



This is to certify that the dissertation entitled

### NATIVE AND NONNATIVE DIFFERENCES IN THE PERCEPTION AND PRODUCTION OF VOWELS

presented by

# **DENNIE HOOPINGARNER**

has been accepted towards fulfillment of the requirements for the

degree in

Doctoral

Linguistics

fam Major Professor's Signature

11 May 2004 Date

MSU is an Affirmative Action/Equal Opportunity Institution

### LIBRARY Michigan State University

PLACE IN RETURN BOX to remove this checkout from your record. TO AVOID FINES return on or before date due. MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
APR 31 04 2006		
NO VO 18 10 2006 0		
		6/01 c:/CIRC/DateDue.p65-p.

# NATIVE AND NONNATIVE DIFFERENCES IN THE PERCEPTION AND PRODUCTION OF VOWELS

By Dennie Hoopingarner

# A DISSERTATION

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

## DOCTOR OF PHILOSOPHY

Department of Linguistics and Germanic, Slavic, Asian, and African Languages

### ABSTRACT

### NATIVE AND NONNATIVE DIFFERENCES IN THE PERCEPTION AND PRODUCTION OF VOWELS

By

### Dennie Hoopingarner

That nonnative speakers typically speak with a foreign accent is uncontroversial, but the extent to which production reflects the linguistic systems of second language learners is less clear. The relationship between perception and production of nonnative speakers has not been examined in a thorough and objective way.

In a perception task, a computer program presented participants with a continuum of synthesized speech samples that represented the vowel space. For each of 11 monothongal English vowels, participants chose the sounds that they judged was the best exemplar of the vowel. Native and nonnative speaker groups participated in the study. In a production task, participants read a word list that contained the same English vowels. Acoustic analysis software was used to extract the formant values of the vowels.

Results indicate that nonnative speakers show more variation in their language systems, and that there is less consistency between perception and production among nonnative speakers than among native speakers. In addition, there is evidence that the native language influences both the perception and production of the second language. To my family. Thank you for your support, love and understanding.

### ACKNOWLEDGEMENTS

I would like to thank the members of my committee, Alan Beretta, Yen-Hwei Lin, Susan Gass and Grover Hudson, for their help as I worked on this dissertation. The document is stronger as a result of their input. Grover was especially helpful as chair of the dissertation committee, reading through many versions of the document, and providing comments and suggestions. While working with me, he managed to address issues with the organization and wording of my writing, as well as big-picture issues of theory and analysis.

Thanks also to Dennis Preston, for our many informal talks, usually in his or my office, or in the hallway in between. Dennis shared his experience in synthesized speech and acoustical analysis, and provided insight into variation within a speech community, that helped in my analysis. The staff at the Language Learning Center, under the management of Michael Kramizeh, configured the computer lab into a customized environment for my data collection.

My friends and professional colleagues both at Michigan State University and at other institutions were supportive of my efforts to complete this dissertation, and I thank you all for your kind words of encouragement.

I would not have been able to get as far as I have without the support and love of my family. They put up with me hiding in my study for many, many evenings while I was writing. My wife Stacy, and my sons Ian and Evan, made it possible for me to dedicate the time necessary to completing the dissertation, and gave me the reason to persevere. Thank you for believing in me and for encouraging me to keep going.

iv

# TABLE OF CONTENTS

LIST OF	TABLES	vii
LIST OF	FIGURES	ix
CHAPT	ER 1	
INTROL	DUCTION	1
1.1	Scope of the Dissertation	1
1.2	Context of the study	2
1.2.1	Competence and performance in a speech community	2
1.2.2	Acceptability tasks in first and second language	4
1.2.3	Limitations of previous studies	5
1.3	Importance of this study	6
1.3.1	Comparison between native and nonnative speakers	6
1.3.2	Objective measure of perception and production	6
1.4	Theoretical background	7
1.4.1	Perception of speech	7
1.4.1.1	Problems of speech perception	7
1.4.1.2	Categorical perception	9
1.4.1.3	Normalization	10
1.5	Research questions and hypotheses	15
1.6	Summary	19
CHAPT	ER 2	
REVIEW	V OF PREVIOUS STUDIES	20
2.0	Introduction	20
2.1	First language transfer and	
	adult second language acquisition	21
2.2	Speech perception	25
2.2.1	Language and general cognition: Voice Onset Time	25
2.2.2	Language-specific: Perceptual magnets	27
2.2.3	The influence of language experience	33
2.3	Perception and production of second-language vowels	38
2.3.0	Introduction	38
2.3.1	Perception of first and second language vowels in adults	39
2.3.1.1	First language vowel perception	39
2.3.1.2	Perception and production of L2 vowels	42
2.4	Acoustic comparison of Korean and English Vowels	49

# CHAPTER 3

METHO	DD	
3.0	Introduction	56
3.1	Participants	56
3.1.1	Native speakers	56
3.1.2	Nonnative speakers	57

3.2	Methodology	
3.2.1	Rationale for the design of the instrument	
3.2.2	Synthesized speech samples	59
3.2.3	The vowel matrix	59
3.3	Procedure	61
3.3.1	Introduction	61
3.3.2	Task 1: Perception	61
3.3.3	Task 2: Production	63
3.4	Data	64

# CHAPTER 4

RESU	LTS	66
4.0	Introduction	66
4.1	Native speakers	67
4.1.1	Task 1: Perception	67
4.1.2	Production	74
4.2	Nonnative speakers	82
4.2.1	Perception	82
4.2.2	Production	
4.3	Perception-production differences	
4.3.0	Introduction	
4.3.1	Intra-group perception versus production	
4.3.2	Cross-group perception and production	

# CHAPTER 5

DISCUS	SION AND CONCLUSIONS 104
5.0	Introduction
5.1	Review of research questions and hypotheses 104
5.2	Evaluation of hypotheses 106
5.2.1	Hypothesis 1
5.2.2	Hypothesis 2 109
5.2.3	Hypothesis 3110
5.2.4	Summary 110
5.3	Additional findings of the study 111
5.3.1	Production less variable than perception111
5.3.2	Conflation of /a/ and /ɔ/ in native speakers' perception 111
5.4	Conclusions 112
5.4.1	Research question 1 113
5.4.2	Research question 2114
5.4.3	Research question 3 115
5.4.4	Research question 4116
5.5	Areas for further research
Reference	ces

# LIST OF TABLES

Table 15: Means for the nonnative speaker production task 92
Table 16: The results of a pairwise comparison from an ANOVA test shows the source of variation between pairs of vowels in nonnative speakers' responses to the production task for the F1 formant (Top) and F2 formant (Bottom). Several pairs of vowels $(/i/-/I/, /e/-/a/, /\epsilon/-/a/, /a/-/o/, /a/-/o/, /o/-/u/,$
/o/-u/, and /u/-/u/) showed no significant difference between them
Table 17: Average values for perception and production, native and nonnative speakers
Table 18: The results of ANOVA tests comparing perception and production among native speaker subjects for F1 (Top) and F2 (Bottom). Only the vowel pair $/a/$ - $/3/$ shows no significant difference (p>.05) between pairs for both F1
and F2
Table 19: The results of ANOVA tests comparing perception and production among native speaker subjects for F1 (Top) and F2 (Bottom). The vowel pairs /i/-/I/, $/e/-/e/,/e/-/æ/,/e/-/a/,/o/-/o/,and /u/-/u/ show no significantdifference (m. 05) between pairs for both E1 and E2$
unterence $(p>.05)$ between pairs for both F1 and F2
Table 20: Standard deviations for F1 and F2 values of vowels in the perception and production tasks for native and nonnative speakers. For each vowel, the first line is the perception value, and the second line is the production value. The column "PB" is the values from Peterson and Barney (1952). The value for the group with the greatest variation is in boldface
Table 21: The results of ANOVA tests comparing perception between native and nonnative speaker subjects for F1 (Top) and F2 (Bottom). The vowel pairs $/e/-/\epsilon/$ , $/a/-/3/$ , $/a/-/3/$ show no significant difference (p>.05) between pairs for both F1 and F2
Table 22: The results of ANOVA tests comparing production between nativeand nonnative speaker subjects for F1 (Top) and F2 (Bottom). The vowelpairs $/I/-/e/$ , $/a/-/a/$ , $/o/-/u/$ show no significant difference (p>.05) betweenpairs for both F1 and F2.102
Table 23: Mean and standard variation values on the perception task for 11simple vowel phonemes of English, native and nonnative speakers. The highervalue of native and nonnative speakers is shown in boldface.107

# LIST OF FIGURES

Figure 1: The F1 and F2 values of vowel samples collected from adult male, adult female, and child speakers in Peterson and Barney's (1952) study plotted on a grid shows considerable variation within each vowel, and significant overlap between vowels
Figure 2: The formant frequency values F1 and F2 of vowels can be plotted onto an XY graph. Values for F2 are plotted along the X axis, and values for F1 are plotted on the Y axis
Figure 3: The X-Y grid maps formant frequency values of vowels in a way corresponding to positions of the vowels in the traditional vowel quadrangle 15
Figure 4: The Ontogeny Model (Major 1987)
Figure 5: Visual representation of tokens from Kuhl's 1991 study. Filled areas on the matrix indicate the F1-F2 values of vowel sounds. The areas labeled "P" and "NP" are the F1-F2 values of the token selections from Peterson and Barney (1952). The label "P" indicates the prototype sound that was closer to the sound produced by more subjects in Peterson and Barney's study. The sound labeled "NP" is the nonprototype sound, a sound that was produced by one subject, but deviates from the average of all subjects In the study to a greater degree than the "P" sound. The filled areas surrounding the prototypes are the F1-F2 values of vowel sounds that were synthesized for this study. The F1-F2 values of these sounds differ incrementally from the sample sounds
from Peterson and Barney
Figure 6 The average F1 and F2 values of American English and Korean vowels plotted on X-Y grids. The X axis is the F2 value, and the Y axis is the F1 value
Figure 7: The respective positions of the F1 and F2 values of the common vowels between English and Korean, /i e $\varepsilon$ a $\Rightarrow$ o u/, plotted on the same graph 53
Figure 8: The vowel matrix
Figure 9: Graphic representation of the vowel area . The gray shaded area represents F1 and F2 values that were not represented by sounds in the perception task, but that could possibly be produced by participants during the production task
Figure 10: Average of Native Speaker Perception
Figure 11: Native speaker choices for perception task for each vowel

Figure 12: Average Native Speaker Perception and Production	76
Figure 13: Native speaker performance on production task	77
Figure 14: Average Nonnative Speaker Perception	83
Figure 15: Nonnative speaker choices for perception task	84
Figure 16: the Nonnative vowel areas: perception task	88
Figure 17: Average Nonnative Speaker Perception and Production	90
Figure 18: Nonnative speaker performance on production task	91

#### Chapter 1: Introduction

#### 1.1. Scope of the Dissertation

This study examines the acquisition of a second language phonology. The goal of this study is to determine the extent to which nonnative speakers of English have acquired the vowels of English. I compare the perception and production of eleven monothongal English vowels by native and nonnative speakers of English. I then explore the relationship between perception and production within a speech community, and describe differences in perception and production between native and nonnative speakers of English.

Many studies of second-language phonology concentrate on the spoken output of nonnative speakers. This study adds the dimension of intuition data, acquired via a judgment task, as a means of measuring the acquisition of the second language vowel system. In a study of phonology, the use of intuition data, the type usually gathered in the form of tasks in second language acquisition research, has the advantage over production data in that foreign accent is eliminated as a factor in participants' responses.

This study seeks to measure participants' intuitions about eleven monothongal vowels in English, and compare that intuitional data to their performance on a production task. By comparing participants' intuitions with their productions, foreign accent can be isolated as an element of linguistic performance separate from linguistic competence. In order to establish baseline data with which to compare nonnative speakers, data from native speakers was also gathered. In this way, first-language competence and performance can be compared to second-language competence and performance.

#### 1.2. Context of the study

#### 1.2.1. Competence and performance in a speech community

Linguistic theory distinguishes between competence and performance in linguistic behavior among native speakers of a language. At least since Chomsky (1965:4) explicitly distinguished linguistic "competence" from "performance," the focus of linguistics has been on competence. The topic of inquiry is thus the abstract, internal system that a linguistic adult attains through the process of language acquisition. Chomsky (1995:14) characterized linguistic competence as "some array of cognitive traits and capacities, a particular component of the human mind/brain."

Although instantiated within the mind/brain of individuals, the grammars of members of a speech community have been assumed to be homogenous (Saussure 1916:19, Chomsky 1965:3). While individuals may exhibit different language performance, the underlying structure of the language system is uniform. It is commonly accepted that linguistic competence is a more reliable indicator of underlying structure than performance. Linguistic competence is unaffected by factors external to language, such as slips of the tongue or psychological factors such as memory and attention. These external factors can influence linguistic performance, however.

Phonological competence being more well-formed than performance is well documented in the literature on child language phonology. The so-called "fis phenomenon" (Berko and Brown 1960) is an example of how competence and performance can differ in an individual's grammar. It describes the condition in which a

child cannot produce a particular sound correctly (in this case, the sound /š/), but can distinguish it when uttered by others. This condition is developmental in nature, and is documented in the first-language literature, but it may be applicable to second language acquisition in that it indicates a system that is in transition.

