation to the second consonant of the cluster, while ZY responses are for mistaken repetitions in which the first consonant shares the place of articulation of neither the first nor the second consonant in the original stimulus.

Bursts	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	X	Y	Z	Total
LC	594	0	11	8	0	0	3	41	0	0	55	0	712
LD	557	0	3	3	5	1	0	73	0	3	58	1	704
CL	584	1	21	0	3	0	1	28	0	6	47	0	691
CD	579	0	6	1	0	2	1	67	0	0	20	0	676
DL	577	0	5	2	6	0	0	22	0	3	79	0	694
DC	565	0	0	5	1	1	0	100	0	0	40	0	712
Total	3456	1	46	19	15	4	5	331	0	12	299	1	4189

No Burst	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	X	Y	Z	Total
LC	158	1	12	2	0	0	0	44	0	7	494	0	718
LD	179	1	18	2	0	0	1	63	0	1	440	0	705
CL	216	0	19	3	0	0	0	48	0	2	428	0	716
CD	206	0	24	2	0	0	0	36	0	0	437	0	705
DL	267	3	10	0	0	0	0	57	0	1	370	3	711
DC	225	0	9	3	1	1	1	66	0	3	406	0	715
Total	1251	5	92	12	1	1	2	314	0	14	2575	3	4270

Nasals	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	X	Y	Z	Total
LC	391	7	239	0	0	0	0	59	0	0	17	0	713
LD	400	0	186	0	0	0	0	93	0	0	32	0	711
CL	227	6	343	0	0	0	0	111	0	0	27	0	714
CD	257	1	324	1	0	0	0	89	0	0	36	0	708
DL	361	0	197	2	0	1	0	111	3	0	29	3	707
DC	320	2	294	1	0	0	0	61	0	0	30	0	708
Total	1956	16	1583	4	0	1	0	524	3	0	171	3	4261

Table 6.22: Raw tallies of listener response types to stimuli with heterorganic (XY) clusters

Bursts	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	X	Y	Z	Total
LC	82.5	0.0	1.5	1.1	0.0	0.0	0.4	5.7	0.0	0.0	7.6	0.0	98.9
LD	77.4	0.0	0.4	0.4	0.7	0.1	0.0	10.1	0.0	0.4	8.1	0.1	97.8
CL	81.1	0.1	2.9	0.0	0.4	0.0	0.1	3.9	0.0	0.8	6.5	0.0	96.0
CD	80.4	0.0	0.8	0.1	0.0	0.3	0.1	9.3	0.0	0.0	2.8	0.0	93.9
DL	80.1	0.0	0.7	0.3	0.8	0.0	0.0	3.1	0.0	0.4	11.0	0.0	96.4
DC	78.5	0.0	0.0	0.7	0.1	0.1	0.0	13.9	0.0	0.0	5.6	0.0	98.9
Total	80.0	0.0	1.1	0.4	0.3	0.1	0.1	7.7	0.0	0.3	6.9	0.0	97.0

No Burst	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	X	Y	Z	Total
LC	21.9	0.1	1.7	0.3	0.0	0.0	0.0	6.1	0.0	1.0	68.6	0.0	99.7
LD	24.9	0.1	2.5	0.3	0.0	0.0	0.1	8.8	0.0	0.1	61.1	0.0	97.9
CL	30.0	0.0	2.6	0.4	0.0	0.0	0.0	6.7	0.0	0.3	59.4	0.0	99.4
CD	28.6	0.0	3.3	0.3	0.0	0.0	0.0	5.0	0.0	0.0	60.7	0.0	97.9
DL	37.1	0.4	1.4	0.0	0.0	0.0	0.0	7.9	0.0	0.1	51.4	0.4	98.8
DC	31.3	0.0	1.3	0.4	0.1	0.1	0.1	9.2	0.0	0.4	56.4	0.0	99.3
Total	29.0	0.1	2.1	0.3	0.0	0.0	0.0	7.3	0.0	0.3	59.6	0.1	98.8

Nasals	XY	XX	YY	YX	XZ	YZ	ZX	ZY	ZZ	X	Y	Z	Total
LC	54.3	1.0	33.2	0.0	0.0	0.0	0.0	8.2	0.0	0.0	2.4	0.0	99.0
LD	55.6	0.0	25.8	0.0	0.0	0.0	0.0	12.9	0.0	0.0	4.4	0.0	98.8
CL	31.5	0.8	47.6	0.0	0.0	0.0	0.0	15.4	0.0	0.0	3.8	0.0	99.2
CD	35.7	0.1	45.0	0.1	0.0	0.0	0.0	12.4	0.0	0.0	5.0	0.0	98.3
DL	50.1	0.0	27.4	0.3	0.0	0.1	0.0	15.4	0.4	0.0	4.0	0.4	98.2
DC	44.4	0.3	40.8	0.1	0.0	0.0	0.0	8.5	0.0	0.0	4.2	0.0	98.3
Total	45.3	0.4	36.6	0.1	0.0	0.0	0.0	12.1	0.1	0.0	4.0	0.1	98.6

Table 6.23: Percentages of listener response types to stimuli with heterorganic (XY) clusters

Figure 6.7 recapitulates the data in Table 6.23 in graphical form. This figure clearly shows that there are considerably more YY responses than ZY responses for nasal-stop heterorganic clusters. YY responses accounted for 31.6% of all responses to these stimuli, while speakers produced ZY responses only 12.1% of the time. The 99% confidence interval for the difference between these proportions is 24.5 $\pm$ 2.3%; the proportion of YY responses was thus significantly greater than the proportion of ZY responses. This bias towards mistakenly producing homorganic clusters indicates that the articulatory difficulty inherent in producing heterorganic clusters motivates speakers to simplify their articulatory task. The preferred repair mechanism, in other words, is not to simply replace the place of articulation of the first consonant with any other place of articulation, but preferably with a place of articulation that matches that of the second consonant in the cluster. This phonological repair process thereby decreases the difficulty of articulating the sequence accurately--at least in the case of nasals and stops without bursts.

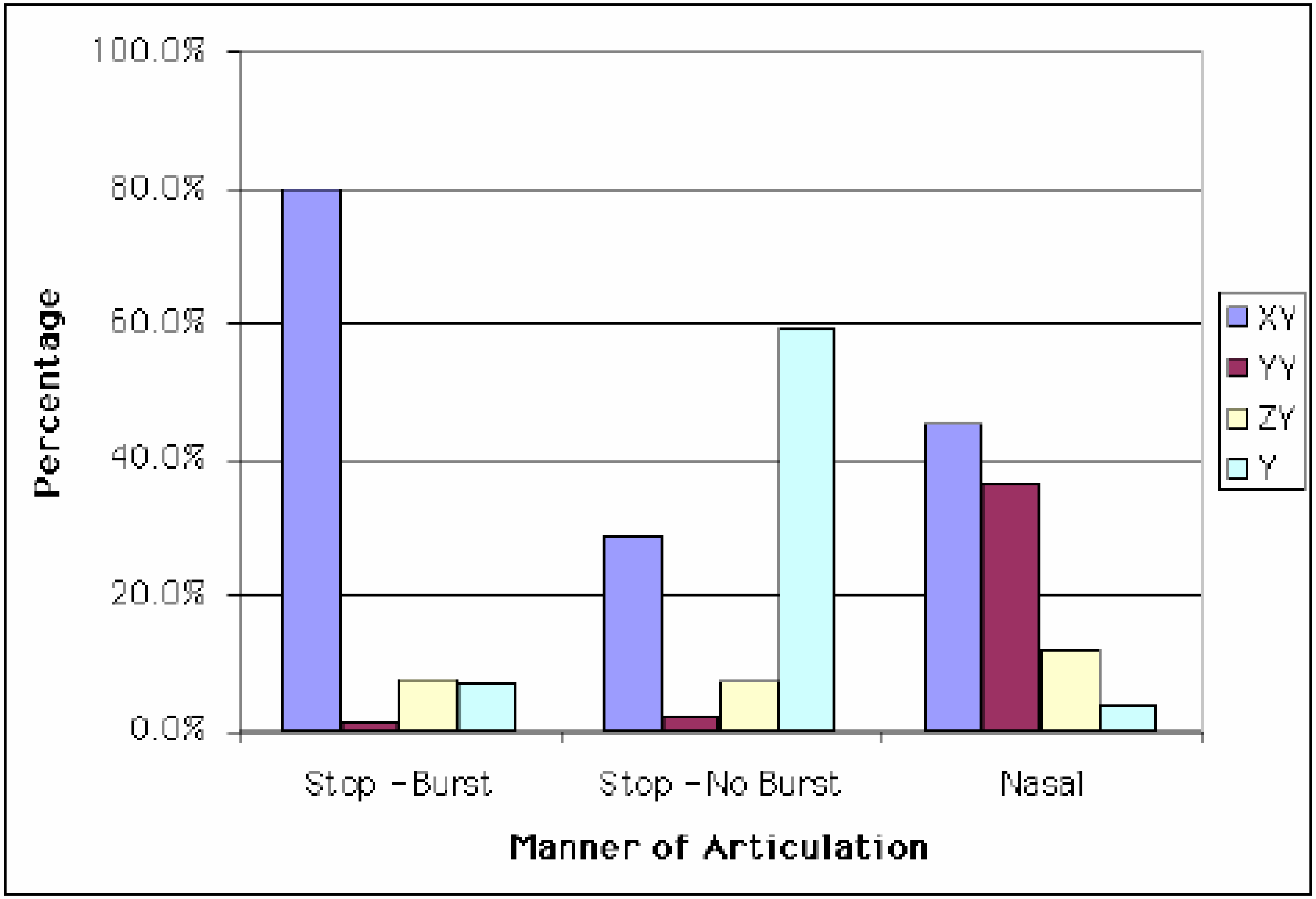


Figure 6.7: Percentages of response types to heterorganic (XY) stimuli

Analyzing the accuracy of spoken repetitions in the repetition/imitation task therefore provides a strong argument that articulatory constraints motivate the cross-linguistic susceptibility of nasals to place assimilation. The A-prime data shows that speakers can reproduce the place of nasals and—to a lesser extent—stops without bursts more accurately in homorganic clusters than in heterorganic clusters. Analyzing the patterns of mistaken responses in the repetition/imitation task also shows that speakers prefer to reduce the articulatory difficulty of producing heterorganic clusters by producing homorganic clusters instead (rather than alternative heterorganic clusters). Place assimilation therefore seems to occur in order to increase the ability of speakers to reproduce consonant clusters accurately—or, as the case may be, spare them the effort of having to produce consonant clusters which are difficult to say.