Although it is a common assumption that competence should be the focus of linguistic theory, it can be very difficult to accurately determine the underlying structure based on the output. Production data is often riddled with errors. In second-language phonology, the difference in performance between native and non-native speakers is especially salient. Foreign accent is clearer marker of a speaker as non-native than syntax or morphology errors. First-language influence on the second-language sound system is a commonly-cited source of foreign accent. While the existence of foreign accent is not controversial, it is an open question whether non-nativelike production is an indicator of nonnative-like competence. In other words, the inability to produce the second language in a native-like manner is not necessarily indicative of the nature of the speaker's underlying phonological competence in the second language.

#### 1.2.2. Acceptability tasks in first and second language

The predominant method used to study the grammar of a language is to call upon the intuition of speakers of the language. A commonly-used instrument in investigations into second language acquisition of syntax is the acceptability task (Schachter 1989, White 1989, Broselow and Finer 1991, Al-Banyan 1996, Liu and Gleason 2002), in which the participant indicates whether a sentence in the target language is grammatical or not. There have been objections raised to this method. Sorace (1996) points out that this kind of task forces the participant to make a categorical, binary decision about a form, when the actual judgment of the form may be gradient. White (2003:17) also characterizes as a "myth" the assumption that acceptability tasks provide a direct reflection on second-language linguistic competence. Nonetheless, the acceptability task remains a commonly used tool for measuring linguistic structure is the least likely to be influenced by outside influences.

In the area of theoretical phonology, the use of acceptability judgments is also the means to discover underlying structure. In fact, it is argued that Optimality Theory (Prince and Smolensky 1993), as a method of analysis, codifies the process by which a native speaker arrives at a judgment about a phonological form by ranking constraints. However, in the field of second language phonology, little word has been done to measure participants' intuitions linguistic forms. Rather, the focus has been on measuring output and making inferences about the internal representation based on that output. In a sense, research that uses second language speakers' output as the only source of data may amount to a study of foreign accent, which may not necessarily represent the speaker's

internal system accurately. Studies in the role of the native language in speech perception, which is discussed in Chapter 2, indicate that a given speech stream can be perceived differently by speakers of different languages. Speech perception is thus part of an individual's grammar.

#### 1.2.3. Limitations of previous studies

As mentioned in the section above, perception and production often operate differently, and so focusing only on output may not give a complete picture of the linguistic system. Consistently in studies of second language phonology, the topic of study has largely been output. Nonnative-like production is often attributed to the influence of the first language on the second. Understanding the differences between the two languages was the goal of Lado's (1957) Contrastive Analysis, based on contrast between two languages. This method of analysis was applied as a method to predict errors in pronunciation. Eckman's (1977) Markedness Differential Hypothesis refined the prediction mechanism to include universal tendencies of phonological complexity. He used production data to evaluate this model.

Most studies of this nature drew their conclusions solely on the basis of an analysis of the participants' output. The measurement of the production data has typically been done using subjective measures. Many studies rely on the judgments of human raters to evaluate production (Eckman 1987, Broselow and Finer 1991, Eckman and Iverson 1994, Major and Faudree 1996). Few studies use computer-based speech analysis

tools to extract the acoustic properties of the speech that participants produced, such as in the procedure used by Flege (1987, 1991) and Flege et at (1997).

1.3. Importance of this study

1.3.1. Comparison between native and nonnative speakers

This study seeks to acoustically define specific vowel categories for both native and nonnative speakers of English, both in identification of vowels, and in production. The phonological phoneme inventory is in some respects a closed system. A given language has a finite number of consonant and vowel contrasts, and so this aspect of the system is quantifiable. Although differences between vowel inventories and systems in different languages has been documented for first languages in studies such as Schwartz et al (1997), there has been little empirical work to systematically map a nonnative speaker's entire vowel system, and compare it to that of a native speaker. The contribution of this study is the addition of this dimension of phonological competence.

1.3.2. Objective measure of perception and production

Another potential contribution of this study is the objective nature of the analysis. As was mentioned in section 1.1.2.3. above, in the bulk of second-language phonology studies, the data was analyzed after it was transcribed. Doucherty and Foulkes (2000: 113) point out potential problems with transcribed data, chief among which, there is always the danger that the data is not transcribed reliably. Even the best-trained phoneticians can inadvertently filter speech sounds through their own phonological systems. This can result in transcriptions that lack accuracy. The current study seeks to avoid that problem by using fully objective measures.

1.4. Theoretical background

#### 1.4.1. Perception of Speech

#### 1.4.1.1. Problems of speech perception

A prerequisite for oral linguistic communication is that the listener understand what the speaker says. That a native speaker of a given language can understand that language when it is spoken by another speaker is an empirical fact, and yet an empirical model for the process by which that occurs still eludes us. In a discussion of speech perception, Chomsky and Miller (1963) discuss two problems that underlie the phenomenon of speech perception: linearity and invariance. The problem of linearity is that the speech stream is not a progression of clearly delineated and acoustically recoverable phonemes, but is instead a fluid mixture of coarticulated sounds. Consonants and vowels are combined together in the speech stream through coarticulation. Hockett (1955: 210) made the analogy that recovering phonological elements from the speech stream is similar to trying to reassemble colored Easter eggs that have been crushed as they move along a conveyor belt. Liberman (1970) pointed out that it is impossible to take a tape recording of a CVC syllable, and cut the tape to isolate the consonants and the vowels. This is because consonants are represented in the speech stream as formant contours around vowels, inseparable from the vowel sound. The seminal study by Delattre et al (1955) showed that consonants can be synthesized by manipulating the first two formants in a vowel sound. By creating contours up or down from the value of the formants that produce the vowel sound, a consonant-like sound was produced. Consonant onset sounds corresponding to /b/, /d/ and /g/ were synthesized in this manner. Participants in the study reported hearing a consonant onset. Delattre et al concluded that the process of speech perception hinges on the formant frequencies of the utterance, and furthermore, that the first two formants, F1 and F2, carry enough information to distinguish both consonants and vowels. Even though consonants and vowels may not be objectively present in the speech stream, listeners are still able to perceive them.

Linearity is also instantiated in the phonological environment of the utterance. The shape of the vowel tract for the production of a given phoneme will differ depending on what sounds are to be articulated before and after it. Katamba (1989: 19) illustrates how coarticulation influences speech production with the example of the English phrase "car keys." The /k/ sound is produced by the tongue interacting with different parts of the oral cavity as the speaker anticipates the following vowel sound. The vowel /a/ is produced by opening the oral cavity, and so the /k/ in the word "car" is produced by the back part of the tongue touching the rear part of the soft palate. The vowel /i/ is produced with a raised tongue, and so the /k/ is "keys" is produced with the tongue more forward.

The second problem of speech perception is invariance, or speaker variation. An acoustic event in speech is not related in a one-to-one fashion to the phonemic element

that it invokes. One factor that causes variation in utterances is idiosyncratic characteristics of the vocal tract, such as length and shape, that make each speaker's voice distinctive. The length of the vocal tract determines the fundamental frequency of the speaker's voice. Since not every speaker has a vocal tract of the exact same length, different speakers uttering the same vowel sound will produce a sound with a different pitch. Changing the shape of the oral cavity by positioning the jaw, tongue and lips, produces different vowel sounds. Not every speaker will position their oral cavity exactly the same, and so the vowel sounds of different speakers will be more or less different.

The problems of speech perception can be approached within the framework of two theoretical concepts. First, using the concept of categorical perception, we can seek to define the acoustic threshold of what we perceive as a phonological element. This process seems to account for the linearity problem of speech perception. The "acoustic space" can be divided up into categories. Any elements of the speech stream that "fit" into a particular acoustic section are interpreted as a single segment. Secondly, the invariance problem can be addressed by models of speaker normalization. This theory attempts to describe how listeners compensate for inter-speaker differences in voicing and vowel quality.

#### 1.4.1.2. Categorical perception

Categorical perception can be used to explain the ability to discriminate between some sounds, and the inability to discriminate between others. The human brain has the ability to gather sounds of language into perceptual categories. Within a category,

variants of a vowel may differ in a measurable, absolute way, and yet the human brain can not distinguish the difference. Two variants of vowels that fall in different categories, however, can differ by the same small amount as between sounds in the same category, and participants can hear the difference between them.

The evidence of the use of categorical perception in language use has strong empirical support (Liberman et al 1957, Fry et al 1962, Lisker and Abramson 1964, Eimas et al 1971, Scoles 1968, Kuhl 1980, Gass 1984, Repp 1984, Werker and Tees 1984, Kuhl 1987, Kuhl 1991, Werker 1994). However, as Kuhl (1987) points out, the effects of categorical perception are not limited to language, and not limited to humans. It may be worth questioning whether categorical perception is linguistic in the modular sense of Fodor (1983), or whether, as part of general cognition, it is an ability that humans share with other animals (Harnad 1987). Nevertheless, it does seem to play a part in language, seems to be a part of the innate human cognitive system, and changes with language acquisition as the infant begins to specialize its sensitivity to salient differences in the native language (Kuhl et al 1992).

#### 1.4.1.3. Normalization

Every speaker's voice is different. Vowel productions vary across speakers. For example, the voices of male speakers, as a general rule, have lower F0 frequencies than those of female speakers. Differences in length of the vowel tract results in a distinct audio quality for each speaker. The seminal study by Peterson and Barney (1952) demonstrated that age and gender are the source of much variation in the production of vowels, largely due to physiological differences between speakers. Figure 1 shows the measurements of vowels collected from speech samples from speakers of different age and gender. There is considerable variation among speakers, and considerable overlap among vowels.



Figure 1: The F1 and F2 values of vowel samples collected from adult male, adult female, and child speakers in Peterson and Barney's (1952) study plotted on a grid shows considerable variation within each vowel, and significant overlap between vowels.

In addition, the acoustic values of vowels produced by any given speaker can vary as a result of context. However, the variation between speakers does not prevent the listener from understanding. When hearing the same utterance by different speakers, the listener perceives the utterance as the same in some sense. In order to understand spoken language, listeners must perceptually remove this speaker-specific variation and perceive the utterances using an objective norm. The process of removing the speaker-specific differences is referred to as normalization. Native speakers seem to compensate effortlessly and automatically for variations in the acoustic productions of different speakers.

The problem of understanding such normalization, as pointed out by Hindle (1978), comes to light when comparing two different speakers. The listener must transform the formant values of the two speakers so that they coincide for comparison. A common assumption is that vowel identification can be done via an objective process, using relative positions in the vowel space, similar to Joos' (1948:68) concept of a "template." Joos suggested that, in speech perception, the listener divides the vowel space into zones. Vowels that fall within a given zone are interpreted as that vowel. This division of the vowel space is a listener's template. The process of normalization, according to Joos, consists of mapping the speaker's utterances onto the vowel space perceived in such zones.

Later studies of speech perception (Ladefoged and Broadbent, 1957, Gerstman 1968, Lehiste and Meltzer 1973, Assman et al 1982) support the position that listeners use a subjective standard to perceive a given speaker's vowels. Studies of normalization techniques by Neary (1989) and Miller (1989) also support the position that listeners do not use objective measures to perceive vowels in terms of absolute formant values.

Normalization is the process of identifying vowel sounds in the speech stream by means of a norm. This norm is a system of vowels, which Kuhl (1991) calls "prototypes." An assumption of the process of normalization is that listeners have an internal system that they are referring to and trying to map the speech stream to. Members of the same speech community have the same internal system, and so normalization does not interfere with communication. Nonnative speakers of the language may or may not have the same

internal system, and so normalization may or may not cause comprehension problems. Part of the task of learning a second language is developing a system of vowels that is used in comprehending spoken language. One of the aims of this study is to explore to what extent second language learners have developed a system to use when normalizing English speech.

#### 1.4.2. Measuring vowels

The source-filter theory of vowel production (Fant 1960) sees vowel production as a combination of two factors. The vocal folds restrict the flow of air through the pharynx, producing vibration and sound. The sound passes through the oral cavity, where it resonates. The shape of the oral cavity makes the sound resonate at specific frequencies. The combination of the resonating frequencies produces a vowel sound. The speaker can change the shape of the oral cavity by altering the position of the jaw, tongue and lips, to produce different resonant frequencies that are perceived as different vowel sounds.

Analyzing vowels in terms of the source-filter theory allows us to isolate the factors that make vowels distinctive. A given speaker will produce a source that is consistent among utterances, because the quality of the source is determined by physiological factors such as gender and size of the vocal tract. The speaker's filter, which changes with the shape of the oral cavity, determines which vowel is produced. The wide variety of vowel sounds that can be produced is attributed to the ability to vary the position of the tongue, lips and jaw. It is the filter that makes vowel sounds distinctive, and so vowels are identified by analyzing the filter.

We can measure vowel sounds in terms of their component formant frequencies. Peterson (1951) and Delattre et al (1955) showed that vowels can be identified solely on the basis of the formant frequencies F1 and F2. Different values of the formant frequencies produce different vowel sounds. The values of the formants F1 and F2 for vowels can be plotted onto a graph to show their relative values. Figure 2 shows such a graph.



Figure 2: The formant frequency values F1 and F2 of vowels can be plotted onto an XY graph. Values for F2 are plotted along the X axis, and values for F1 are plotted on the Y axis.

Values for F2 are plotted along the X axis, and values for F1 are plotted on the Y axis. This grid differs from other grids in that the origin point in the vowel graph is in the upper-right hand corner, rather than in the lower-left hand corner. By formatting the graph in this manner, when vowels are plotted on the graph, their relative positions are consistent with the traditional vowel quadrangle, as shown in Figure 3. High front vowels such as /i/ have a relatively low F1 value and a relatively high F2 value. Plotting those vowels on this grid would put them in the upper-left hand quadrant, corresponding to their position on the vowel quadrangle. Similarly, high back vowels such as /u/, with their

low F1 and F2 values, would fall in the upper-right hand quadrant. Vowels like /a/ have a high F1 value and a medium-range F2 value.



Figure 3: The X-Y grid maps formant frequency values of vowels in a way corresponding to positions of the vowels in the traditional vowel quadrangle.

Studies by Peterson and Barkey (1952) and Hillenbrand et al (1995) showed that the frequency range for adult male speakers is 300 - 750 Hz for F1, and 1000 - 2400 for F2. These values would be higher for female voices, due to the shorter length of the vowel tract.

1.5 Research questions and hypotheses

This study examines second language speech perception and production, comparing the intuitions of native speakers and nonnative speakers intuitions about vowel phonemes in English, and comparing perceptions to production of the vowels by nonnatives. It seeks to map the English vowel system for both native and non-native speakers for perception and production. The specific research questions to be addressed by the dissertation are as follows:

- To what extent can we determine the acoustic properties of the ideal vowel phoneme of each monothongal vowel in English based on native speakers' intuitions?
- 2. To what extent do nonnative speakers of English agree on the ideal vowel phoneme of each monothongal vowel in English?
- 3. Are the intuitions of nonnative learners of English similar to the intuitions of native speakers with regard to the vowel phonemes of English?
- 4. How similar is production of vowels in English by both native speakers of English and nonnative learners of English to their respective identification of an ideal vowel?

The motivation behind research question (1) is the theory that native speakers have established phonemic categories for the vowels of their native language, and that they can use their intuitions to indicate whether or not a given token of speech falls within the category for a given vowel (Liberman 1970). This study asks whether the intuitions of native speakers converge on an acoustic value for a given vowel. The degree of consensus among the native speaker group can be compared to the degree of consensus among nonnative speakers.

The basis of research question (2) is the supposition that native speaker members of a speech community have a common linguistic system, including vowel categories. Nonnative speakers, as learners of the language, are assumed to have a language system that is different from native speakers'. It is unclear, however, to what extent nonnative speakers have a speech community, with vowel categories that are homogenous to an analogous degree that is found among native speakers. Research question 2 addresses this issue. Using the same task for native and nonnative speakers, this study measures the intuitions of nonnative speakers about the English monothongal vowels.

Research question (3) asks whether, if there is a speech community of nonnative speakers, the intuitions of native speakers are similar to those of nonnative speakers. This question is motivated by the assumption that an important part of acquiring a second language is acquiring the internal system of the language. Intuitional data can provide valuable insight into an individual's grammar.

Research question (4) builds on two assumptions. The first is that there is a difference between linguistic perception and production. The vowel sound that a speaker produces may be different from what the same speaker would identify as the correct vowel sound. The second assumption is that nonnative speakers' foreign accents are not necessarily an accurate reflection of their second-language perceptual vowel systems. The first language may influence a nonnative speaker's production of a vowel sound without influencing the perception.

This study seeks to address these questions by comparing the perception and production of English vowels by two groups of participants: native speakers from Michigan and nonnative speakers of English from Korea. With regard to research questions 1 and 2, it is expected that native speakers, as members of the same speech

community, will agree on vowel categories. Nonnative speakers will show more variation than the native speaker group.

The comparison of native and nonnative speaker participant groups, as indicated in research question 3, will show the extent to which nonnative speakers share the same intuition about vowel categories in English as native speakers do. This comparison will eliminate foreign accent as a factor, and will thus be a better indicator of the degree to which the participants have acquired the English vowel system.

Finally, a comparison of perception and production will reveal differences between the two groups. Research question 4 examines the role of accent in linguistic performance. Ideally, there would be no difference between perception and production. In the real world, we expect to see other factors influencing linguistic performance. Comparing native and nonnative speakers, we can expect to see a closer match among native speakers than nonnative speakers. Foreign accent is expected to emerge as a factor among nonnative speakers, resulting in a greater dissimilarity between perception and production. Native speakers are expected to show more consistency between perception and production.

The specific hypotheses of the study are as follows.

- Given a perception task in which participants identify an English vowel from a continuum of F1-F2 combinations (the "perception task"), there will be less variation among native speakers than among nonnative speakers.
- In the comparison of the performance of native speakers and nonnative speakers on the perception task and a "production task" in which participants produce words containing the English vowels, there will be less variation between the two

groups' performance on the perception task than there will be on the production task.

3. In a perception and a production task, there will be less variation between the two tasks among native speakers than among nonnative speakers.

# 1.6 Summary

Native speakers of a language have a grammar that is embodied in the speaker's mind/brain in the form of a linguistic competence. The linguistic competence of members of a speech community is relatively homogenous. An individual's language system influences both speech perception and production. Due to factors unrelated to competence, linguistic performance can differ among speakers. Performance is visible and measurable, but linguistic theory is more concerned with competence. Since it is a part of the mind/brain, competence cannot be measured directly. The acceptability task is one of the most popular ways to explore competence. It is commonly used in studies of second language syntax, but not in second language phonology.

This study seeks to use both performance and competence data on both native and nonnative speakers, to try to gain a more accurate picture of second language phonology. Fully objective measures will increase the accuracy of reporting and analysis.

Chapter 2: Review of previous studies

### 2.0. Introduction

In this chapter I review studies concerning four areas. The first area is factors that influence adult second language acquisition. A great deal of research points to qualitative differences between language acquisition in adults and children. As the participants of this study are adults, it is relevant to acknowledge these differences and how they can affect the results of the study.

The second area is speech perception. Speech perception is an important part of this study, and so it is relevant to define its domain. I review studies that examine the phenomenon of speech perception as a linguistic phenomenon as opposed to a general cognitive function. I then review studies that point to differences between perception of language in general as compared to perception of language-specific structures. Finally, I examine studies that explore the role of language experience on second language speech perception.

The third area that I discuss is the relationship between perception and production in second language acquisition. The potential for the two to differ in first language performance is well known, but is an area that has received relatively little attention in second language research. This study focuses on the difference between perception and production in second language, and the review of the literature is intended to give some background to the topic.

The fourth area is an acoustic comparison between the vowels of Korean and English. The nonnative speaker participants of this study are Korean, and the native speaker participants are American. This area uses data from a study that collected

production data from native Korean and American English participants producing vowels in their native languages. The data from that study can provide reference material for this study.

### 2.1. First language transfer and adult second language acquisition

As Gass (1996) points out, although the details of the role of the native language in second language acquisition have been a continuous subject of debate, the influence of the native language has not been in dispute. A common term for first-language influence on the second language is "transfer." White (1989) takes the position that UG is available to learners, but acknowledges the role of the first language in the formation of the second-language grammar. Leather and James (1996: 275-276) take as a base assumption that listeners use the phonetic categories of their native language when labeling auditory stimuli, a position taken earlier in the last century by Trubetzkoy (1969 [1939]: 52-53). Trubetzkoy saw the first language as a filter through which the second language was perceived:

"The phonological system of a language is like a sieve through which everything that is said passes. Only those phonic marks that are relevant for the identity of the phoneme remain in it... when [a person] hears another language spoken he intuitively uses the familiar 'phonological sieve' of his mother tongue to analyze what is said. However, since this sieve is not suited for the foreign language, numerous mistakes and misinterpretations are the result. The sounds of the foreign language receive an incorrect phonological interpretation since they are strained through the 'phonological sieve' of one's own mother tongue."

Work on loanwords also suggests that the native language is operant in the process by which a foreign word is altered to fit the structure and constraints of the adopting language. As foreign words become part of the native language, something of their structure often changes as the words are reinterpreted by the adopting language's phonology. Silvermann (1992) proposed a two-stage process by which a word is first perceived by native speakers of the adopting language, and then a structure for the word is abstracted out inductively. In Kenstowicz' (2003) Optimality Theory analysis of Fijian loanwords, the process of nativization consists of trying to maintain faithfulness to salient aspects of the foreign word by making use of repair strategies in the native language grammar.

Odlin (1989), in his discussion of language transfer, distinguishes transfer effects in various aspects of phonology. The characteristics of the native language phonological categories, phonemes, rules and syllable structure can all influence the interlanguage phonology.

Even phonological categories that are similar in both the first and second languages often differ at the subphonemic level. Flege (1987) describes phonetic difference between the Dutch and English phonemes /i/ and /u/. Dutch /i/ has lower F2 and higher F1values than English /i/, placing it lower in the acoustic space than English /i/. Dutch /u/ has a higher F2 value than English /u/, making it closer phonetically to English /u/ than to English /u/. Flege found that Dutch learners of English produced /u/ with less accuracy than native speakers, due to the influence of the Dutch system. Riney and Takagi (1999) investigated the influence of the differences in the voice onset time (VOT) in /p t k/ in Japanese and English had on the production of the English consonants by native speakers of Japanese. They found that although native-like VOT correlated to overall proficiency in English production, VOT values did not become more English-like over time. That is, participants who had native-like VOT did not seem to develop the
native-like pronunciation over time. Riney and Takagi attributed the lack of change over time to the close similarity between Japanese and English VOT values. Flege (1987) claims that second-language sounds that are perceived by the learner as equivalent to a native-language sound will be produced as the native-language sound. Riney and Takagi claim that this "equivalence classification" is operational in their study, and that it tended to prevent native-like acquisition by their participants. The authors suggest that the influence of the L1 phonology was so strong that it impeded the L2 pronunciation of many participants.

Phonemic differences between languages can also influence interlanguage. Studies by Sheldon and Strange (1982), Yamada et al (1997), Takagi (2002) and many others have documented the chronic inability of Japanese learners to correctly distinguish and produce the English segments /r/ and /l/. Even with intensive discrimination training, learners still do not uniformly develop a native-like system for distinguishing between the two phonemes. The difficulty of the learners seems to stem from the fact that in the acoustic space where Japanese has only one phoneme, English has two. The influence of the first language seems to interfere with the establishment of new phonemic categories in the second language. Wayland (1997) studied the acquisition of Thai consonants, vowels, and tones by native English speakers, and found significant native-language effects for all categories.

Major's (1987) Ontogeny Model of second language phonology development is an effort to account for developmental patterns in the interlanguage, while explaining the role of the first language in the second language system. According to the model, the nature of errors in the second language will change over time. In the early stages of

acquisition, the influence of the first language will dominate the phonological system, causing interference-like errors. These errors will decline over time, as the learners acquire the second language phonology. Developmental errors will be nonexistent in the beginning, will increase in frequency, and then decrease. Figure 4 below illustrates the model.



Figure 4: The Ontogeny Model (Major 1987)

Major tested the Ontogeny Model with Brazilian learners of English and their acquisition of English /r/, final consonant clusters, and voiced and unvoiced word-final obstruents, and tentatively concluded that the learners followed the sequence outlined in his model. This model has been criticized for being vague in some aspects (James 1988), but Major does make one strong claim: that there is no fundamental difference between first and second language acquisition with regards to the progress of developmental errors. The difference between adults and children, Major claims, is the initial state. While children start with no established language system, adults start with an intact language system, and it is this native language system that interferes with the development of the second language phonology. A revision of the model proposed by Major (2001) is the Ontogeny Phylogeny Model, which adds the role of the L2 as a factor to the other two factors of L1 and developmental influences, and makes a clearer statement that "development" indicates universal patterns of language acquisition.

# 2.2. Speech Perception

### 2.2.1. Language and general cognition: Voice Onset Time

A common topic for second-language speech research is the acquisition of voice onset time (VOT). Adults can discriminate between voiced and voiceless segments in their native language in a categorical manner, as demonstrated by Lisker and Abramson (1964). Presented with a collection of synthetic speech segments whose only difference was a gradual variation in the voice onset time, participants could identify each segment as voiced or voiceless with great consistency and uniformity of responses. The ability to discriminate differences in it is evident in very young infants. Eimas et al (1971) tested infants as young as 1 month old, and found that they could distinguish VOT differences in a categorical manner. The ability to detect differences in VOT seems to be innate.

The language-specific values for VOT must be learned, however. Lasky et al (1975) examined the ability of infants to detect cross-language VOT values. Infants from Spanish and English-speaking environments were tested. The two languages have different VOT values to distinguish voiced and voiceless onsets. English onset VOT is more consistent with other languages. Voiced onsets have a typical VOT of 20 milliseconds, and voiceless onsets have an average VOT of 40 milliseconds. Spanish, on the other hand, has a more marked system of dividing the VOT space. Voiced onsets in Spanish have a pre-voiced VOT of –20 milliseconds, and voiceless onsets have a VOT of 20 milliseconds.

The infants listened to a series of sounds that varied from [ba] to [pa] in gradient steps. Using a head-turn method, the infants indicated when they heard a change in the category from /b/ to /p/. The results from testing the responses of infants showed that infants 4-6 months old from both language backgrounds responded in a uniform manner. Even infants in a Spanish environment indicated the change at the VOT values of 20 and 40 milliseconds, even though in the language of their environment, VOT times ranged from –20 to 20 milliseconds. The result of this study indicates that there are innate phonetic values that may provide a basis from which language-specific values can be built, and at the age of 4-6 months, the language-specific values have not yet been set.