Analyzing the repetition/imitation response data in terms of articulatory variability provides further insight into the inherent difficulty of producing certain consonant clusters. Following Munson (2000), articulatory variability was analyzed by calculating the token-to-token variability in the duration of consonant clusters in all recorded repetitions from the repetition/imitation task. The durations of these consonant clusters was measured from the offset of the initial vowel (acoustic marker #2) to the onset of the burst for the second consonant (acoustic marker #5) in the intervocalic consonant cluster. The timing information for these acoustic landmarks was calculated and recorded using a customized Praat script, as described in the methods section above. Measuring the duration of the entire consonant cluster, therefore, simply involved subtracting the time of the onset of the first consonant from the time at which the second consonant in the cluster was released.

The variability in the duration of these consonant clusters was determined by measuring differences in duration across tokens of the same consonant cluster--as determined by the transcriptions of the recorded responses. The variability in the duration of /tk/ sequences, for instance, was not measured across all responses to /tk/ stimuli--nor was it measured across only /tk/ responses to /tk/ stimuli. The variability in the duration of /tk/ sequences was simply measured across all responses which were transcribed with a /tk/ cluster in them. Following Munson (2000), token-to-token variability in such consonant cluster durations was calculated by dividing the standard deviations of these durations over the mean of those durations:

(6.4) $Var_i = \sigma_i / \mu_i$

The durations of consonant clusters in some responses were unusually long. In general, these long durations seemed to result from lapses in concentration on the part of the speakers, rather than from genuine difficulties with the production of the cluster. For this reason, these durations were systematically excluded from consideration in the calculation of token-to-token variability for each cluster type. This was done by first calculating the standard deviation and mean of cluster durations across all responses for a particular cluster type. Outlying durations were then excluded from the calculation of variability for that particular cluster if they were greater than two standard deviations (in either direction) from the mean of durations for that cluster type. After excluding these responses, the standard deviation and mean for the durations of the remaining productions of the cluster were calculated once again for that cluster type. The variability of token-to-token duration for that particular cluster type was then determined by taking the ratio of these secondary calculations of the standard deviation and mean for that cluster.

Tables 6.24-6.26 list the resultant variability ratios for each cluster type, as broken down by place sequence, manner of articulation of the first consonant, and language of the listener/repeater. Importantly, this table only includes variability values for heterorganic clusters that were produced by the participants in the experiment. It excludes homorganic clusters, since it was impossible to make equitable comparisons between both cluster types for all manners of articulation (i.e., stop-stop and nasal-stop clusters). There were very few homorganic clusters in which the first consonant was a stop without a burst, for instance; in any event, most of these would have been indistinguishable from a singleton consonant with the same place of articulation. Where homorganic stop-stop clusters and singleton stops may differ, though, is in their duration, and the duration variability of both of these sound sequences taken together would, therefore, be unrealistically high in comparison to other stop-stop and nasal-stop sequences.

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
Dutch	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	0.128	35	0.090	73	0.081	55	0.098	63	0.095	58	0.095	83
2	0.085	29	0.093	37	0.081	30	0.116	45	0.096	49	0.138	49
3	0.187	50	0.135	41	0.144	71	0.157	41	0.159	18	0.164	10
4	0.098	68	0.079	43	0.104	48	0.076	23	0.095	38	0.109	32
5	0.133	40	0.157	47	0.137	99	0.093	44	0.110	61	0.205	56
6	0.171	41	0.137	41	0.146	73	0.160	70	0.173	15	0.141	66
7	0.148	77	0.153	56	0.112	77	0.121	43	0.164	37	0.168	74
8	0.188	42	0.140	42	0.148	80	0.122	47	0.152	33	0.135	45
9	0.151	31	0.088	28	0.107	33	0.119	12	0.143	19	0.118	26
10	0.122	35	0.151	16	0.095	62	0.124	25	0.075	25	0.110	32

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
English	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	0.101	12	0.141	19	0.144	44	0.161	13	0.189	18	0.168	39
2	0.163	14	0.091	24	0.139	80	0.120	34	0.110	28	0.086	27
3	0.186	52	0.157	83	0.149	33	0.143	38	0.147	80	0.147	87
4	0.251	13	0.192	18	0.160	49	0.142	8	0.170	27	0.136	40
5	0.086	21	0.176	13	0.094	88	0.186	47	0.096	22	0.201	15
6	0.139	38	0.111	34	0.115	34	0.083	31	0.078	25	0.101	26
7	0.130	40	0.131	23	0.121	54	0.070	38	0.117	59	0.134	63
8	0.206	26	0.141	52	0.149	35	0.148	26	0.148	80	0.144	55
9	0.121	16	0.182	31	0.147	20	0.137	16	0.136	12	0.124	24
10	0.162	22	0.166	5	0.123	76	0.117	37	0.148	45	0.127	54

Table 6.24: Consonant cluster duration variability, for productions with stops with bursts, by cluster type and repeater (N = number of tokens for cluster and repeater)

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
Dutch	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	0.195	2	0.156	3	---	0	0.120	8	0.088	3	0.134	4
2	0.156	5	0.130	27	0.156	7	0.090	14	0.160	11	0.234	13
3	0.146	22	0.202	31	0.096	8	0.174	21	---	0	0.213	7
4	---	0	---	0	---	0	0.080	26	---	0	---	0
5	0.158	4	0.129	29	0.297	4	0.120	43	0.197	9	0.120	21
6	0.183	3	0.148	4	0.124	8	0.128	23	0.031	2	0.170	16
7	0.191	5	0.147	10	---	0	0.101	9	0.082	8	0.190	18
8	0.087	10	0.180	23	0.202	5	0.127	22	0.118	4	0.194	17
9	0.147	13	0.123	18	---	0	0.103	9	0.137	9	0.162	15
10	0.146	15	0.111	50	0.139	4	0.055	21	0.133	18	0.128	20

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
English	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	0.147	43	0.149	46	0.147	14	0.110	36	0.083	43	0.138	56
2	0.143	21	0.112	27	0.242	17	0.202	21	0.182	9	0.175	17
3	0.187	10	0.135	16	---	0	0.215	6	0.092	10	0.194	27
4	0.176	28	0.144	31	0.218	6	0.165	22	0.152	10	0.143	15
5	0.092	20	0.104	13	0.055	3	0.115	7	0.165	11	0.139	26
6	0.160	22	0.117	17	---	0	0.104	12	0.089	13	---	0
7	0.112	8	0.093	20	0.142	2	0.240	14	0.147	6	0.185	3
8	0.157	37	0.151	26	0.271	3	0.172	10	0.215	10	0.135	18
9	0.184	3	0.146	6	---	0	0.145	5	0.115	18	0.108	9
10	0.235	27	0.035	6	0.266	4	0.169	62	0.201	8	0.100	28

Table 6.25: Consonant cluster duration variability, for productions with stops without bursts, by cluster type and repeater (N = number of tokens for cluster and repeater)

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
Dutch	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	0.148	23	0.132	32	0.155	2	0.228	4	0.084	44	0.108	39
2	0.207	33	0.166	59	0.236	11	0.143	5	0.135	22	0.097	23
3	0.263	14	0.233	14	0.101	18	0.162	47	0.106	22	0.153	22
4	0.103	19	0.077	9	0.087	16	0.061	15	0.092	32	0.128	43
5	0.109	11	0.179	12	0.104	17	0.188	29	0.102	51	0.223	54
6	0.176	16	0.129	7	0.212	20	0.137	18	0.098	9	0.251	4
7	0.140	43	0.166	62	0.130	41	0.164	15	0.070	8	0.161	20
8	0.227	13	0.168	31	0.157	25	0.175	28	0.223	13	0.197	20
9	---	0	0.157	9	0.198	18	0.202	12	0.136	12	0.226	4
10	0.135	42	0.133	33	0.210	11	0.059	2	0.156	13	0.148	18

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
English	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	0.140	23	0.130	27	0.152	76	0.118	37	0.115	17	0.134	29
2	0.165	25	0.136	40	0.198	24	0.133	25	0.143	17	0.063	12
3	0.159	44	0.159	62	0.184	21	0.185	13	0.134	39	0.156	32
4	0.078	20	0.168	18	0.146	14	0.098	2	0.125	43	0.162	42
5	0.128	15	0.130	18	0.105	39	0.151	28	0.084	32	0.099	20
6	0.219	21	0.149	25	0.198	18	0.125	21	0.129	44	0.109	33
7	0.117	8	0.130	15	0.134	40	0.109	33	0.140	45	0.153	42
8	0.146	43	0.198	48	0.163	35	0.120	6	0.125	24	0.255	19
9	0.159	4	0.195	2	0.182	25	0.190	25	0.146	14	0.182	18
10	0.149	32	---	0	0.111	37	0.211	45	0.080	19	0.168	20

Table 6.26: Consonant cluster duration variability, for productions with nasals, by cluster type and repeater (N = number of tokens for cluster and repeater)

A repeated measures analysis of variance was therefore run on the variability data for the heterorganic consonant clusters only. This ANOVA considered only language of speaker (Dutch, English), manner of articulation of the first consonant (nasal, stop with burst, stop without burst) and cluster type (labial-coronal, labial-dorsal, coronal-labial, coronal-dorsal, dorsal-labial, dorsal-coronal) as possible influences on the variability of consonant cluster durations; the results of this ANOVA are given in Table 6.27. This analysis revealed only one significant effect on duration variability: an interaction between manner of articulation and the language of the speaker ($F = 6.483$, $df = 2,8$, $p = .021$). The post-hoc t-tests on this interaction, given in Table 6.28, reveals how particularly interesting the direction of this effect is. For the English speakers, there are no significant differences in duration variability between any of the three manners of articulation. Duration variability is slightly less for stops without bursts than it is for the other two manners of articulation, but this difference is not statistically reliable. For the Dutch speakers, on the other hand, the duration variability for stops without bursts <u>is</u> significantly lower than the variability in both nasal-stop clusters and no-burst stop clusters. Figure 6.8 presents graphically how variability interacts with the language of the speaker and manner of articulation.

	F	df	Sig.
Manner	1.234	2,8	0.341
Manner * Listener	6.483	2,8	**0.020**
Cluster	3.555	5,5	0.095
Cluster * Listener	1.62	5,5	0.305
Manner * Cluster			
Manner * Cluster * Listener			
Listener	0.054	1,9	0.821

Table 6.27: Results of ANOVA on output XY duration variability

Manner * Listener

Manner	Listener	Mean Var.	T-Tests
Burst	Dutch	0.127	0.307
	English	0.140	---
No Burst	Dutch	0.139	0.346
	English	0.151	---
Nasal	Dutch	0.154	0.198
	English	0.145	---

Listener * Manner			T-Tests	
Listener	Manner	Mean Var.	vs. No Burst	Vs. Nasal
Dutch	Burst	0.127	0.154	**0.008**
	No Burst	0.139	---	0.051
	Nasal	0.154	---	---
English	Burst	0.140	0.222	0.559
	No Burst	0.151	---	0.516
	Nasal	0.145	---	---

Table 6.28: T-tests on significant effects in output duration variability ANOVA

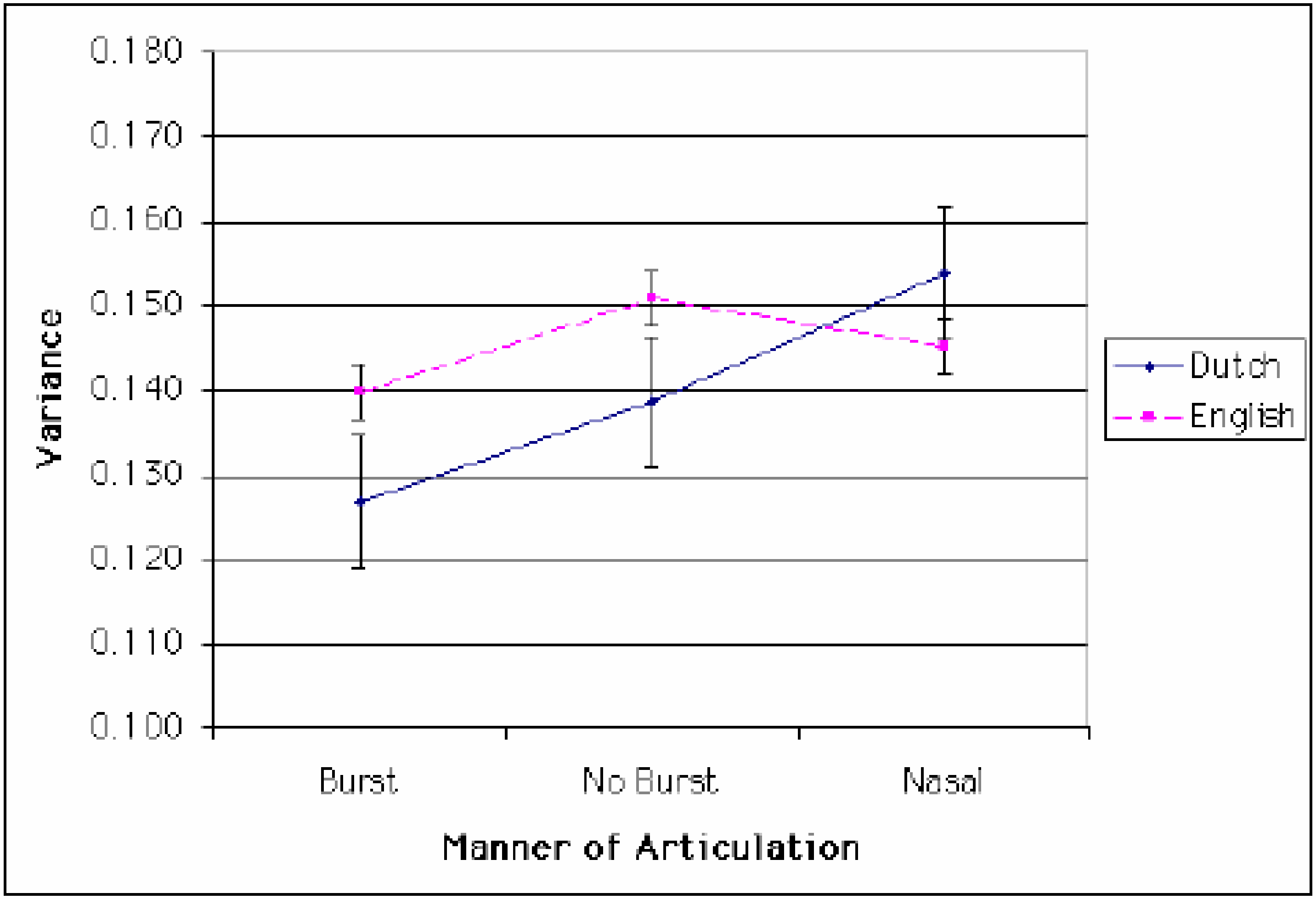


Figure 6.8: Variation in consonant cluster duration

Post-hoc t-tests also reveal that, for Dutch speakers, the difference in duration variability between nasals and stops without bursts is also on the cusp of significance ($p = .051$). The overall pattern of interactions therefore seems to indicate that Dutch speakers show less variability in stop-stop productions than in nasal-stop productions, while English speakers are equally variable at producing both manners of articulation in heterorganic clusters.

This interaction between language and manner of articulation corresponds to some extent with the phonological differences between Dutch and English. In English, both nasals and stops undergo place assimilation in casual speech, while, in Dutch, only nasals may undergo this process. The manner of articulation of the targets of place assimilation in both languages corresponds to the manners of articulation which speakers produce most variably in heterorganic clusters in each language. This correspondence between phonology and articulation thus provides more evidence that articulation might motivate certain segments to undergo place assimilation; however, in this case, the motivation is language-specific, rather than universal. The fact that analyzing articulatory difficulty in these two different ways has led to conflicting conclusions indicates that these two measures—repetition accuracy and repetition variability—may capture completely independent aspects of "articulatory difficulty."

Analyzing the variability of formant transitions also reveals another interesting disjunction between independent measures of articulatory difficulty. Formant transition variability was measured in much the same way as duration variability. The transitions themselves were calculated by measuring the difference in formant values for all of the first four formants between the midpoint of the initial vowel and the offset of that same vowel. The timing of the offset of the vowel had been marked (acoustic marker #2) using a customized Praat script, as described in the methods section above. The timing of the onset of this vowel was also recorded during the same segmentation process (acoustic marker #1), so the midpoint of the first vowel was simply calculated by adding half of the difference between first vowel offset and onset to the time of the onset of that vowel. Following Zsiga (1994), the formant transitions were then calculated by subtracting formant values at vowel midpoint from the corresponding values at vowel offset; these values would be negative in the case of falling formant transitions and positive in the case of rising formant transitions.

The primary transition cues to place of articulation are in the second and third formants, so this analysis only considered token-to-token variability in the transitions for these two formants. The token-to-token variability for these two formant transitions was calculated in the same way as duration variability was calculated for each particular consonant cluster. Only formant transitions into heterorganic clusters were considered in this analysis. Formant transition variability was also calculated for particular consonant clusters, as they were transcribed for each spoken response. Outlying transition values were also excluded from the token-to-token variability calculations. In order to do this, standard deviations and means were calculated for each group of responses for a particular cluster type, and transitions which were more than two standard deviations distant—in either direction--from these preliminary means were then excluded from the analysis. Standard deviations were then calculated for the second and third formant transitions in all remaining responses, for each cluster type. These secondary calculations of the standard deviation were thereafter taken to be the token-to-token variability for the two formant transitions for each consonant cluster type. Tables 6.29-6.34 list all the variability values that were calculated in this way, for second and third formant transitions, respectively, broken down by cluster type, manner of articulation of the first consonant, and language of the repeater.