The language-specific value of VOT is one aspect of the second language that must be learned. Gass (1984) examined second language learners' perceptions of VOT by using a forced-judgment task with synthesized stimuli. Participants indicated whether the sound that they heard began with a voiced or a voiceless bilabial segment. Compared with native speakers of English, whose judgments typically show a clear categorical distinction consistent with studies done with native speakers (Lisker and Abramson 1964, Liberman 1970), the nonnative speakers showed a fuzzy, continuous distinction, with no clear categorical boundaries between segments. Gass suggests that the participants are influenced by the categorical boundaries in their native language. Their judgments seem to reflect a system in flux. The participants could have been in the process of developing an interlanguage system, and so their responses showed that they had not yet developed clear distinctions of the VOT boundaries. Another factor that may have influenced the participants is their relative short time in an English-speaking environment (half of the participants had been in the US between 2 days and 1 month). Work has also been done in measuring the ability to detect VOT differences in non-human animals. Kuhl and Miller (1975) found that chinchillas could detect VOT in a manner almost identical to that of humans. Other evidence that animals can perceive speech with patterns similar to humans was found in studies by Kluender et al (1987) and Kluender and Lotto (1994). These studies suggest that the speech perception mechanisms may not be modular and specific to language, but are part of the general cognitive processing capabilities of the nervous system.

# 2.2.2. Language-specific: Perceptual magnets

Kuhl's (1991) seminal study introduced what she termed "perceptual magnets" for speech perception. Kuhl proposed a model of native language perception that differentiates categorical perception from the linguistic process of phonemic perception. Although categorical perception has been shown in nonlinguistic modules of perception in humans and in nonhuman animals, Kuhl attempted to determine whether phonemic perception was only found in humans. While not refuting categorical perception as a real phenomenon, Kuhl's study was an attempt to isolate the linguistic component of vowel perception.

For her study, Kuhl chose two tokens of the vowel /i/ from Peterson and Barney's (1952) study. She termed one of the tokens from the Peterson and Barney study a "prototype," because it was close to the average value for /i/ in Peterson and Barney's study. The other token vowel from the Peterson and Barney study was termed a "non-prototype" of the vowel /i/, because although a token of /i/ with its values was found in the study, it was an outlier token in the study. Kuhl then synthesized two series of 32

vowels for each token, each of which differed from the token in measured distances based on values of the first and second formants. The relative values of the stimuli are shown in Figure 5 below. The "P" and "NP" indicate the relative positions of the prototype and nonprototype sounds, and the variant sounds, as they would appear on an F2-F1 X-Y grid. As mentioned in section 1.4.2, the F1 and F2 formant values are indicators of vowel height and frontness, respectively.



Figure 5: Visual representation of tokens from Kuhl's 1991 study. Filled areas on the matrix indicate the F1-F2 values of vowel sounds. The areas labeled "P" and "NP" are the F1-F2 values of the token selections from Peterson and Barney (1952). The label "P" indicates the prototype sound that was closer to the sound produced by more subjects in Peterson and Barney's study. The sound labeled "NP" is the nonprototype sound, a sound that was produced by one subject, but deviates from the average of all subjects In the study to a greater degree than the "P" sound. The filled areas surrounding the prototypes are the F1-F2 values of vowel sounds that were synthesized for this study. The F1-F2 values of these sounds differ incrementally from the sample sounds from Peterson and Barney.

One group of adult participants listened to the group of sounds that were based on the prototype vowel. Using a goodness rating task, participants evaluated each sound for the goodness to fit for the vowel /i/ on a scale of 1 to 7. Kuhl found that participants showed a clear preference for the prototype over the other sounds that were variants of the prototype. This indicated that participants agreed on a "best fit" paradigm of the vowel /i/. In the task, the more a sound stimulus varied from the prototype sound, the worse it was ranked by participants.

A second group of adult participants performed the same judgment task, but the sounds that they listened to were based on the nonprototype sound. Although the nonprototype sound was an instantiation of a sound that was actually uttered by a subject in the Peterson and Barney study, participants showed a stronger preference for sounds that were closer in value to the prototype sound, the sound that was closer to the average value of utterances from the Peterson and Barney study. This finding supports the suggestion of the first task, that participants had an intuition about the best fit for /i/. Even though the prototype sound was not available to the second group of participants, they nevertheless showed preference for those sounds that were closer to the prototype sound in the first task.

Kuhl's conclusion from the study was that adults have established a clear phonological category for the vowel /i/. This vowel category was part of the participants' linguistic system. Furthermore, they could call on their linguistic competence to evaluate other sounds and indicate how similar those sounds were to their linguistic category for the vowel sound. Kuhl also noted that the response patterns of all the subjects were consistent within the group. This suggested to Kuhl that the "internal standard for the vowel /i/" is quite similar among speakers of the same speech community.

Another task in Kuhl's study used the same set of stimuli. This task measured whether participants could detect differences between sounds. A tenet of categorical perception is that within categories, differences between tokens are difficult to detect, while across category boundaries, differences are easier to detect. Participants listened to a pair of sounds, and indicated whether the sounds were the same or different. It was in this task that the linguistic factor seemed to influence participants' judgments. Participants had more difficulty detecting differences between sounds that were both close to the prototype. They had less difficulty detecting differences between sounds that were farther away from the prototype. Kuhl dubbed this the "perceptual magnet effect." Kuhl found that participants tended to identify sounds that were close to the prototype as

being the same, but were able to distinguish between sounds that were not close to the prototype. Kuhl proposed that the prototype acted as a magnet, drawing to it sounds that were close to it in the acoustic space, and consequently making them harder to distinguish.

Human infants aged 6-7 months and monkeys were tested on the same stimuli, to see if they could detect the gradual differences in the sounds. Participants listened to two sounds, and indicated whether the sounds were the same or different. The responses of human infants were gathered using a head-turn technique (Kuhl 1981). Monkeys were trained to press a button for a food reward when they heard a different sound. The results for human infants mirrored those for adults. Kuhl concluded from this that human infants also have a mental prototype for the vowel /i/. Rhesus monkeys could detect differences between stimuli, but their response patterns were quite different from those of human participants. Kuhl concluded that since humans and monkeys used different bases to judge the phonetic material, that only humans showed evidence for mental prototypes for linguistic segments.

Kuhl's finding has important theoretical implications for speech perception. A nagging problem with theories of speech perception is that non-human animals, which are not supposed to be capable of linguistic behavior as humans are, exhibit the categorical perception effects that supposedly indicated linguistic behavior. The concept of a perceptual magnet captures the distinction between discrimination between categorical perception, which may be a general cognitive function, from phonemic identification, which should be linguistic and restricted to humans. There is no doubt that

categorical perception underlies the perceptual magnet effect, but Kuhl's study indicates that there is more to perceptual magnets than categorical perception.

Kuhl concluded from this study that phonetic categories are unique to humans, since they were not evident in the response patterns of monkeys. Furthermore, phonetic categories influence categorical perception in humans. The presence of phonetic categories in infants could be due to an innate endowment, but as Kuhl points out, sixmonth-old infants have already been exposed to a considerable amount of language, and it is possible that categories have been formed by that age. The effect of language experience on category formation is discussed in section 2.2.3 below.

Kuhl and Iverson (1995) suggest that the acoustic space is divided up into "natural magnets," which they indicate can be seen as innate phonemic categories. During language acquisition, as infants gain more exposure to the language of their environment, the acoustic space is re-divided, and new phonemes ("magnets") are formed. Some magnets may even disappear. The well-documented difficulty that Japanese speakers have in discriminating the English /r/ and /l/ distinction (e.g., Sheldon and Strange 1982, Yamada, Tohkura, and Kobayashi 1997, Gordon, Keyes and Yung 2001) could be explained as the development of a perceptual magnet in Japanese speakers that blurs the phonetic distinction between the two sounds.

A study by Iverson et al (2003) explored the influence of first language on speech perception. Their hypothesis was that participants would be less able to detect contrasts that do not exist in their L1 than contrasts that are found in their L1. The target of their investigation was the distinction of /r/ and /l/ in English, and its acquisition by native speakers of Japanese. They employed as stimuli 18 synthesized tokens of a CV syllable.

The stimuli were synthesized to vary the F2 and F3 values in equal steps, forming a series of sounds that varied between /ra/ and /la/. Participants, native speakers of English, German and Japanese, listened to the stimuli in pairs, and indicated whether the sounds were the same or different.

Results showed that Japanese participants are less sensitive to the distinction between /ra/ and /la/. A particular difference between native speakers and Japanese speakers was sensitivity to differences in the F3 formant. English-speaking participants, but not Japanese speakers, indicated that stimuli with different F3 values were different sounds. Iverson et al concluded that native and nonnative speakers attended to different aspects of the sound. Changes in the F2 formant seemed to be the determining factor for Japanese participants to make their choice in labeling the sound, in contrast to the F3, which native speakers seemed to use. They explain their results by suggesting that the participants' Japanese phonological systems were "mis-tuned" to the English contrast. They had developed perceptual categories for Japanese, and used these categories to interpret the English sounds. Because their systems did not attend to the indicative factor, F3 values, they did not accurately perceive the nonnative contrast.

## 2.2.3. The influence of language experience

The generalization of some aspects of speech perception to general auditory processing strategies has forced a refinement of theories of speech perception. Kuhl's research suggests that phonetic categories (Kuhl's "prototypes") are formed early in life. Her study suggested that categories were formed at least by 6 months of age. The formation of language-specific prototypes was the focus of Werker and Tees (1984).

They studied infants' abilities to distinguish native and nonnative phonetic contrasts. The difference between [t] and the retroflex [t] that is found in Hindi could be detected by infants younger than 10 months regardless of whether the distinction is present in the language of the environment. Infants aged 10-12 months of age were less able than 6 month-old infants from the same language environment to distinguish between [t] and the retroflex [t], while 10-12 month old infants from Hindi-speaking environments could detect the difference. Werker and Tees conclude that the loss in ability to detect nonnative contrasts is a result of language experience.

Werker and Lalonde (1988) replicated the approach used by Werker and Tees (1984). They asked whether adult speakers of English and Hindi would categorize the same continuum of sound differently. Their stimuli consisted of eight synthesized sounds that varied the onset formants to form a /ba/ to /da/ continuum. Using an ABX matching task, Werker and Lalonde were able to determine where participants' categorical boundaries fell. The English-speaking participants divided the continuum into two categories, labial (/ba/) and alveolar (/da/), but the Hindi speakers divided the continuum into three categories: labial (/ba/), dental (/da/) and retroflex (/da/). They then tested two groups of infants using the same stimuli. Infants aged 6-8 months, regardless of the language of their environments, could discriminate the contrasts that Hindi speaking adults indicated. Infants ages 11-13 months from an English-speaking environment could not detect the Hindi contrasts. Werker and Lalonde conclude that language experience in the first year of life causes a developmental change the way that spoken language is perceived. Infants lose sensitivity to some contrasts as they develop perceptual categories.

Kuhl et al (1992) tested the ability of 6-month old infants in English and Swedish environments to detect native and nonnative vowels. A background assumption of this study was that by the age of 6 months, perceptual categories have already been formed, and these categories will affect the perception of speech. Kuhl's perceptual magnet effect predicts that the categories that are formed as a result of linguistic experience serve as prototypes. The effect of the perceptual magnet effect is a reduction in the ability to discriminate between small variations from the prototype. The vowels that are near the prototypes will not be distinguished as different from the prototype. There were two assumptions of this study. One was that the prototypes for /i/ are different for English and Swedish speakers (a comparison of American and Swedish adults' prototypes of /i/ was measured in Willerman and Kuhl 1996, reflecting a measurable difference). The other assumption was that the Swedish participants had in their vowel systems the high front vowel /y/, and the English speakers did not.

Kuhl et al chose good examples ("prototypes") of English /i/ and Swedish /y/, and synthesized variants of them by altering the F1 and F2 values. The F1 and F2 values of the stimuli varied in measured steps from the prototypes. Participants were American and Swedish infants. Both groups of participants listened to the stimuli based on both languages' prototypes. The prediction was that the /i/ prototype would be a magnet for the American but not the Swedish infants, and the /y/ prototype would be a magnet for the Swedish but not the American infants.

Using a head-turn technique (Kuhl 1981), infants listened to stimuli in 2 second intervals, and, during a learning phase, learned that when the stimulus changed, a toy bear pounded a drum. Infants turned their heads to anticipate the performance of the toy bear.

The American infants did not show signs of detecting variation from the English /i/ prototype to the degree that the Swedish infants did. This finding was predicted by the perceptual magnet theory. A prototype sound acts as a magnet, reducing discriminability of similar sounds. If Swedish infants did not have the same prototype, then they would not show reduced ability to detect differences. Similarly, Swedish infants could not detect variation from the /y/ prototype to the degree that the American infants could. The Swedish infants showed signs that they had developed a prototype for /y/, and thus displayed perceptual magnet effects. The results from this study give evidence that by age 6 months, infants had been affected by the language of their environments, and they had developed linguistic categories for those vowels.