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
Dutch	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	455.6	34	417.1	71	570.5	55	408.9	62	366.8	58	478.3	81
2	604.6	30	381.6	41	447.9	35	630.9	44	535.6	45	651.2	51
3	684.3	51	704.3	42	315.6	78	479.1	39	776.1	26	397.8	16
4	609.8	66	618.5	41	308.0	59	407.6	21	162.8	54	517.8	50
5	926.4	41	846.8	50	575.3	107	778.1	43	905.7	106	833.8	96
6	426.6	40	684.1	41	488.2	76	761.6	65	581.0	14	759.7	67
7	954.3	102	856.9	71	608.9	108	788.1	43	728.4	39	871.6	90
8	515.0	41	843.7	42	413.0	78	583.8	46	666.6	38	674.3	56
9	383.1	30	226.5	30	508.6	32	529.1	11	564.5	28	802.2	28
10	498.9	39	369.6	16	529.5	67	525.5	24	1059.1	31	732.8	46

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
English	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	143.6	11	486.6	22	283.4	94	597.3	14	555.4	26	516.9	52
2	358.8	19	506.8	23	568.7	91	584.9	36	516.5	35	510.8	33
3	435.0	61	584.9	89	387.3	43	732.6	40	501.5	95	567.2	103
4	792.3	14	589.4	17	637.3	59	313.5	8	641.5	63	688.6	73
5	348.7	23	498.4	19	442.9	109	400.8	46	792.2	37	1048.7	25
6	207.4	41	96.0	33	343.4	43	283.8	33	371.9	60	615.9	50
7	520.5	40	485.7	27	803.3	87	639.3	48	636.2	86	683.1	86
8	698.6	27	296.3	55	447.5	50	429.6	24	836.1	94	568.9	68
9	900.7	18	783.5	31	486.1	42	1084.0	23	330.6	15	511.1	30
10	572.4	23	589.7	5	395.9	99	414.2	39	432.8	53	314.4	59

Table 6.29: F2 transition standard deviations, by cluster type and repeater, for stops with burst productions (N = number of tokens for cluster and repeater)

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
Dutch	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	500.7	37	471.1	71	602.9	56	361.0	62	428.8	63	427.1	84
2	582.5	31	518.8	44	708.4	34	464.2	42	485.5	49	591.0	52
3	508.7	52	436.4	41	533.4	85	549.0	41	763.9	25	506.2	16
4	459.5	69	445.9	41	403.3	59	475.9	23	283.6	55	395.6	54
5	510.9	41	537.1	48	531.6	110	686.4	43	571.8	103	543.2	93
6	357.0	43	461.2	42	574.0	77	534.9	69	541.8	14	476.2	69
7	654.3	103	547.1	74	553.3	115	604.3	45	353.2	38	489.1	86
8	404.4	42	737.3	43	372.4	78	313.0	44	399.2	37	487.4	56
9	608.6	31	375.4	30	556.2	33	484.6	11	501.0	28	736.4	27
10	641.4	41	477.5	16	557.8	69	702.1	24	688.1	29	600.6	46

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
English	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	343.8	13	344.6	23	279.1	95	407.0	13	401.1	26	589.5	53
2	437.3	19	631.9	24	865.9	93	878.1	36	723.0	36	583.7	32
3	442.3	60	536.5	94	474.8	43	501.5	39	440.9	95	317.1	102
4	434.8	13	417.9	17	463.1	58	160.5	7	463.6	64	479.4	75
5	790.7	23	460.3	18	458.8	112	472.7	50	493.9	36	783.4	24
6	363.1	41	257.4	34	258.2	45	136.9	32	331.8	63	418.9	56
7	758.9	43	647.3	29	624.3	86	589.1	49	496.3	91	527.5	93
8	838.4	27	713.1	59	644.8	51	786.4	26	924.7	94	524.1	70
9	761.8	18	597.4	31	662.5	45	511.3	23	522.7	16	343.9	29
10	466.1	24	138.2	5	462.9	101	332.8	38	346.1	52	299.2	58

Table 6.30: F3 transition standard deviations, by cluster type and repeater, for stops with burst productions (N = number of tokens for cluster and repeater)

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
Dutch	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	211.4	2	290.1	3	---	0	767.1	8	986.2	3	211.9	4
2	1162.4	5	738.5	27	454.3	7	340.6	13	775.9	11	470.7	12
3	582.5	23	911.1	31	914.0	8	586.8	19	---	0	250.1	6
4	---	0	---	0	---	0	310.4	25	---	0	---	0
5	91.7	4	732.1	30	868.9	4	414.4	41	908.4	9	737.3	21
6	988.4	3	532.8	4	139.4	7	691.9	22	51.8	2	1125.4	17
7	832.6	5	691.0	10	---	0	1032.0	9	1107.0	8	815.9	18
8	519.6	9	562.6	22	935.0	5	597.3	22	1029.0	4	345.2	17
9	207.0	12	347.4	18	---	0	145.0	9	569.1	9	687.7	14
10	645.2	16	516.1	48	161.7	4	875.4	23	659.9	18	918.4	19

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
English	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	483.3	40	643.3	44	395.5	14	645.9	36	526.7	45	440.4	53
2	589.1	21	471.2	27	682.2	17	852.2	23	626.2	9	658.2	17
3	306.9	11	426.1	16	---	0	564.1	6	521.7	10	523.1	27
4	724.8	29	585.6	32	230.2	5	454.6	21	534.8	11	768.8	16
5	135.1	19	456.3	12	246.0	3	68.8	6	810.3	10	425.5	24
6	305.6	21	220.9	17	---	0	225.1	11	244.4	12	---	0
7	164.9	7	550.2	20	76.5	2	675.7	14	357.0	6	311.9	3
8	489.1	34	667.7	26	218.6	3	794.4	10	348.8	11	651.9	18
9	466.1	3	389.9	6	---	0	329.4	5	703.6	19	1057.1	9
10	616.5	26	292.5	6	195.0	4	497.0	61	404.3	8	153.1	26

Table 6.31: F2 transition standard deviations, by cluster type and repeater, for stops without burst productions (N = number of tokens for cluster and repeater)

Cluster Dutch	Dor-Cor Var.	N	Dor-Lab Var.	N	Cor-Dor Var.	N	Cor-Lab Var.	N	Lab-Dor Var.	N	Lab-Cor Var.	N
1	171.5	2	207.8	3	---	0	504.2	7	925.3	3	303.5	4
2	637.4	5	569.3	25	403.1	6	492.5	14	472.5	10	557.8	13
3	568.2	22	650.9	31	463.4	8	578.3	20	---	0	402.8	7
4	---	0	---	0	---	0	307.0	26	---	0	---	0
5	196.8	4	629.8	30	638.8	4	507.2	42	668.9	9	667.2	22
6	90.9	3	393.6	4	293.3	7	419.1	23	93.8	2	276.7	15
7	587.5	5	592.6	11	---	0	649.4	9	574.8	8	540.9	18
8	391.2	9	700.7	23	957.7	5	374.2	21	562.8	4	346.5	18
9	504.5	13	589.7	18	---	0	155.4	8	699.0	9	506.7	14
10	429.1	15	531.1	50	681.4	4	537.8	22	753.6	18	540.6	19

Cluster English	Dor-Cor Var.	N	Dor-Lab Var.	N	Cor-Dor Var.	N	Cor-Lab Var.	N	Lab-Dor Var.	N	Lab-Cor Var.	N
1	416.7	44	671.6	45	459.1	14	643.1	38	443.8	45	465.2	53
2	608.5	20	411.8	25	778.9	17	697.7	22	802.9	9	499.5	16
3	270.6	9	441.8	16	---	0	443.3	6	474.9	10	362.1	26
4	461.7	27	367.7	32	567.5	6	469.2	22	397.6	11	519.0	17
5	1198.7	20	919.5	12	183.8	3	195.8	7	657.0	10	551.1	26
6	265.0	21	393.3	16	---	0	120.2	11	342.6	13	---	0
7	368.4	8	884.3	21	226.8	2	560.1	13	396.2	6	212.4	3
8	736.7	36	786.5	28	482.7	3	1017.8	9	618.1	10	765.7	18
9	943.3	3	705.3	5	---	0	470.1	5	705.8	19	645.2	8
10	502.3	26	500.5	7	359.0	4	401.4	62	343.7	7	280.6	27

Table 6.32: F3 transition standard deviations, by cluster type and repeater, for stops without burst productions (N = number of tokens for cluster and repeater)

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
Dutch	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	481.8	20	200.2	28	9.5	2	233.4	4	164.7	38	258.5	36
2	830.4	34	534.2	52	672.4	6	132.9	5	424.0	21	768.6	22
3	100.3	11	215.0	14	190.0	5	518.6	47	293.5	15	387.6	15
4	206.8	18	608.0	9	902.6	4	113.5	15	266.6	12	359.1	18
5	398.7	10	342.7	8	778.7	3	433.0	26	1087.1	13	564.4	13
6	75.5	15	663.1	7	110.7	13	200.7	17	308.1	9	403.6	4
7	950.4	18	946.6	46	622.2	5	463.2	14	460.1	4	848.0	4
8	510.7	12	392.6	30	214.8	23	323.8	28	273.6	5	636.6	8
9	---	0	703.1	6	53.9	16	79.0	12	354.3	2	481.3	4
10	522.6	33	583.4	32	937.3	5	579.1	2	608.0	9	815.8	5

Cluster	Dor-Cor		Dor-Lab		Cor-Dor		Cor-Lab		Lab-Dor		Lab-Cor	
English	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N	Var.	N
1	346.8	21	549.5	23	272.2	20	306.0	36	188.2	7	505.0	16
2	382.9	22	459.0	38	343.3	12	448.6	22	508.2	8	678.6	7
3	438.9	34	286.3	43	950.9	10	893.4	11	555.1	20	597.3	11
4	643.6	19	851.0	18	167.2	3	2.3	2	636.1	5	218.5	8
5	1201.2	13	477.1	10	907.7	15	581.0	21	1089.1	17	932.9	12
6	268.7	15	553.1	25	637.2	5	436.7	19	886.8	5	642.3	3
7	421.9	5	473.2	11	306.7	10	804.4	26	538.3	10	755.0	14
8	729.0	41	969.3	41	669.7	16	226.6	6	788.2	10	603.9	5
9	131.2	2	11.0	2	---	0	398.5	17	489.1	9	667.3	10
10	519.3	30	---	0	505.7	12	462.5	43	377.5	11	153.5	12

Table 6.33: F2 transition standard deviations, by cluster type and repeater, for nasal productions (N = number of tokens for cluster and repeater)