An implication of Kuhl's (1981) study is that language-specific categories seem to be formed earlier for vowels than for consonants. Polka and Werker (1994) had similar findings. They examined the development of perception of L1 vowels and the accompanying loss of ability to detect nonnative vowel contrasts in Canadian infants. Variants of the German high front vowel /y/ were presented in a CVC syllable /dyt/. Variants of the high back rounded vowel /u/, found both in Canadian English and German, were also prepared in the same CVC environment (/dut/). Two groups of infants, aged 6-8 months and 10-12 months, were presented with the stimuli and tested for discrimination using the head-turn method described in Kuhl (1981). Comparing the results of this vowel-discrimination study with those of infant consonant-discrimination studies, the authors found that the 6-8 month old infants responded with more languagespecific responses than was found in that age group in discrimination tasks involving consonants. Running the same experiment with 4-month old infants showed more

language-neutral results, the kind shown in 6-8 month old infants for consonants. The results of this study suggest that the shift in discrimination ability from language-neutral to language-specific happens at an earlier age for vowels than for consonants. Later research by Bosch and Sebastian-Gallés (1997) with 4-month old infants supports the finding that at that age, infants are already able to discriminate between their native language and other, even closely related, languages.

In summary, research on the effect of language experience on speech perception shows that from an early age, the native language changes the way that speech is perceived. The process of language acquisition seems to include the development of perceptual categories. These categories affect the way that the native language is perceived, and also reduces sensitivity to nonnative contrasts. This process of specialization to a language system results in an internal language system that is specialized for one language. This system may assist members of the speech community to disregard the individual differences of speakers, which will aid speech perception. However, specialization to the acoustic properties of one language may also impede the acquisition of a second language. The process of second language acquisition may require the learner to alter the existing categories, and may require the development of new categories. Measuring whether a learner of a second language has categories similar to those of native speakers is one way to measure second language acquisition.

The research also illustrates methods by which speech perception can be tested. In many studies of speech perception, the use of synthesized speech is a way to control the stimuli that is presented to participants. A goodness of fit task in which adult participants

compare the stimuli with their internal systems is a common method to investigate phonological competence.

#### 2.3. Perception and production of second-language vowels

### 2.3.0. Introduction

The relationship between vowel perception and production is central to this study. The relationship is analogous to the competence-performance relationship in linguistic theory. Part of an individual's linguistic competence is the vowel categories in the individual's phonological system. As the discussion of the research in section 2.2 above suggests, the native language of the perceiver influences how vowels are perceived in speech. Part of the task of the adult language learner is to adjust the perceptual categories of the native language to accommodate the categories of the second language. In other words, the language learner has to change or develop a new phonological competence.

A salient part of an individual's linguistic performance is speech production. But just like other instances of linguistic performance, factors other than competence can influence speech performance. This is especially evident in second-language speech. Foreign accent is an indicator that the learner has not mastered the phonology of the second language. However, just as definitive conclusions about linguistic competence should not be made based on performance, so also must we exert caution about reaching conclusions about second-language phonology based on second-language speech. This section discusses research that explores the relationship between vowel perception and production. First, I briefly review research into the nature of vowel perception in general.

Following that is a discussion of research in second-language vowel perception. Finally, I review studies of the relationship of second-language vowel perception and production.

### 2.3.1 Perception of first and second-language vowels in adults

#### 2.3.1.1. First language vowel perception

Ladefoged and Broadbent (1957) synthesized vowels in a /bVt/ context, and preceded the word with the phrase "Please say what this word is..." Participants were asked to identify the vowel in the last word, choosing between *bit*, *bat*, *bet* or *but*. In tokens in which the first formant of the "carrier phrase" that preceded the /bVt/ word was shifted up or down, but the /bVt/ word was the same, participants identified the vowel in the /bVt/ word differently. The quality of the vowels in the preceding phrase seemed to influence the perception of the following word. Their study suggests that perception of vowels is greatly dependent on context. Ladefoged and Broadbent suggest that the preceding acoustic signal, which they term a "carrier signal," gives a crucial reference of vowel quality to the listener. This study illustrates the phenomenon of speech normalization, as discussed in section 1.4.1.3 above. Listeners need to adjust their perception based on the qualities of the speaker's voice. The carrier phrase in this study influenced listeners to normalize the speech stream up or down, and the /bVt/ word was normalized with the carrier phrase.

Fry et al (1962) explored participants' ability to recognize phoneme boundaries and intra-category discrimination of vowels. They produced 13 synthetic vowels that varied on a continuum that covered the vowels /1/, / $\epsilon$ / and / $\alpha$ /. Participants were presented with the stimuli using a "forced-choice ABX" method. Two sounds were

presented ("A" then "B"), followed by a third sound ("X"). Participants were asked whether the third sound ("X") was the same as the first ("A") or the second ("B"). The stimuli were presented in isolation (different from Ladefoged and Broadbent (1957), which included a carrier signal) with no context. There was significant overlap in participants' responses. Unlike studies of categorical perception, in which there is considerable agreement among participants for consonant boundaries, no such general agreement was found for the vowel boundaries. The results of this study indicated that boundaries between vowels not clear-cut. Fry et al concluded that the perception of vowels is continuous rather than categorical. In the discrimination task, there was no evidence of categorical perception effects.

The study by Fry et al complements Ladefoged and Broadbent (1957) in showing that vowels are relative in nature. When presented in a context, in relation to other vowels in a phrase, listeners are able to make consistent judgments about vowels, as in the Ladefoged and Broadbent study. However, if there is no point of reference, as in the Fry et al study, listeners do not make categorical judgments about the vowel sound. In explaining their results, Fry et al point to Joos' (1948:68) "template" model of speech perception that suggests the listener mentally maps the speech signal into vowels within the acoustic space (see the discussion of Joos' model in the discussion of normalization in section 1.4.1.3 above). In the absence of context, listeners are unable to normalize a vowel sound.

Scholes (1967a) asked whether participants would associate synthetic vowels with phonemes in their native languages. Participants indicated whether each of 69 synthesized vowels was a representation of a sound in their language. The vowel sounds

were synthesized from a continuum of F1 and F2 values over a normal frequency range for human speech. Participants chose from a list of words in their native language as a match for each of the synthesized vowels. The stimuli were presented in a scrambled order. With almost no exception, participants were able to match each synthesized vowel with a vowel sound in their language. The methodology allowed the mapping of boundaries for each vowel for each participant. There was significant agreement among speakers of the various languages as to the category boundaries for the vowels. There was some overlap, however. Some stimuli were identified as different vowel phonemes by some speakers. This study showed that there is some speaker-specific variation in the perception of vowels, similar to that found by Fry et al (1962). While there was not complete agreement among all participants with regard to vowel category boundaries, the overall level of agreement among participants indicates that the identification technique reflects the homogeneity of a language community's vowel phoneme categories.

In a followup study, Scholes (1968) had nonnative English speaker participants give words in their native language that represented the vowel sound that they heard in the synthesized sounds. The synthesized vowel sounds provided an objective environment in which to compare vowel phoneme patterns of various languages. Participants from the same language background gave words with largely similar phonemes in their native languages. Scholes claimed that this task supported the results of the 1967 study, that participants can associate native-language vowel phonemes with the synthetic stimuli. Participants' second task was to listen to the vowel sounds again, and give an English word that used the vowel sound that they heard.

Scholes then mapped the F1 and F2 values of the words collected from the two tasks onto an X-Y grid. The areas covered by the vowel sounds of the two groups of words overlapped significantly. Scholes claimed that this showed similar categorization between vowels in their native language and in English. He concluded that the non-native speakers of English seemed to perceive English according to their native language phonemes, and that their native language categories influenced their performance in a second-language task. Scholes found that the categorization of synthetic vowels by nonnative speakers is the same for the native language and nonnative language, as long as the L1 and the L2 have counterparts. Nonnative speakers will hear the nonnative language through their native language system.

Scholes' findings are consistent with those of the studies in section 2.2.3 that show evidence for a specialization of the perceptual system for the phonetic categories of a particular language. Evidence for the influence of the first language on the second can be seen through performances on tasks that make use of perceptual categories. Participants in Scholes' studies tended to identify second language vowels in terms of vowel categories in their first languages.

### 2.3.1.2. The interrelationship of perception and production of L2 vowels

Flege's (1995) Speech Learning Model represents an attempt to account for the differential performance in perception and production in second language phonology. Specifically, it addresses a phenomenon in second language phonology that on the surface may seem puzzling: that sounds in the second language that are quite different from sounds in the native language are acquired with more accuracy than sounds that

have close correlates in the native language. Contrastive Analysis (Lado 1957:59) predicts that areas of the native and target languages that are similar would be acquired with greater ease than areas in which the two languages are not similar. Stockwell and Bowen (1965:9-18) present a systematic method of predicting difficulty based on a comparison of the phonological systems of two languages. Following that claim, then, we would predict that phonological segments in the target language that correlate to those in the native language should then be easier for learners to acquire. However, empirical evidence shows just the opposite.

Flege and Hillenbrand (1984) compared production of the English onset /t/ by French-speaking learners of English. Acoustical comparison of the onset in the two languages showed that VOT values for the segment /t/ differ phonetically in French and English. Native speakers of French, when speaking English, tend to use French VOT values to produce an English /t/. The interpretation of this phenomenon using the Speech Learning Model is that the French learners of English perceive the English /t/ as similar enough to the French /t/, so no adjustment of the phoneme is necessary.

Flege (1987) had measured the performance of native speakers of English in producing the French segments /t/, /u/ and /y/. The first two segments Flege classified "similar," because although not phonetically the same, English and French both have those segments. The segment /y/ is a "new" sound for Anglophone learners of French. Flege found that their production of this new sound was more nativelike than their production of the French categories that had similar counterparts in English.

What is interesting in these examples is that the Contrastive Analysis Hypothesis would have predicted that the new sound /y/ would have given the American learners of

French more difficulty than the segments that were similar to segments in the learners' native language. The similar segments should have been easier to learn, and so the participants' performance should have been better. The answer that Flege gives to the contradiction to the Contrastive Analysis Hypothesis is that learners will establish new categories for the segments in the second language that are perceived by the learner as new, and for sounds that are classified as similar to categories in the native language, the learner will simply substitute those categories.

This model of second-language speech is an interesting corollary to Eckman's (1977) Markedness Differential Hypothesis, which was a refinement of Lado's (1957) Contrastive Analysis Hypothesis. Lado predicted that in comparing the L1 and L2, those elements that are similar to the learner's native language will be easy to learn, and those elements that are different will be difficult. Flege's SLM predicts that those categories in the L2 that are similar to the L1 will be more accented, and those elements that are new in the L2 have the potential to be acquired with native-like accuracy.

Bohn and Flege (1997) tested this hypothesis in a study involving the acquisition by native speakers of German of the English vowel  $/\alpha$ . German has the vowels /a/, /e/, and  $/\epsilon/$ , but not  $/\alpha/$ . The authors asked whether adult learners of a second language could acquire a new category, and how their perception and production would compare.

Two groups of nonnative speakers were compared to a group of native speakers. One group, which they labeled "experienced" learners, has a mean length of stay of 7.5 years. The other group of nonnative speakers, the "inexperienced," had a mean length of stay of 0.5 years. For the production task, the two groups of German learners of English read short sentences in English, ending in words that contained the  $/\epsilon/$  or /æ/ sound.

Acoustic analysis of the production showed that the speech of the more experienced group of nonnative speakers was more similar to the native speakers, and the speech of the group of less experienced participants was less native-like. There was little overlap between the two vowels in the production of the native speakers, but no distinction between the two vowels by the inexperienced nonnative speakers. Bohn and Flege conclude from this that the inexperienced speakers do not distinguish a distinct category for the  $/\alpha$  sound. They suggest that length of exposure to the second language could have a greater effect on the production than on the perception of the second language.

For the perception test, the authors synthesized a gradient of 33 vowels that varied between  $|\varepsilon|$  and  $|\varpi|$  on a continuum, varying F1, F2 and F3, plus the duration of the segments. Using a forced-response method, participants indicated whether each sound corresponded to the word *bet* or *bat*. Items were presented in random order. Native speakers responded with a categorical pattern, indicating a clear separation between  $|\varepsilon|$ and  $|\varpi|$  in their vowel systems. Overall, the nonnative group that had had more exposure to English responded in a more native-like manner than the nonnative group with relatively less exposure on the production test. The responses of the experienced group showed a more distinct differentiation between the two vowels than those of the less experienced group. However, the pattern of perception by both groups of nonnative speakers differed from that of the native speakers in that it reflected a more continuous pattern, with no clear separation between the segments, than was found in the native speakers.

A closer analysis of the data by the authors indicated that nonnative speakers referred to different acoustic cues than native speakers. The pattern of responses among

native speakers showed that this groups detected a category change along the continuum of F1, indicating vowel height. Vowel height was the salient factor that separated vowels for native speakers. The pattern of responses for nonnative speakers did not show such a trend. There was no clear categorical distinction among nonnative speakers that correlated with F1. Instead, the response pattern of nonnative speakers indicated that they associated vowel duration with category change. For native speakers of English, duration of the segment was not a factor in making a judgment, but "inexperienced" German learners of English seemed to rely on duration as the prime factor in making their judgments. This finding indicates that the participants had developed categories for the second language vowels that use different criteria for distinguishing the vowels than those that are used by native speakers.

An interesting aspect of this study is the apparent disconnect between perception and production. The experienced nonnative group produced the vowels to a more nativelike degree of accuracy than the inexperienced group, yet their performance on the perception task showed differences suggestive of how the vowel is represented in the participants' grammars. This result was similar to that of the VOT study by Gass (1984), which indicated that perception and production could be disconnected among nonnative speakers.

The indication that the nonnative speakers in Bohn and Flege's study relied on a different acoustic aspect of the vowel than native speakers for identifying vowels parallels the findings of the study of Japanese learners of English by Iverson et al (2003), in which relying on changes in F3 instead of F2 led nonnative speakers to inaccurate judgments.

Ingram and Park (1996) investigated the perception and production of Australian English vowels by Korean and Japanese learners of English. In the first task, participants listened to recordings of speakers reading /hVd/ words containing the Australian English vowels /i/, /1/, /e/, / $\alpha$ / and / $\alpha$ /. Participants performed a forced-choice identification task to identify the vowels. Japanese participants could correctly identify all of the vowels with a high degree of accuracy (92% - 100%), but Korean participants responded with a much more mixed pattern of judgment for the vowels /e/ and / $\alpha$ /, with accuracy measures between 46% and 54% for those vowels. Korean participants with more exposure to English did better on the task of identifying /e/, but misidentified several tokens of / $\alpha$ / as /e/, and vice versa.

The authors surmised that the two groups of participants were using different strategies for arriving at their judgments. Their hypothesis was that Japanese participants were using duration as a factor in perception, possibly because length is contrastive in Japanese, but not in Korean. The influence of the L1, the authors suggest, helped the Japanese arrive at the correct judgment. Since long and short vowels are not distinctive in Korean, the authors continue, length did not influence Korean participants' judgments. Korean participants relied on other cues to distinguish vowels. The study did not include a discussion of the formant values of the vowels, focusing instead on vowel duration as a factor in perception and production.

The second task was reading aloud the words that were in the first task. Participants' voices were recorded and analyzed for vowel duration. The results showed a clear distinction between the two groups. Japanese participants produced vowels with internally-consistent values, in lengths that differed across segments. The authors assume

that the participants were transferring the moraic vowel length from Japanese to the Australian English target vowels. The Korean participants, on the other hand, produced vowels in a pattern of lengths that was closer to that of the target language.

The third task was a "native rating" of the L2 vowels. Participants listened to tokens spoken by two native speakers of Australian English, and were told to write the vowel in their native language that best represented the vowel sound that they heard. Additionally, they were told to indicate whether the vowel that they heard was long or short. The Japanese participants categorized the vowels consistently for each of the two speakers. Their responses did not indicate sensitivity to inter-speaker differences. The Korean participants' judgments seemed to have been more speaker-dependent, however. They rated the same vowel from the two speakers differently. While Japanese participants responded to the stimuli in a categorical way, Korean participants responded to absolute differences in token vowel duration.

The authors assumed that the Korean participants were using phonetic cues to make their judgments. Japanese participants, however, seemed to have normalized the speakers' vowels for length, compensated for speaker differences, and responded on the basis of phonological cues. Since vowel length is distinctive in Japanese, but not Korean, the authors conclude that the influence of the participants' native languages was at work in their perception of the second language. Japanese subjects used processing strategies for speaker normalization that are operational in their native language. The responses were very consistent among that group, a pattern that would be expected among members of the same speech community using their native language competence.

## 2.4. Acoustic Comparison of Korean and English vowels: Yang 1996

Yang (1996) compared the production of English and Korean vowels, with the purpose to compare the acoustic properties of the vowel phonemes of the two languages. The data for the study was samples of vowel utterances from native speakers of English and Korean who read word lists containing the vowel phonemes of the languages. Participants' voices were recorded for later acoustic analysis. The 20 English speaker participants were from the South or Southwest of the United States. The 20 Korean participants were from Seoul and spoke standard Korean.

Table 1 and Table 2 show the average F1 and F2 values for the Korean vowels /i y i e  $\emptyset \varepsilon$  a ^ o u/ and English vowels /i I e  $\varepsilon x$  a  $\vartheta \vartheta \vartheta \upsilon$  u/ from the study as produced by the participants from the two language backgrounds.

Korean	F1	F2
i	343	2517
У	356	2409
i	426	1596
e	570	2173
ø	531	2006
3	634	2067
a	862	1583
^	687	1246
0	476	987
u	396	1001

Table 1: The average values of F1 and F2 for Korean vowels as produced by the 20 native speaker participants in Yang (1996).

English	F1	F2
Ι	338	2572
Ι	438	2193
e	495	2309
3	581	2072
æ	756	1901
а	710	1169
ə	647	1486
Э	720	1083
0	513	1167
υ	469	1409
u	375	1452

Table 2: The average values of F1 and F2 for American English vowels as produced by the 20 native speaker participants in Yang (1996).

Figure 6 shows the same values plotted on an X-Y grid, showing their respective positions within the vowel space.





For purposes of comparison, Y ang assumed that the two languages had vowels in common. In the literature, vowels that are written with the same phonetic symbols across languages are often acoustically similar, and should cluster when graphed according to their F1 and F2 values. Yang compared the vowels that are written /i e  $\varepsilon$  a o u/. We can also include in the comparison the Korean vowel /^/ and the English /ə/, because the two vowels have similar acoustic properties.

Figure 7 below shows where the values of each of these vowels lie within the vowel space. The Korean vowel  $/^/$  is labeled  $/^/$  for purposes of comparison.



Figure 7: The respective positions of the F1 and F2 values of the common vowels between English and Korean, /i e  $\varepsilon$  a  $\Rightarrow$  o u/, plotted on the same graph

For many of the vowels, the relative positions of the corresponding vowels are quite similar. The vowel /i/ for example is very similar between languages. Larger differences are found in the vowels /a/, /ə/ and /u/.

An interesting aspect of the differences between the vowel systems is the relative values for |e| and  $|\varepsilon|$ . The Korean vowels |e| and  $|\varepsilon|$  are lower and more back than their

English counterparts. There is greater similarity between the Korean /e/ and the English  $|\epsilon|$  than there is between the English and Korean /e/ or  $|\epsilon|$ .

The average F1 and F2 values of the common vowels is given in the Table 3 below.

	F1		F2	
Vowel	English	Korean	English	Korean
i	338	343	2572	2517
e	495	570	2193	2173
ε	581	634	2072	2067
a	710	862	1169	1583
ə	647	687	1486	1246
0	513	476	1167	987
u	375	396	1452	1001

Table 3: average F1 and F2 values for the vowels that are common for Korean and English in Yang (1996).

There are two aspects of Yang's study that are relevant to this study. The first is that the vowels that are indicated as corresponding vowels do indeed seem generally to correspond. When acoustic measurements of tokens of the vowels that use the same notation in the literature are mapped on a grid using the F1 and F2 values as x and y coordinates, the respective positions of the vowels in the acoustic are similar across the two languages.

The second significant aspect of this study is that it shows that while the vowel systems of English and Korean have some similarities, they are still different. Through an examination of the formant values of corresponding vowels across languages, Yang found statistically significant differences between the two language systems. A series of t-tests revealed cross-language differences in nearly every pair.

## 3.0 Introduction

As mentioned in Chapter 1, this study is an investigation into the relationship between perception and production of American English vowels by native and nonnative speakers. This study investigates perception in a way that differs from other perception studies. This chapter describes the participants, research instrument, and data collection procedure.

## 3.1. Participants

## 3.1.1. Native Speakers

Ninety-eight male and female native speakers of English participated in the study. They were recruited from an undergraduate sociolinguistics course at Michigan State University, and received partial course credit in return for their participation. The majority of the participants ranged in age from 18 to 22 years old. Of those participants, six were not from Michigan, and the data from five others was not collectable because of technical difficulties. The data from eighty-seven participants was usable, and is reported in this study.

Age	Number of	
	participants	
18	1	
19	15	
20	40	
21	22	
22	7	
over 22	2	
Total	87	

Table 4: Age of native speaker participants

3.1.2. Nonnative Speakers:

Twenty-seven Korean nonnative speakers of English participated in the study. They were recruited from English as a Second Language classes that were offered by Michigan State University for members of the community. The participants ranged in age from 18 to 42 years old. One participant did not report an age. Their length of residence in the US ranged from a few months to 12 years, with the majority less than one year. One participant did not report the length of time studying English or residing in the U.S.
Age	Years studied English	Months in the U.S.	Gender
18	6	just	F
18	4	10 months	F
19	10	2 months	F
19	5	2 years	М
19	6	4 months	F
20	4	4 months	F
20	8	12 months	М
21	1	7 months	F
21	9	4 months	F
22	not reported	not reported	N/A
22	10	8 months	F
22	10	9 months	F
22	3	3 months	F
22	3	4 months	F
22	6	24 months	М
23	1	12 months	М
23	5	7 months	F
24	6	3 months	М
24	8	12 months	N/A
25	7	4 months	М
27	3	8 months	F
36	14	12 years	F
36	22	7 months	М
37	10	9 months	М
38	5	12 months	М
40	3	8 months	F
42	6	4 months	М

Table 5: Nonnative speaker participants' age, years studying English and length of stay inthe US.

# 3.2 Methodology

# 3.2.1. Rationale for the design of the instrument

The instrument used in this study was designed to gather data in a non-subjective way. The hypotheses being tested required the availability of a continuum of vowel sounds that varied in controlled, incremental steps. Synthesized speech samples were the most effective way to present participants with such a continuum.

Using a computer-based instrument facilitated objectivity of data collection. The data was stored in a central database, and so administering the instrument on a networked computer system was the most practical option. The instrument that was developed was able to accommodate a large number of participants working independently

#### 3.2.2. Synthesized speech samples

The synthesized speech sounds used in this study consisted of 306 artificial vowels each of 500 milliseconds duration, synthesized with the software system *Praat*, version 4.0 (Boersma and Weenick 1999-2000), and presented in a web-based software program written in Macromedia Flash MX.

The acoustic properties of the speech sounds were based on those used in a study by Frieda et al (1999). The speech sounds approximated a male speaker with a voice pitch that began at 300 Hz, and fell to 130 Hz over the duration of each sound. The audio files were created at a 22,050 Hz sample rate, and compressed using the Speech codec of Flash MX at 22 kHz.

### 3.2.3. The vowel matrix

A 17x18 cell matrix represented the range of F1 and F2 value combinations. F1 values ranged in 50 Hz increments from 200 to 1000 Hz in 17 steps with a constant bandwidth of 50. F2 ranged in 125-Hz increments from 475 Hz to 2600 Hz in 18 steps at a bandwidth interval of 100 Hz. For all samples, F3 was held constant at 3000 Hz, B3=150. The 306 vowels were presented to participants in a matrix as shown below. Clicking on any of the squares in the matrix played the vowel with the F1 and F2 values

59

that corresponded to that square's relative position in the matrix. By clicking on adjacent squares, participants could hear the continuum of vowel sounds that was represented in the matrix.



Figure 8: The vowel matrix

The matrix represents a 2-dimensional continuum of vowel sounds. Vowel height is represented in the vertical axis, and vowel frontness is represented by the horizontal axis. The sounds were arranged so that the F1 and F2 values of sounds were lowest at the origin point. In the section of chapter 1 on measuring vowels (section 1.4.2), I discussed how arranging the F1 and F2 values in this way aligns the values with the positions of vowels in the traditional vowel quadrangle. Cells further to the left had a higher F2 value. Cells further below the origin had a higher F1 value. 3.3. Procedure

## 3.3.1. Introduction:

The data collection took place via the Internet. The only requirements were that the participants' computers have a web browser with the Flash 6 plugin, be connected to the Internet, and have the capability to play and record sound. The advantage of developing the instrument in Flash was that the program was playable on any platform or operating system that supported the Flash player, and so there were no cross-platform differences in the presentation or functionality of the instrument. This lent more flexibility in finding a location to administer the instrument.

Computer laboratories at various locations at Michigan State University were used to collect data. The researcher was present at all data collection sessions to answer questions that participants had about the study or about their tasks, and to assist with computer problems.

All participants performed the same tasks in the same order. All items in each task were presented in the same order.

## 3.3.2. Task 1: Perception

The first task was to identify a specific vowel sound. The software first collected some demographic data, including place of birth, age, and native language. This was to identify the two groups to be compared: natives of Michigan and Korean learners of English. Next, it presented the participants with a brief introduction to the vowel quadrangle, and to computer-generated speech. Some brief tasks in vowel identification were presented in the form of a game. The purpose of this was to assist them in navigating the software and familiarize them with synthesized speech sounds. The few participants that had problems with the software were able to overcome them before beginning the actual task.

This orientation phase of the program gradually introduced the participants to the concept of the vowel space, and how this was modeled in the vowel grid that was the medium of the first task. Just before beginning the task, participants were able to explore the vowel grid at their leisure, familiarizing themselves with the way the vowel quality gradually changed with their movement of the cursor through the gradient of F1 and F2 values presented in the matrix.

62

participant had a unique record in the database, which was tagged with an identification number.

Vowel	Word list
i	eat, he, bead, sleep, peak
Ι	hit, bid, lip, pick, slip
e	hate, late, paid, race, lake
3	pet, head, tell, red, wet
æ	hat, dad, sap, tap, rat
a	rod, hot, dock, sod, nob
ə	but, duck, cut, hush, rub
С	caught, law, draw, paw, pause
0	blow, so, rose, toe, post
U	book, hood, look, push, would
u	who, loose, you, blue, spoon

Table 6: Word lists from task 1.