Cluster Dutch	Dor-Cor Var.	N	Dor-Lab Var.	N	Cor-Dor Var.	N	Cor-Lab Var.	N	Lab-Dor Var.	N	Lab-Cor Var.	N
1	324.0	20	286.6	30	191.9	2	112.3	4	220.6	39	237.6	35
2	513.8	33	566.9	55	745.8	6	350.3	5	394.7	23	656.5	22
3	297.0	11	342.1	14	529.9	5	445.1	47	209.5	15	270.9	15
4	699.6	20	413.4	8	813.6	4	416.2	16	290.0	11	469.2	19
5	478.0	10	745.9	9	717.0	3	375.6	27	578.7	13	919.7	14
6	436.3	16	774.2	7	483.4	14	463.7	17	290.4	8	843.2	4
7	361.6	18	629.3	46	693.9	5	385.1	14	204.2	4	255.9	4
8	484.3	12	629.2	32	431.7	23	534.2	29	802.7	5	636.7	9
9	---	0	914.7	6	205.8	17	273.4	11	133.4	2	1996.8	4
10	468.0	36	522.7	32	846.3	5	824.0	2	973.6	9	1326.5	5

Cluster English	Dor-Cor Var.	N	Dor-Lab Var.	N	Cor-Dor Var.	N	Cor-Lab Var.	N	Lab-Dor Var.	N	Lab-Cor Var.	N
1	387.4	22	495.4	22	235.1	20	211.2	36	125.1	7	481.1	16
2	437.8	22	494.2	39	498.6	12	777.8	25	367.7	8	834.0	8
3	539.0	35	638.9	48	864.3	10	511.8	10	475.4	21	424.0	11
4	554.7	20	523.3	18	271.0	3	5.7	2	733.9	5	137.2	8
5	441.0	13	351.0	10	703.2	15	582.8	22	854.7	17	465.9	11
6	588.8	17	346.5	26	462.6	5	308.7	19	404.5	5	209.0	3
7	786.7	5	402.4	10	430.0	10	633.0	25	362.4	11	424.5	14
8	569.4	43	816.5	42	705.5	16	229.3	7	626.0	10	1707.1	5
9	486.1	2	66.5	2	---	0	433.1	17	217.0	9	243.8	10
10	462.6	28	---	0	428.3	12	390.5	42	198.0	10	151.7	12

Table 6.34: F3 transition standard deviations, by cluster type and repeater, for nasal productions (N = number of tokens for cluster and repeater)

Tables 6.35 and 6.36 show the results of two repeated measures ANOVAs, which were run on the formant variability data. These two ANOVAs considered only the effects that language of speaker (English, Dutch), cluster type (labial-coronal, labial-dorsal, coronal-labial, coronal-dorsal, dorsal-labial, dorsal-coronal) and manner of articulation (stop with burst, nasal) might have on the formant transition variability values. These ANOVAs excluded the formant variability data for stops without bursts because of insufficient data for this manner of articulation. Both ANOVAs yielded a significant main effect for manner of articulation ($F = 7.875$, $df = 1,12$, $p = .016$ for F2 variability; $F = 5.367$, $df = 1,12$, $p = .039$ for F3 variability). Somewhat surprisingly, post-hoc t-tests indicate that the direction of this effect--for both formant values--is the opposite of what might be expected: transition variability is higher for nasals than it is for stops with bursts. These results suggest that speakers of both Dutch and English can produce nasal place of articulation in heterorganic clusters more consistently than stop place of articulation. However, a significant interaction ($F = 11.35$, $df = 1,12$, $p = .003$) between manner of articulation and the language of the speaker emerges in the analysis of F2 transition variability. Post-hoc t-tests on this effect (Table 6.37) indicate that only Dutch speakers exhibit greater variability in stop transitions than in nasals--for English speakers, there is no significant difference.

	F	df	Sig.
Manner	7.875	1,12	**0.020**
Manner * Listener	13.35	1,12	**< .001**
Cluster	2.04	5,8	0.177
Cluster * Listener	1.274	5,8	0.362
Manner * Cluster	0.621	5,8	0.689
Manner * Cluster * Listener	0.738	5,8	0.616
Listener	0	1,12	0.985

Table 6.35: Results of ANOVA on output F2 variability

	F	df	Sig.
Manner	5.367	1,12	**0.040**
Manner * Listener	0.038	1,12	0.848
Cluster	1.649	5,8	0.252
Cluster * Listener	0.291	5,8	0.905
Manner * Cluster	3.312	5,8	0.065
Manner * Cluster * Listener	0.357	5,8	0.864
Listener	0.104	1,12	0.753

Table 6.36: Results of ANOVA on output F3 variability

Manner	Mean Var.	T-Test
Stops	323.0	**<u>0.003</u>**
Nasals	254.6	

Manner * Listener

Manner	Listener	Mean Var.	T-Tests
Stops	Dutch	346.7	0.157
	English	299.2	
Nasals	Dutch	230.7	0.061
	English	278.1	

Listener * Manner

Listener	Manner	Mean Var.	T-tests
Dutch	Stops	346.7	**<u>< .001</u>**
	Nasals	230.7	
English	Stops	299.2	0.519
	Nasals	278.1	

Table 6.37: T-tests on significant effects in output F2 variability ANOVA

Manner	Mean Var.	T-Test
Stops	320.5	**<u>0.001</u>**
Nasals	253.1	

Table 6.38: T-test on significant effect in output F3 variability ANOVA

The patterns in formant transition variability thus seem to complement the patterns in consonant cluster duration variability: where duration varies more, transitions vary less, but where transitions vary more, durations vary less. In the case of second formant transitions, this even corresponds to the language-specific differences found in duration variability: there are significant differences in variability between stops and nasals for Dutch speakers, but not for English speakers. This complementarity suggests a trade-off between speakers' ability to produce place of articulation consistently in heterorganic clusters and their ability to produce the entire consonant sequence fluently. Success in producing place of articulation correctly in such clusters, that is, apparently comes at the cost of being able to coordinate the entire gestural sequence for the consonant cluster fluently.

It may, therefore, be concluded that speakers do have the ability to produce the place of articulation of nasals and stops correctly in heterorganic consonant clusters. Speakers cannot, however, coordinate the gestures required to produce these clusters very consistently if and when they achieve articulatory accuracy in producing all of the places of articulation. As an alternative to attempting a sequence of articulatory tasks which are this difficult to coordinate, speakers may allow the first stop or nasal in the cluster to undergo place assimilation, and produce a comparatively simpler homorganic cluster instead. Since Dutch speakers apparently do not have difficulty coordinating the gestures required for heterorganic stop-stop clusters, they are under less pressure to make assimilatory substitutions for these sound sequences. Stops, therefore, do not undergo place assimilation in casual speech in Dutch, even though nasals do.

6.4 Conclusion

The results of the previous four studies therefore seem to show that the cross-linguistic susceptibility of nasals to place assimilation is motivated exclusively by constraints on articulation, regardless of perception. The results of this final, repetition/imitation study showed, however, that these articulatory constraints have both universal and language-specific aspects. Speakers of both Dutch and English had more difficulty reproducing heterorganic nasal-stop clusters accurately than they did heterorganic stop-stop clusters. This articulatory fact corresponds to the broader, cross-linguistic fact that nasals undergo place assimilation more often than stops do. However, measures of articulatory variability across repetitions differed between the two groups of speakers. Heterorganic stops-stop sequences and heterorganic nasal-stop sequences were equally difficult for the English speakers to produce consistently, while Dutch speakers had more success producing the stop-stop sequences consistently. Because of the Dutch speakers' lack of difficulty in producing these clusters--apparently--heterorganic stop-stop sequences do not undergo place assimilation in Dutch. Thus, even though articulatory constraints can, apparently, operate as an "external" influence on phonology, these constraints are not determined solely on a language-independent basis. They may, in fact, be influenced by language-specific knowledge or experience.

The next, and last, chapter in this dissertation examines the influence that the frequency of language-specific experiences in English and Dutch might have on the interactions between phonological processes and external forces such as articulation and perception.

CHAPTER 7

CODA AND CONCLUSION: THE EFFECTS OF FREQUENCY AND LANGUAGE-SPECIFIC EXPERIENCES ON PATTERNS IN PLACE ASSIMILATION

One aspect of the assimilatory processes in both Dutch and English which the articulatory data cannot account for is the fact that only coronal stops and nasals--in both languages--may undergo this process. If anything, this phonological asymmetry seems to be most closely related to the perception of place of articulation in both languages. Both the AX discrimination experiment and the magnitude estimation study yielded results showing the lowest levels of discriminability or perceptual distance for coronal consonants of both manners of articulation. The tendency for stops and nasals of this particular place of articulation to undergo place assimilation may, therefore, result from perceptual influences as Kohler (1990), Hura et al. (1992) and Jun (1995) envisioned them: the perceptually weakest segments are those which are the most susceptible to phonological change.

Theoreticians such as Kohler (1990), Hura et al. (1992), and Jun (1995), however, based their analysis of perceptual influences on phonology on the assumption that perceptual considerations had extra-linguistic, auditory/acoustic origins and were, therefore, the same for all speakers of all languages. Jun (1995), for instance, argued that coronals were more susceptible, cross-linguistically, to processes of place assimilation because the cues to their place of articulation were particularly weak in a post-vocalic context. These cues are weak because coronal gestures themselves are typically rapid. The resultant vowel-to-consonant transitions for coronals are therefore relatively short, and provide listeners with less perceptual information than either labial or dorsal transitions, which are comparatively longer. Evidence has emerged since Jun (1995), however, that coronals do not always have weaker place cues than other places of articulation in experimental settings (cf. Winters 2001). Furthermore, in some languages, stops and nasals of other places of articulation may undergo place assimilation to the exclusion of coronals. Odden (1987), for instance, notes that only nasal dorsals undergo place assimilation in Chukchi. Similarly, Stemberger (1992), citing data from Merlett (1981), describes a process of place assimilation in Seri which targets only labials, to the exclusion of coronals and dorsals. Assimilatory patterns in which three places of articulation--labial, coronal, and dorsal--individually undergo place assimilation in a language to the exclusion of all other places of articulation in that language have, therefore, been attested to in at least one of the world's languages (Hume & Tserdanelis 2003). It is impossible for all three of these places of articulation to individually undergo place assimilation because they all have the universally weakest place cues of any place of articulation. If perception does, in fact, motivate place assimilation in these various cases, the perceptually weakest place of articulation would have to change from language to language. The perceptual influences on phonology, that is, would not be universal but, rather, language-specific.

The phonological evidence thus suggests that perceptual abilities may be modified by linguistic knowledge or experience, just as the production data in chapter 6 showed that articulatory abilities may vary from language to language. (These perceptual and articulatory abilities may, in turn, influence phonological processes in the language from which they emerged.) Such language-specific effects on perception and articulation presumably arise from either language-specific experiences or knowledge of a particular language. A growing body of research (Bybee 2000, Vitevitch & Luce 1999, Frisch 1996, Pierrehumbert 2000) has, in fact, argued that much--if not all--language-specific knowledge is rooted in the frequency of certain language-specific experiences. It is therefore interesting to note the relationship that exists between the frequency of certain places of articulation in English and Dutch and their likelihood to undergo place assimilation in those two languages.

Tables 7.1-7.4 list the type and token frequencies, in both English and Dutch, for the nasal and stop sounds that were tested in the aforementioned perception and production experiments. These frequencies were calculated by counting the relevant number of monomorphemic items in the CELEX database (Baayen et al. 1995). The CELEX database for English consists of 52,588 unique lemmas extracted from dictionaries and literature by the CELEX Center for Lexical Information in Nijmegen, the Netherlands; the corresponding database for Dutch includes 124,136 lemmas. Both databases are essentially text files, listing the orthographic representations of each lemma along with unique identifiers, phonological information, and a raw frequency count for that lemma in the source materials. For example, the Dutch database contains the following line of information for the monomorphemic noun, "oorlog" (English “war”):

(7.1)

71859\oorlog\8285\'or-lOx\[VVC][CVC]\[o:r][lOx]\'or-lOx\[VVC][CVC]\[o:r][lOx]\o:r-lOG\o:r.lOG

The only relevant information in this database entry for the calculation of the frequency totals in Tables 7.1-7.4 is the basic phonetic transcription of the lemma (i.e., [o:r][lOx]), along with its frequency count (8285). From this information, it is possible to count the number of lemmas (or items) in which certain sound sequences occurred--i.e., their "type" frequency--as well as the number of times in the CELEX corpora that each such sound or sound pattern occurred--i.e., their "token" frequency. The tables in 7.1-7.4 list both of these frequency counts for both language databases.

Dutch Types	Total	Word Begin	Word End
p	1603	606	262
t	2809	504	831
k	2392	1037	414
m	1490	437	260
n	1775	142	397
ŋ	382		70

Table 7.l: Type frequencies of relevant stop and nasal sounds, in CELEX Dutch database

Dutch Tokens	Total	Word Begin	Word End
p	836941	306650	190438
t	3977344	988383	1642772
k	2092910	900006	493089
m	1358746	690483	215198
n	5606034	351633	3910268
ŋ	411448		95856

Table 7.2: Token frequencies of relevant stop and nasal sounds, in CELEX Dutch database

English Types	Total	Word Begin	Word End
p	1201	557	256
t	2241	482	895
k	1786	693	406
m	1165	378	268
n	1789	155	625
ŋ	228		59

Table 7.3: Type frequencies of relevant stop and nasal sounds, in CELEX English database

English Totals	Total	Word Begin	Word End
p	481522	246516	125952
t	2080757	463669	1196780
k	711899	267812	315028
m	856631	359795	388417
n	2096671	272474	852609
ŋ	138544		77982

Table 7.4: Token frequencies of relevant stop and nasal sounds, in CELEX English database

In order to calculate these numbers, the phonetic transcription and frequency count information for each lemma was extracted from the CELEX database by a customized Perl script. This Perl script also removed extraneous brackets, dashes, and stress markers from the transcription of each lemma, as well as colons (representing vowel length), thereby leaving only sound segment and count information, as below:

(7.2) orlOx 8285

Before calculating type frequency statistics across the arrays of such information for both languages, the collection of lemmas in each language was reduced to only those which had unique phonetic transcriptions. There were two distinct lemmas in the English database, for instance, which had the phonetic transcription /wO:r/. The CELEX entries for these two lemmas are given below:

(7.3) 50735\war\6484\1\P\'w$R\[CVVC]\[wO:r*]

(7.4) 50736\war\38\1\P\'w$R\[CVVC]\[wO:r*]

The first /wO:r/ lemma had a frequency count of 6,484 while the second had a frequency count of 38. The customized Perl script reduced these two entries to one, and combined the two frequency counts—6,484 and 38--into 6,522 for the sole, remaining entry. This same Perl script then counted how many of these unique phonetic transcriptions contained particular sounds or sound sequences of interest. The basic sounds of interest have already been listed in Tables 7.1-7.4 and include all of the stop and nasal consonants which were tested in the production and perception experiments described in the previous chapters. Counts were made not only of the number of unique transcriptions which contained these consonants in any context, but also of the number of lemmas in which these consonants appeared at the beginning of a word or at the end of a word. The customized Perl script also made corresponding counts for the token frequencies of each of these sounds in each of these contexts. Calculating these frequencies simply involved summing up the CELEX frequency counts for each lemma in which they appeared.

The customized Perl script also made counts of the total number of phonetically unique, monomorphemic lemmas in both languages (8,283 in Dutch, 6,605 in English), as well as the sum of all these lemmas' frequency counts, by language (15,896,626 in Dutch, 11,516,848 in English). Dividing the raw frequency counts in Tables 7.1-7.4 by these type and token totals enables equitable comparisons to be made between the frequency counts for the various sounds and sound patterns in the two different languages. Tables 7.5 and 7.7, for instance, list the percentage of types in the English and Dutch databases, respectively, which contain /p/, /t/, /k/, etc., in the various phonological environments. Likewise, Tables 7.6 and 7.8 list the corresponding percentages of tokens in both databases which contain /m/, /n/ and /ŋ/, etc., in various contexts.

Type Percents	Total	Word Start	Word End
p	19.4%	7.3%	3.2%
t	33.9%	6.1%	10.0%
k	28.9%	12.5%	5.0%
m	18.0%	5.3%	3.1%
n	21.4%	1.7%	4.8%
ŋ	4.6%	0.0%	0.8%

Table 7.5: Percentages of types with relevant stop and nasal sounds, in CELEX Dutch database

Token Percents	Total	Word Start	Word End
p	5.3%	1.9%	1.2%
t	25.0%	6.2%	10.3%
k	13.2%	5.7%	3.1%
m	8.5%	4.3%	1.4%
n	35.3%	2.2%	24.6%
ŋ	2.6%	0.0%	0.6%

Table 7.6: Percentages of tokens with relevant stop and nasal sounds, in CELEX Dutch database

Type Percents	Total	Word Start	Word End
p	18.0%	8.4%	3.8%
t	33.6%	7.2%	13.4%
k	26.8%	10.4%	6.1%
m	17.5%	5.7%	4.0%
n	26.8%	2.3%	9.4%
ŋ	3.4%	0.0%	0.9%

Table 7.7: Percentages of types with relevant stop and nasal sounds, in CELEX English database

Token Percents	Total	Word Start	Word End
p	4.2%	2.1%	1.1%
t	18.1%	4.0%	10.4%
k	6.2%	2.3%	2.7%
m	7.4%	3.1%	3.4%
n	18.2%	2.4%	7.4%
ŋ	1.2%	0.0%	0.7%

Table 7.8: Percentages of tokens with relevant stop and nasal sounds, in CELEX English database

Comparing these type and token percentages across the two languages reveals a great number of similarities between English and Dutch. One important similarity is the preponderance of coronal stops and nasals at the end of words in both languages. Coronal stops occur at the end of 13.4% of all unique English lemmas, while 8.4% of these lemmas end in coronal nasals. Similarly, 10% of all unique Dutch lemmas end in coronal stops, and 4.8% of Dutch lemmas end in coronal nasals. For stops and nasals at the ends of words in both English and Dutch, then, coronals are the most frequent place of articulation. Tables 7.9 and 7.10 present this generalization in a different way. They show the conditional probabilities that, given the presence of either a nasal or a voiceless stop at the end of a word in either English or Dutch, the word will end with either a labial, coronal or dorsal place of articulation. In all cases, the conditional probability of those stops or nasals having a coronal place of articulation is greater than 50%.

	p	t	k
Types	17.4%	55.1%	27.5%
Tokens	8.2%	70.6%	21.2%
	m	n	ŋ
Types	35.8%	54.6%	9.6%
Tokens	5.1%	92.6%	2.3%

Table 7.9: Conditional type and token probabilities for place of articulation at the ends of Dutch lemmas, given either a voiceless stop or nasal manner of articulation

	p	t	k
Types	16.4%	57.5%	26.1%
Tokens	7.7%	73.1%	19.2%
	m	n	ŋ
Types	28.2%	65.7%	6.2%
Tokens	29.4%	64.6%	5.9%

Table 7.10: Conditional type and token probabilities for place of articulation at the ends of English lemmas, given either a voiceless stop or nasal manner of articulation

It is perhaps not coincidental that only segments with the most predictable place of articulation, at the ends of words, undergo place assimilation in casual speech in English and Dutch. The predictability of the coronal place of articulation in this context may bias listeners towards perceiving a coronal stop or nasal, given enough evidence for either of those manners of articulation, even when there is little evidence for the coronal place of articulation in the acoustic signal. The results of Winters (2001, 2002) show, in fact, that English listeners are biased towards coronal responses in a place identification task.