# 3.3.3. Task 2: Production

The software program prompted the participants to read a short word list, presented in writing, into their computer's microphone. The list of words was *heat*, *hit*, *hate*, *pet*, *hat*, *hot*, *hut*, *pause*, *toe*, *book*, *who*. The two considerations in choosing the words on the list were familiarity to the participants, and the acoustic properties of the vowels produced when speaking them. These words were chosen because they were thought to be more familiar to nonnative speakers. The word list was in an order that circumscribed the vowel quadrangle. This was done in an effort to maximize contrast between the vowel sounds in the words by the participants when they read the list. The word list was presented in the same order to each participant, on the assumption that attempting to maximize dispersion was preferable to counter-balancing. Because certain onset consonants, such as voiced bilabials, produce a formant transition in the following vowel, the word list avoided the use of voiced onsets. This decreased the likelihood that the vowel sound produced by the participants was influenced by its phonetic environment.

The words were presented in an order that followed the outline of the vowel quadrangle. The vowel sound of each word on the list was adjacent to a word whose vowel sound was adjacent to it in the vowel space. This was done with the intention of maximizing contrast between vowels.

The audio was captured using the Flash MX microphone control, and was streamed to the server in real-time using the Flash MX Communication Server. Participants used Telex headsets with a built-in boom microphone. The audio was sampled digitally at 22,050 Hz with a 16-bit sample rate. The audio file for each participant was tagged with the participant's unique identification number, which linked the participants' responses on both tasks to the demographic data.

After the participants clicked a button indicating that they had finished recording, the program informed them that they had completed the task, and thanked them for their participation. Native speaker participants completed both tasks in 10 to 20 minutes. Nonnative speakers finished both tasks in 10 to 45 minutes, with most completing within 25 minutes.

# 3.4 Data

Since participants' responses to the identification task were recorded to a database, they were readily available for recovery and analysis. Each square on the vowel grid was

64

internally tagged with a unique code, which would be used to recover the participants' choices for each vowel.

The software program Praat was used to extract the F1 and F2 values for each vowel from the production data of each participant. The values were taken from the earliest point in each syllable where the vowel became stable, typically between 75 and 125 milliseconds after the beginning of the onset. These values were entered into the database as well. Several native speaker participants did not speak clearly or loudly enough, and so their audio files could not be used. Some other participants chose not to record their voices. Audio data from 46 of the 87 native speaker participants could be used.

### 4.0 Introduction

In this chapter, I first show the results of both the perception and production tasks in graphic and numeric format. I examine the data in terms of within-group homogeneity and perception-production differences, and then look at between-group differences. Next, I show a statistical comparison of within-group and within-task results.

Using custom software that accesses the data from the server and plots it onto a grid, the average F1 and F2 values of participants' responses were plotted in the X-Y grid of Figure 9. Values for F2 make up the X axis values, and F1 values are arranged on the Y axis. By positioning the 0 point of both axes in the upper-right hand corner, instead of the lower-left hand corner, the vowels can be plotted into positions analogous to the traditional vowel diagram. The F2 parameter reflects frontness of the vowel, and the F1 represents height. Fronter vowels fall to the left of the origin point, and higher vowels towards the top.

The data were stored automatically into a database on the server during data collection. Scripts written in the PHP programming language ran on the server to retrieve the data, and to calculate averages. The results were collected by software written in Macromedia Flash, which displayed the data in graphical form.



Figure 9: Graphic representation of the vowel area. The gray shaded area represents F1 and F2 values that were not represented by sounds in the perception task, but that could possibly be produced by participants during the production task.

The statistical procedures used for analysis largely consist of heteroscodastic t-

tests. According to Howell (1995:246-247), the t-test is the best choice for comparing the means of two independent groups. ANOVA was also used to identify the component factors of variation.

4.1 Native Speakers

4.1.1 Task 1: Perception

Figure 10 shows the average judgments of the 87 participants who identified themselves as native speakers from Michigan. The positions of the vowels relative to each other are in a pattern acoustically consistent with the English vowel system, as in Ladefoged (1993:212).



Figure 10: Average of Native Speaker Perception

This positioning of the vowels indicates that in this study, as in the studies by Scholes (1967b, 1968), the participants were able to associate the synthesized sounds with vowel sounds in their own language systems. The data from this task was consistent with other acoustic measurements of English vowels.

The average numbers shown in Figure 10 belie the variation in responses. Figure 11 shows the choices of all native speaker participants for each vowel sound. Trends are clearer for some vowel sounds than for others. The responses for /i/, /a/ and /u/ appear to cluster together more than those for /a/ and /e/, for example. This variation may be a

reflection of natural variation within the speech community. This kind of wide variation was also found in the study by Frieda et al (1999). In that study, participants chose the synthesized vowel sound that best matched a particular vowel sound. There was considerable variation among the responses. It could be claimed that synthesized vowels are not natural-sounding enough to make a clear distinction, or that confusion or fatigue on the part of the participants influenced their responses. However, in light of the fact that this is not the only study that found this kind of variation, it seems more likely that this phenomenon is a reflection of inter-speaker differences.



Figure 11: Native speaker choices for perception task for each vowel

Table 7 shows the means of F1 and F2 for the vowels that were tested for in task 1.

Standard deviations are given below the means, in parentheses.

Vowel	F1	F2
I	346	2320
	(176)	(459)
Ι	529	2123
	(137)	(450)
e	632	2220
	(135)	(480)
3	652	2156
	(172)	(405)
æ	803	2138
	(153)	(422)
a	815	1283
	(163)	(391)
ə	652	1331
	(171)	(390)
Э	779	1291
	(150)	(429)
0	613	909
	(178)	(356)
U	508	1067
	(175)	(346)
u	409	953
	(184)	(365)

Table 7: Means for the native speaker perception task. Standard deviations for each mean are under the mean, in parentheses.

Two pairs of vowels (/e/-/ $\epsilon$ /, /a/-/ $\sigma$ /) had F1 and F2 values that are very close together, within 40 Hertz for the F1 values and within 60 Hertz for the F2 values. When plotted onto the vowel space, these pairs of vowels were very close together. This similarity of position suggests the possibility that the participants did not distinguish between the vowel categories. Two-tailed t-tests were performed to determine if the differences between the means were significant or not. Heteroscedastic t-tests were performed to test two samples with unequal sample sizes. The t-test for the  $|e|/|\epsilon|$ 

distinction yielded t(175) = 0.96, p = 0.34 for F2, and t(171) = -0.89, p = 0.38 for F1. The t-test for the /a/-/ɔ/ distinction yielded t(179) = -0.14, p = 0.89 for F2, and t(179) = 1.56, p = 0.12 for F1. In the t-test to see if there was a difference between the means, the null hypothesis was that there would be no difference. In all four t-tests, the observed value was less than the critical value of t, 1.97, and so we cannot reject the null hypothesis. We cannot assume that there is any significant difference between the F1 or F2 values for the pairs of vowels in question.

To examine the system of native speakers as a whole, an ANOVA test was performed that included a post-hoc pairwise comparison test using the Tukey method. This test revealed the level of variation between vowels in the perception test. Table 8 shows the results of the pairwise comparison test for F1 and F2.

	Ι	e	ε	æ	a	ə	э	0	υ	u
i	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.249
Ι		0.001	0.000	0.000	0.000	0.000	0.000	0.026	0.999	0.000
e			0.999	0.000	0.000	0.999	0.000	0.999	0.000	0.000
ε				0.000	0.000	1.000	0.000	0.870	0.000	0.000
æ					1.000	0.000	0.996	0.000	0.000	0.000
a						0.000	0.922	0.000	0.000	0.000
ə							0.000	0.870	0.000	0.000
С								0.000	0.000	0.000
0									0.001	0.000
υ										0.002
	I	e	3	æ	а	ə	Э	0	υ	u
i	0.050	0.861	0.208	0.101	0.000	0.000	0.000	0.000	0.000	0.000
I		0.891	1.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
e			0.994	0.963	0.000	0.000	0.000	0.000	0.000	0.000
3				1.000	0.000	0.000	0.000	0.000	0.000	0.000
æ					0.000	0.000	0.000	0.000	0.000	0.000
a						0.999	1.000	0.000	0.018	0.000
Э							1.000	0.000	0.001	0.000
Э								0.000	0.011	0.000
0									0.253	1.000

Table 8: The results of a pairwise comparison from an ANOVA test shows the source of variation between pairs of vowels in native speakers' responses to the perception task in the F1 formant (top) and F2 formant (bottom). Only two pairs of vowels ( $/e/-/\epsilon/$  and  $/a/-/\mathfrak{O}/$ ) showed no significant difference between them.

0.735

υ

Vowel pairs that showed no significant difference were those with a p value less than 0.05. In evaluating this data, it is necessary to look at both the results for F1 and F2 for each vowel pair. One formant will only show one dimension of the vowel, and it would be inaccurate to say that vowels are not distinct on the basis of only one formant. For example, the F1 values of /I/ and /U/ are very close, showing a p value of 0.999.

However, we cannot conclude from this that the vowels are not distinct. In fact, we would expect the F1 value for those two vowels to be very close, because they are both high vowels, and the F1 formant indicates vowel height. The difference between these two vowels is revealed by the F2 formant. The comparison of the F2 values for those vowels shows a p value of 0.000. Again, this is to be expected, since /1/ is a front vowel, and / $\upsilon$ / is a back vowel, and F2 indicates vowel frontness. For purposes of this comparison, then we assume that vowels are not distinct only if the p value for both F1 and F2 are over 0.05. Under that assumption, only the vowel pairs /e/-e/ and /a/-J/s/ showed no significant difference between both the F1 and F2 values for native speakers' perception. This ANOVA test indicates that all the other vowels that were identified by the native speakers in the perception task are distinct.

## 4.1.2 Production

Figure 12 shows the average values for perception, from Figure 10 above, along with the average values for production. As with the perception task, the averages of the F1 and F2 values for the 87 participants are consistent with previous theoretical and empirical work. For most vowels, there seems to be a close match between native speakers' perception and production. The two exceptions are /e/ and /æ/. Those two segments are higher and more forward in production than they are in perception. This phenomenon is consistent with the Northern Cities vowel Shift (Labov et al 1973, Labov 1994, Labov et al 1997). The shift involves the raising of low front vowels into the range occupied by mid and high front vowels. The acoustic properties of vowels on American English are changing as the vowel system undergoes a vowel chain shift. A chain shift is a repositioning of vowels within a system. This chain shift involves a

74

shifting of the positions of six low and mid vowels. As one vowel moves into the acoustic space occupied by another vowel, the other vowel moves as well. The cascade of movement results in a reorganization of the vowel system. Labov (1994: 178) characterizes the Northern Cities Shift as "one of the most vigorous sound changes now in progress in the United States."

This shift has resulted in a measurable change in the pronunciation of vowels in the affected areas of the United States. According to Labov et al (1997), Michigan is in the area affected by the vowel shift. The major trends in the shift are summarized in Table 9 below.

Original	Post-shift
/a/	/æ/
/ɔ/	/a/
/æ/	/ɛ/
/1/	/ɛ/
/ɛ/	/ə/
/ə/	/ɔ/

Table 9: Main effects of the Northern Cities Shift. Column 1 shows the vowel that is affected. Column 2 shows the vowel space that the shifting vowel moves into (from Labov et al 1997).

The research program examining this vowel shift centers on measuring production data. While the shift in pronunciation is well established, the relationship between the pronunciation of the vowels and the perception, if any, has been unclear. This vowel shift is changing the quality of vowels that native speakers of English in Michigan produce. The result of this shift will be a difference between native speakers' intuitions about the acoustic values of vowels, and those that they actually produce. There will be, in other words, a widening competence-performance distinction among speakers in areas affected by this vowel shift. In that light, the results of the production task are not surprising. What is interesting is that the shift in the low front vowels is only evident in production, not in perception.



Figure 12: Average Native Speaker Perception and Production

Figure 13 shows the aggregate production data for each vowel. The production of native speaker participants is much more uniform than their perception data, with very few within-group outliers.



Figure 13: Native speaker performance on production task

Several tokens of /i/ and /a/ fell outside the limits of the graph. Two factors may have influenced the values found for /i/. The gender of participants may have been a contributing factor. There were several female participants, while the synthesized vowels were in the range of a male speaker. Since the typical female vocal tract is shorter than that of the typical male, the formant frequencies of utterances by females are generally higher.

The other factor could have been the co-called "hyperspace effect" hypothesis of Johnson (2000), according to which participants choose hyperarticulated vowels in an apparent effort to maximize vowel contrast. Although there seemed to be a preference for /i/ sounds with very high F2 values, this effect seems to be limited to front vowels. If there was preference to hyperarticulate vowels to maximize vowel contrast, then we would expect to see a preference for back vowels to have lower F2 values. In fact, however, there was no evidence of preference for the other high vowel in this study, /u/, to have very low F2 values. In fact, only one participant chose the 450 Hz F2 value for /u/, which was the back-most value available to participants, and only four made an F2 selection that was less than 1000 Hz. In production, the minimum F2 value was 727 Hz, and only 5 participants produced an F2 less than 1000 Hz.