Coronals which have undergone place assimilation in casual speech in English and Dutch may therefore still be perceived as having a coronal place of articulation, due to a frequency-based perceptual bias in English and Dutch listeners. Place assimilation in this case would not have a significant effect on the listeners' ability to perceptually reconstruct the speaker's intended utterance. The fact that listeners might not be able to reconstruct labials or dorsals in this way could prevent these segments from undergoing a similar process of place assimilation in these two languages. In this analysis, casual speech place assimilation should only target the most frequent place of articulation in a language, regardless of what the physical content of that place of articulation actually is.

It is interesting to note that the particular patterns of place assimilation in Dutch and English seem to be the result of the interaction of more than one independent or external force on the phonology of those languages. The direction of assimilation--which, in both of these languages, is regressive--seems to be determined by perception, as numerous perceptual studies (Malecot 1956, Fujimura et al. 1978, Ohala 1990) have shown that place cues are more salient in onset position than in coda position. The particular place of articulation--coronal--which is targeted in both languages seems to be conditioned by the high frequency of these segments at the ends of words in English and Dutch. Lastly, articulatory constraints seem to condition the manners of articulation which are targeted by place assimilation in the two languages--stops and nasals in English, and only nasals in Dutch. Assimilation induces place assimilation more often, cross-linguistically, in nasals than in stops, because heterorganic nasal-stop clusters are more difficult to produce accurately and consistently than stop-stop clusters are (as was shown in Chapter 6). This analysis seems to indicate that there are at least three “external” forces working together to mold the processes of place assimilation which exist in English and Dutch. Whether or not these various forces operate independently of one another may be left to future theoreticians to decide.

Whatever the particular relationship between the individual forces shaping phonological systems, however, this interpretation of the interaction of phonetic forces with place assimilation differs from any of the more formal approaches which preceded it. The ostensible objective of any theoretical treatment of place assimilation is to represent what speakers have to "know" in order to make this process work the way it does. While early approaches characterized this knowledge in terms of logical-formal relationships between the input and output of the process, later treatments refined their representations of this knowledge according to what kinds of assimilations commonly occurred throughout the languages of the world. Some researchers--such as Kohler (1990) and Jun (1995)--suggested that the linguistic knowledge required to make assimilatory processes work is based on universal perceptual and articulatory factors. Perceptual and articulatory forces were, in a sense, indirectly encoded into a speaker's grammar in Jun (1995). Ohala (1990), on the other hand, theorized that knowledge was not required to account for certain patterns in place assimilation; these were the result, instead, of "innocent misapprehensions" on the part of listeners. Rather than having perceptual forces encoded directly in the grammar, then, Ohala (1990) suggested that, in a sense, the grammar of place assimilation itself was agrammatical; it just occurred, without any particular knowledge on the part of the speaker intervening.

Hume & Johnson's (2001) model also excludes perceptual forces from phonological theory per se, but does allow for interaction between phonetic “filters,” such as perception and articulation, and language-specific phonological patterns. The evidence garnered from the results of the perception and production experiments in this study indicate that this may be the appropriate way to conceive of the relationship between the phonetics and phonology of place assimilation processes in English and Dutch. The particular patterns of place assimilation in English and Dutch seem to be most closely related to the variability and accuracy of English and Dutch speakers' productions of heterorganic stop-stop and nasal-stop clusters. While these results revealed an overall, cross-linguistic trend towards better reproduction accuracy for stops over nasals in such clusters, the language of the speaker seemed to have an effect on the variability in the duration of (and formant transitions into) these consonant clusters. Both of these results seem to account for the fact that nasals may undergo casual speech place assimilation in both English and Dutch, but that only stops undergo this process in English. The articulatory "filter" on patterns in place assimilation therefore seems to have both universal and language-specific aspects. These universal aspects may simply reflect characteristics of the articulatory system and may not, therefore, have any relevance for the cognitive representation of language. The language-specific aspects, on the other hand, must have their origins in language-specific experience or knowledge, neither of which can be completely "substance-free" if they can alter the articulatory and perceptual abilities of the users of language.

The fact that more than one "external" force may have an effect on the same process in phonology underscores the fact that none of them should be studied in complete isolation from the rest. Throughout the history of phonology, many researchers have perhaps gone too far in this direction in making the case for the importance of studying perception's influence on phonology. Such efforts have an historical tendency to postulate that perceptual factors alone account for more than they actually can. Such analyses have generally been corrected by subsequent research revealing that only articulatory factors can account for certain patterns in phonology. The research in this dissertation provides another example in this long line of theoretical corrections. It establishes that the cross-linguistic tendency of nasals to undergo place assimilation more often than stops has its origins in articulatory difficulties, rather than perception. This conclusion was only reached, however, after carefully testing the predictions of both perceptual and articulatory explanations of the same phenomenon. It seems that addressing both of these perspectives at once--rather repetitively switching the theoretical focus of phonology back and forth between them--might therefore save the theoretical world much wasted time and a lot of unjustified rhetoric.

The results of this study have also added, in a small way, to the growing body of literature which maintains that it is similarly unwise to study the "external" effects of perception and articulation without considering the role that language-specific knowledge or experience plays in making particular sound patterns easy or difficult for speakers of a language to perceive or say. The fact that speakers of English and Dutch target different segments in place assimilation seems to be motivated by language-specific differences in their abilities to produce these segments in heterorganic consonant clusters. Articulation and perception may not, therefore, only influence phonology as external and independent forces--mere products of the physical system on which phonology is imposed--but also as the product of language-specific development in articulatory and perceptual abilities.

The formal analysis of place assimilation has been consistently refined since the publication of SPE by eliminating its ability to represent assimilatory processes which are cross-linguistically unlikely or unattested. Establishing likely patterns in place assimilation--or cross-linguistic implicational relationships between the targets, triggers, and conditions of assimilation--have thus helped flesh out the theoretical representation of this process. The evidence from this study, however, suggests that place assimilation may also be constrained on a language to language basis by language-specific articulatory and perceptual abilities. Considerations of the interaction between such abilities and the frequency of certain sound patterns in English and Dutch even seem to suggest that language-specific experiences and knowledge may interact with articulation and perception to mold the particular patterns of place assimilation in any given language. Further pursuing knowledge of the apparent interaction between frequency, articulation, perception and phonological processes should provide a clearer picture of the language-universal and language-specific limits on the process of place assimilation.

BIBLIOGRAPHY

Abercombie, D. 1967. *Elements of General Phonetics.* New York: Aldine.

Archangeli, D. 1984. Underspecification in Yawelmani phonology and morphology. PhD Dissertation. MIT.

Baayen, R.H., Piepenbrock, R., and Gulikers, L. 1995. The CELEX Lexical Database (Release 2) [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor].

Bach, E. 1968. Two proposals concerning the simplicity metric in phonology. *Glossa* **4**, 3-21.

Barry, M. 1991. Temporal modelling of gestures in articulatory assimilation. *Proceedings of the XIIth International Congress of Phonetic Sciences* **4**, 14-17.

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

Booij, G. 1995. The Phonology of Dutch. Oxford: Oxford University Press.

Browman, C.P. and Goldstein, L. 1986. Towards an articulatory phonology. *Phonology Yearbook* **3**, 219-252.

Browman, C.P. and Goldstein, L. 1989. Articulatory gestures as phonological units. *Phonology* **6 (2)**, 201-251.

Browman, C.P. and Goldstein, L. Tiers in articulatory phonology with some implications for casual speech. In *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech* (J. Kingston and M. Beckman, eds.), 341-376. Cambridge: Cambridge University Press.

Brown, G. 1977. *Listening to Spoken English.* London: Longman.

Bybee, Joan. 2000. The Phonology of the Lexicon: Evidence from Lexical Diffusion. In *Usage-Based Models of Llanguage* (M. Barlow and S. Kemmer, eds.). Stanford, CA: CSLI Publications, Center for the Study of Language and Information.

Catford, J.C. 1977. *Fundamental Problems in Phonetics*. Bloomington: Indiana University Press.

Cho, Y.-M. 1988. Korean assimilation. *Proceedings of West Coast Conference on Formal Linguistics* **7**, 41-52.

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

Chomsky, N. and Halle, M. 1968. *The Sound Pattern of English*. New York: Harper.

Clements, G.N. 1985. The geometry of phonological features. *Phonology Yearbook* **2**, 225-252.

Clements, G.N. and Hume, E. 1996. The internal organization of speech sounds. In *The Handbook of Phonological Theory* (J. Goldsmith, ed.), 245-306. Cambridge: Blackwell.

Cooper, F. 1950. *Journal of the Acoustical Society of America* **22**, 761-762.

Cooper, F., Delattre, P., Liberman, A., Borst, J. and Gerstman, L. 1952. Some experiments on the perception of synthetic speech sounds. *Journal of the Acoustical Society of America* **24**, 597-606.

Cooper, F., Liberman, A., and Borst, J. 1951. *Proceedings of the National Academy of Science* **37**, 318-325.

de Courtenay, B. 1895. An attempt at a theory of phonetic alternations. In *A Baudouin de Courtenay Anthology.*

Cressey, W. 1974. Homorganic in generative phonology. *Papers in Linguistics* **7**, 69-81.

Delattre, P., Liberman, A. and Cooper, F. 1955. Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America* **27**, 769-773.

Edwards, J., Beckman, M.E. and Munson, B. 2003. The interaction between vocabulary size and phonotactic probability effects on children' s production accuracy and fluency in nonword repetition. ms. Ohio State University and University of Minnesota.

Fowler, C.A. and Dekle, D.J. 1991. Listening with eye and hand: cross-modal contributions to speech perception. *Journal of Experimental Psychology: Human Perception and Performance* **17**, 816-828.

Frisch, S. 1996. Similarity and frequency in phonology. PhD dissertation. Northwestern 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.

Gimson, A.C. 1962. *An introduction to the pronunciation of English.* London: Edward Arnold.

Goldsmith, J. 1976. Autosegmental phonology. PhD Dissertation, MIT.

Goldsmith, J. 1981. Subsegmentals in spanish phonology. In *Linguistic studies in the romance languages* **9** (W. Cressey and D. J. Napoli, eds.). Washington: Georgetown University Press.

Greenberg, J.H. and Jenkins, J.J. 1964. Studies in the psychological correlates of the sound system of American English. *Word* **20**, 157-177.

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

Halle, M. and Vergnaud, J.R. 1980. Three dimensional phonology. *Journal of Linguistic Research* **1**, 83-105.

Harms, R. 1966. Stress, voice and length in southern Paiute. *International Journal of American Linguistics* **2(3)**, 228-235.

Harnsberger, J. and Pisoni, D. 1999. Eliciting speech reduction in the laboratory II: calibrating cognitive loads for individual talkers. *Progress Report No. 23, Indiana University Speech Research Laboratory*.

Hayes, B. 1986. Assimilation as spreading in Toba Batak. *Linguistic Inquiry* **17**, 467-499.

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

Huang, T. 2001. The interplay of perception and phonology in tone 3 sandhi in Chinese Putonghua. *Ohio State University Working Papers in Linguistics* **55**, 23-42.

Hume, E. 1994. *Front Vowels, Coronal Consonants and their Interaction in Nonlinear Phonology*. New York: Garland.

Hume, E. 1998. The role of perceptibility in consonant/consonant metathesis. *Proceedings of West Coast Conference on Formal Linguistics* **17**, 293-307.

Hume, E. and Johnson, K. 2001. A model of the interplay of speech perception and phonology. In *The Role of Speech Perception in Phonology (*E. Hume and K. Johnson , eds.),, 3-26. New York: Academic Press.

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.

Jakobson, R., Fant, G. and Halle, M. 1963. *Preliminaries to Speech Analysis*. Cambridge, MA: MIT Press.

Jones, D. 1956. *An Outline of English Phonetics.* 8th ed. Cambridge: Heffer & Sons.

Joos, M. 1948. Acoustic phonetics. *Language* **24** (Suppl.), 1-137.

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

Kim-Renaud, Y.-K. 1974. Korean consonantal phonology. PhD Dissertation. University of Hawaii.

Kiparsky, P. 1982. Lexical morphology and phonology. In *Linguistics in the Morning Calm* (I. Yang, ed.). Seoul: Hanshin.

Kiparsky, P. 1985. Some consequences of lexical phonology. *Phonology Yearbook* **2**, 83-138.

Kirchner, R. 1998. An effort-based approach to consonant lenition. PhD Dissertation. UCLA.

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

Kohler, K. 1991. The phonetics/phonology issue in the study of articulatory reduction. *Phonetica* **48**, 180-92.

Kohler, K. 1992. Gestural reorganization in connected speech: a functional viewpoint on 'articulatory phonology'. *Phonetica* **49**, 205-211.

Ladefoged, P. 1975. *A Course in Phonetics*. 2nd ed. New York: Harcourt Brace Jovanovich.

Liberman, A. 1957. Some results of research on speech perception. *Journal of the Acoustical Society of America* **29**, 117-123.

Liberman, A., Delattre, P., Cooper, F. and Gerstman, L. 1954. The role of consonant-vowel transitions in the perception of stop and nasal consonants. *Psychological Monographs* **68**, 1-13.

Liljencrants, J. and Lindblom, B. 1972. Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* **48**, 839-862.

Lindblom, B. 1983. Economy of speech gestures. In *Speech Production* (P. MacNeilage (ed.)). New York: Springer Verlag.

Lindblom, B., MacNeilage, P. and Studdert-Kennedy, M. 1983. Self-organizing processes and the explanation of phonological universals. In *Explanations of Linguistic Universals* (B. Butterworth, B. Comrie and O. Dahl, eds.). The Hague: Mouton.

Lindblom, B. and Sundberg, J. 1971. Acoustical consequences of lip, tongue , jaw and larynx movement. *Journal of the Acoustical Society of America* **50**, 1166-1179.

Lindblom, B., Pauli, S. and Sundberg, J. 1974. Modeling coarticulation in apical stops. In *Proceedings of the Speech Communication Seminar*, *Speech Communication*, **2**. (G. Fant, ed.). Stockholm: Almqvist and Wiksell.

MacKay, I.R.A. 1978. *Introducing Practical Phonetics*. Boston: Little, Brown.

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

Mascaro, J. 1983. Phonological levels and assimilatory processes. ms., University of Barcelona.

Marlett, S. 1981. The structure of Seri. PhD Dissertation. University of California, San Diego.

Mielke, J. (To appear). The diachronic influence of perception: experimental evidence from Turkish. *Proceedings of Berkeley Linguistics Society* **29**.

Mohanan, K.P. 1991. On the bases of radical underspecification. *Natural Language and Linguistic Theory* **9**, 285-325.

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.

Munson, B. 2000. Phonological pattern frequency and speech production in children and adults. PhD Dissertation. The Ohio State University.

Nolan, F. 1992. The descriptive role of segments: evidence from assimilation. In *Laboratory Phonology II* (G. Docherty and R. Ladd, eds.), 261-280. Cambridge: Cambridge University Press.

Nosofsky, R.M. 1992. Similarity scaling and cognitive process models. *Annual Review of Psychology*, **43**, 25-53.

Odden, D. 1987. Dissimilation as deletion in Chukchi. In *Proceedings from the fourth Eastern States Conference on Linguistics, Vol. 3* (A. Miller and J. Power, eds.), 235-246. Columbus, Ohio: Ohio State University.

Ohala, J.J. 1981. The listener as a source of sound change. In *Papers from the Parasession on Language and Behavior: Chicago Linguistics Society* (C.S. Masek, R.A. Hendrik, M.F. Miller, eds.), 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.

Ohala, J.J. and Ohala, M.. 1993. The phonetics of nasal phonology: theorems and data. In *Phonetics and Phonology, Vol. 5: Nasals, nasalization and the velum* (M.K. Huffman and R.A. Krakow (eds.)), 225-249.

Paradis, C. and Prunet, J.-F. 1991. Asymmetry and visibility in consonant articulations. In *Phonetics and Phonology, vol. 2: The special status of coronals* (C. Paradis and J.-F. Prunet, eds.), 1-28. San Diego: Academic Press.

Pierrehumbert, J. 2001. Why phonological constraints are so coarse-grained. In SWAP special issue: language and cognitive processes (J. McQueen and A. Cutler, eds.), **16 (5/6)**, 691-698.

Podgorny, P. and Garner, W.R. 1979. Reaction time as a measure of inter- and intra-object visual similarity in letters of the alphabet. *Perception & Psychophysics* **21**, 37-52.

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

Potter, Kopp and Green. 1947. *Visible Speech*. New York: D. Van Nostrand.

Prince, A. and Smolensky, P. 1993. Optimality theory: constraint interaction in generative grammar. ms. Rutgers University and University of Colorado.

Repp, B.H. 1982. Perceptual assessment of coarticulation in two-stop sequences. *Haskins Lab Status Report on Speech Research.* New Haven: Haskins Laboratories.

Rhee, S.-C. 1999. Keeping and losing contrast: spirantization and release in Assamese and other Indic languages. *Studies in Phonetics, Phonology and Morphology* **5**, 357-384.

Sagey, E. 1986. *The representation of features and relations in non-linear phonology*. PhD Dissertation, MIT.

de Saussure, F. 1983. *Course in General Linguistics*. C. Bally and A. Scheheye (eds.), R. Harris (tr.). London: Duckworth.

Shepherd, R.N. 1978. The circumplex and related topological manifolds in the study of perception. In *Theory Construction and Data Analysis in the Behavioral Sciences* (S. Shye, ed.). San Francisco: Jossey-Bass.

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

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

Smith, D. 1978. Temporal aspects of english speech production: a developmental perspective. *Journal of Phonetics* **6**, 37-67.

Smith, D. 1992. Relationships between duration and temporal variability in children's speech. *Journal of the Acoustical Society of America* **91**, 2165-2174.

Steriade, D. 1982. *Greek prosodies and the nature of syllabification*. PhD Dissertation, MIT.

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.

Stevens, K.N. 1989. On the quantal nature of speech. *Journal of Phonetics* **17**, 3-46.

Stevens, K.N. and Blumstein, S.E. 1978. Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America* **64**, 1358-1368.

Takane, Y. and Sargent, J. 1983. Multidimensional scaling models for reaction times and same-different judgments. *Psychometrika* **48 (3)**, 393-423.

Trubetzkoy, N.S. 1969. *Principles of Phonology*. C. Baltaxe (tr.). Berkeley and Los Angeles: University of California Press.

Tserdanelis, G. 2001. A perceptual account of manner dissimilation in Greek. *Ohio State University Working Papers in Linguistics* **55**, 172-199.

Tserdanelis, G. and Hume, E. 2000. Nasal assimilation in Sri Lankan Portugese Creole: implications for markedness. Paper presented at the Montreal Ottawa Toronto Workshop on Phonology. Toronto, February 4-6.

Van Oostendorp, M. 2001. Nasal consonants in variants of Dutch and some related systems. www.neerlandistiek.nl/01/08.

Vitevitch, M. S. and Luce, P. 1999. Probabilistic phonotactics and neighborhood activation in spoken word recognition.

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. 2001. VCCV perception: putting place in its place. *Ohio State Working Papers in Linguistics* **55**, 70-87.

Winters, S. 2002. Perceptual influences on place assimilation: a case study. Manuscript submitted for publication. Ohio State University.

Yoneyama, K., Beckman, M. E., and Edwards, J. 2001. Phoneme frequencies and acquisition of lingual stops in Japanese. Manuscript submitted for publication.

Zipf, G. 1949. *Human behavior and the principle of least effort.* Cambridge: Addison-Wesley.

Zsiga, E.C. 1994. Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics* **22**, 121-140.