The other unexpected finding was the range of F1 values of the vowel /a/. This vowel had higher F1 values than was predicted in the literature. Ladefoged (1993:193) indicates a F1 value for /a/ at 710 Hz. In this study, however, the F1 values for /a/ ranged from 459 Hz to 1279 Hz for production. That F1 values would be as high as they were in the production task was not anticipated in the creation of the perception task of the instrument. The vowel matrix provided vowel sounds with a maximum F1 value of 1000 Hz. In the perception task, the range for /a/ was 250 to 1000 Hz. The high F1 values in production suggest that the perception task should have included a greater range for F1. Ten participants chose the maximum F1 value in the perception task (100 Hz), and 9 participants produced F1 values in excess of 1000 Hz.

78

The means for the vowels are given in Table 10. Standard deviations for each mean are shown in parentheses below the mean.

Vowel	F1	F2
i	305	2673
	(100)	(263)
I	505	2191
	(95)	(255)
e	461	2480
	(88)	(254)
3	738	1912
	(140)	(240)
æ	699	2249
	(183)	(261)
a	879	1489
	(160)	(193)
ə	706	1396
	(139)	(154)
Э	779	1219
	(96)	(223)
0	607	1377
	(102)	(202)
υ	576	1226
	(90)	(318)
u	407	1352
	(57)	(353)

Table 10: Mean formant values for native speakers: production task. The standard deviation for each mean is shown in parentheses below the mean.

F1 and F2 means for the vowels  $\epsilon$  and  $\pi$  appeared to be close together. The F1 values differed by only 39 Hz, and the F2 values differed by 237 Hz. T-tests were performed to determine the significance of the differences. The null hypothesis was that there would be no difference. The t-test for the means for F2 yielded t(91) = -6.51, p = 0.00, and the t-test for F1 yielded t(86) = 1.19, p = 0.24. The observed value for F2 exceeded the critical value of 1.99, but the observed value for F1 did not. We can thus

reject the null hypothesis that there is no difference between the respective F2 values for  $\epsilon/and/ac/$ . There is a statistically significant difference in frontness between these two vowels. However, we cannot reject the null hypothesis that there is no difference between the F1 values for that pair of vowels.

An ANOVA test was performed to show the level of variation among vowels in the production system. The p values for each vowel pair in the post-hoc pairwise comparison is given in Table 11.

	Ι	e	ε	æ	a	ə	Э	0	U	u
i	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
I		0.812	0.000	0.000	0.000	0.000	0.000	0.002	0.140	0.005
e			0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.548
3				0.886	0.000	0.971	0.869	0.000	0.000	0.000
æ					0.000	1.000	0.052	0.013	0.000	0.000
a						0.000	0.004	0.000	0.000	0.000
ə							0.121	0.004	0.000	0.000
Э								0.000	0.000	0.000
0									0.978	0.000
υ										0.000

	Ι	e	ε	æ	а	ə	Э	0	ប	u
i	0.000	0.015	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Ι		0.000	0.000	0.992	0.000	0.000	0.000	0.000	0.000	0.000
e			0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000
3				0.000	0.000	0.000	0.000	0.000	0.000	0.000
æ					0.000	0.000	0.000	0.000	0.000	0.000
a						0.803	0.000	0.565	0.000	0.266
ə							0.038	1.000	0.057	0.999
С								0.106	1.000	0.301
0									0.149	1.000
U										0.385

Table 11: The results of a pairwise comparison from an ANOVA test shows the source of variation between pairs of vowels in native speakers' responses to the perception task in the F1 Formant (top) and F2 formant (bottom). Only one pair of vowels (/o/-/u/) showed no significant difference between them.

As outlined in the section on native speaker perception in 4.1.1, we assume that vowels are not distinct only if the p value for both F1 and F2 are over 0.05. Only one pair of vowels, /o/ and /u/, showed no significant difference between both the F1 and F2

values for native speakers' production. This ANOVA test indicates that all the other vowels that were identified by the native speakers in the production task are distinct.

4.2 Nonnative speakers

## 4.2.1 Perception

Figure 14 shows the average values of the nonnative speakers' choices in the perception task. While the vowels are arranged in the general pattern that indicates the participants identified the synthesized vowels as speech sounds analogous to natural speech, the pattern has noticeable differences from that of native speaker participants. First is the close convergence of /i/ and /I/. The average of F2 for /i/ was 2382 Hz, and 2419 Hz for /I/. The average F1 was the same for both vowels: 335 Hz.

The very low placement of  $\epsilon$  and  $\epsilon$  clusters these very closely with k. It seems that nonnative speakers had difficulty discerning these three vowels in the perception task. The average F1 for  $\epsilon \epsilon$  w were 676, 713, 759, respectively, and the F2 values were 2160, 2174, 2128, respectively.

Nonnative participants also placed / $\upsilon$ / close to /u/, which may mean that they had difficulty differentiating those two vowels as well. The placement of / $\sigma$ / was higher than /a/, near / $\sigma$ /.



Figure 14: Average Nonnative Speaker Perception

Individual participants' choices are shown in the scatterplot diagrams in Figure 15. As with native speakers, there is considerable variation in the positions of the tokens for each vowel, although for each vowel, a general trend can be seen. The grouping of /i/ and /i/ along the left edge of the grid indicates that the participants could have preferred tokens with F2 values that were even higher than those that were available for this task. Fifteen participants, more than half, chose the maximum F2 value for /i/, and fourteen chose the maximum F2 value for /i/. For all the front vowels /i, 1, e,  $\varepsilon$ ,  $\alpha$ /, there is a preference trend toward more fronting, represented by higher F2 values. This could be a result of first language influence.



Figure 15: Nonnative speaker choices for perception task

Table 12 shows the means of F1 and F2 for the vowels that were tested for in the perception task for nonnative speakers.

Vowel	F1	F2
i	335	2382
	(127)	(467)
I	335	2419
	(103)	(284)
e	676	2160
	(182)	(516)
ε	713	2174
	(203)	(484)
æ	759	2128
	(212)	(515)
a	774	1244
	(232)	(446)
ə	743	1378
	(156)	(393)
Э	606	1220
	(185)	(539)
0	615	989
	(193)	(523)
U	400	1072
	(164)	(405)
u	361	1142
	(151)	(527)

Table 12: Means for the nonnative speaker perception task

T-tests were performed to test the difference between the means for several pairs of vowels that varied by less than 50 Hz for F1, and 100 Hz for F2. The results are summarized in Table 13 below. In every t-test, we could not reject the null hypothesis that the means were statistically the same. The vowels /i I/ were virtually the same. Three 2-way t-tests were performed on the combinations of /e  $\varepsilon \alpha$ /. There was no statistically significant difference found in any of the tests. It may be accurate to say that the participants' responses to the perception task revealed a 5-vowel interlanguage system, as shown in Figure 16.

Distinction	Formant	Observed	Significant
		value	
i-I	F1	0.00	NS
	F2	-0.35	NS
e-e	F1	-0.71	NS
	F2	-0.10	NS
e-æ	F1	-1.55	NS
	F2	0.23	NS
£-æ	F1	-0.82	NS
	F2	0.34	NS
ə-a	F1	-0.58	NS
	F2	1.17	NS
u-U	F1	0.91	NS
	F2	0.54	NS
0-0	F1	-0.18	NS
	F2	1.60	NS

Table 13: Between-category t-tests: nonnative speaker perception



Figure 16: the Nonnative vowel areas: perception task

An ANOVA test was performed to show the level of variation among vowels in nonnative speakers' perception. The p values for each vowel pair in the post-hoc pairwise comparison is given in Table 14.

	I	e	3	æ	a	Э	С	0	υ	u
i	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.960	1.000
I		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.960	1.000
e			1.000	0.820	0.625	0.952	0.932	0.974	0.000	0.000
3				0.997	0.974	1.000	0.488	0.625	0.000	0.000
æ					1.000	1.000	0.059	0.101	0.000	0.000
a						1.000	0.023	0.042	0.000	0.000
ə							0.148	0.229	0.000	0.000
Э								1.000	0.001	0.000
0									0.001	0.000
ប										0.999
	r			· · · · · · · · · · · · · · · · · · ·						
	I	e	3	æ	a	ອ	Э	0	υ	u
i	1.000	0.814	0.868	0.654	0.000	0.000	0.000	0.000	0.000	0.000
I		0.628	0.703	0.448	0.000	0.000	0.000	0.000	0.000	0.000
e			1.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
3				1.000	0.000	0.000	0.000	0.000	0.000	0.000
æ					0.000	0.000	0.000	0.000	0.000	0.000
a						0.994	1.000	0.654	0.961	0.999
ə							0.978	0.088	0.375	0.750
э								0.772	0.986	1.000
0									1.000	0.983
U										1.000

Table 14: The results of a pairwise comparison from an ANOVA test shows the source of variation between pairs of vowels in nonnative speakers' responses to the perception task for the F1 formant (Top) and F2 formant (Bottom). Several pairs of vowels  $(/i/-/i/, /e/-/\epsilon/, /e/-/æ/, /\epsilon/-/æ/, /a/-/ə/, /a/-/o/, /o/-/o/, and /u/-/u/)$  showed no significant difference between them.

The results of the ANOVA test were consistent with the t-tests. There was no

difference between several vowel pairs:  $\frac{i}{-1}$ ,  $\frac{e}{-2}$ ,  $\frac{e}{-2}$ ,  $\frac{e}{-2}$ ,  $\frac{a}{-2}$ ,

/3/-/0/, and /v/-/u/.

#### 4.2.2. Production

The difference between perception and production of nonnative speakers is shown in Figure 17. In the production task, there is greater separation of /i/ and /u/, however, /u/is higher than /i/ and is almost as fronted. The positions of /e/ and /ae/ are even closer in the production task than in the perception task, almost completely overlapping. Because of the great variation of /o/ among the participants, its average position is between /e/ and /e/.



Figure 17: Average Nonnative Speaker Perception and Production

The plots of all tokens for the production task in Figure 18 show the variation among tokens, but similar to the production data in native speakers, there was greater uniformity among tokens in the production task than for the perception task. As with the native speakers, there were several instances of high F2 values exceeding the limits of the grid. This again could be because of the gender of the participants. Female voices tend to have higher F2 values than males because of their shorter vocal tract. Also similar to native speakers, there were many instances of a high value of F1 for the vowel /a/. As with the native speakers, the nonnative speaker participants exhibited a preference for a greater range in F1 than anticipated. There were 5 instances of F1 values in excess of 1000 Hz for /a/, and two cases of F1 values over 1000 Hz for /ɔ/. The greater range of F1 in the production task suggests that the perception task would have been able to more accurately reflect the intuitions of the participants if the synthesized vowels in the vowel matrix had included a greater range for F1.



Figure 18: Nonnative speaker performance on production task
Vowel	F1	F2
i	356	2584
	(51)	(369)
Ι	399	2537
	(58)	(389)
e	572	2349
	(97)	(409)
3	705	2106
	(135)	(339)
æ	674	2058
	(121)	(422)
a	879	1368
	(202)	(284)
ə	656	1096
	(132)	(153)
Э	628	1275
	(211)	(536)
0	547	1195
	(118)	(471)
U	437	1175
	(144)	(413)
u	429	1062
	(139)	(264)

Table 15: Means for the nonnative speaker production task. The standard deviation for each mean is shown in parentheses below the mean.

The averages for the vowels /i t/ and / $\epsilon$  æ/ were very close to each other, all within 50 Hz, so t-tests were performed to test for significance of difference. The t-test for the pair /i-t/, the t-test yielded t(48) = 0.44, p = 0.66 for F2, and t(47) = -2.77, p = 0.01 for F1. For the pair / $\epsilon$ -æ/, the t-test yielded t(46) = 0.45, p = 0.66 for F2, and t(47) = 0.85, p = 0.40 for F1. The only case in which the null hypothesis could be rejected was the F1 values of /i-t/, which had an observed value of -2.66, greater than the critical value of 2.01. Those vowels could be distinguished only on the basis of height.

An ANOVA test was performed to show the level of variation among vowels in nonnative speakers' production. The p values for each vowel pair in the post-hoc pairwise comparison is given in Table 16.

	Ι	e	ε	æ	a	ə	э	0	υ	u
i	1.000	0.584	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
I		0.848	0.007	0.001	0.000	0.000	0.000	0.000	0.000	0.000
e			0.537	0.260	0.000	0.000	0.000	0.000	0.000	0.000
ε				1.000	0.000	0.000	0.000	0.000	0.000	0.000
æ					0.000	0.000	0.000	0.000	0.000	0.000
a						0.354	0.999	0.905	0.824	0.196
ə							0.882	0.998	1.000	1.000
э								1.000	0.998	0.720
0									1.000	0.984
υ										0.996

	Ι	e	ε	æ	a	ə	Э	0	U	u
i	0.993	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.644	0.773
Ι		0.001	0.000	0.000	0.000	0.000	0.000	0.013	0.997	1.000
e			0.043	0.292	0.000	0.587	0.954	1.000	0.035	0.019
ε				1.000	0.001	0.981	0.698	0.005	0.000	0.000
æ					0.000	1.000	0.986	0.064	0.000	0.000
a						0.000	0.000	0.000	0.000	0.000
ə							1.000	0.198	0.000	0.000
Э								0.649	0.000	0.000
0									0.189	0.118
υ										1.000

Table 16: The results of a pairwise comparison from an ANOVA test shows the source of variation between pairs of vowels in nonnative speakers' responses to the production task for the F1 formant (Top) and F2 formant (Bottom). Several pairs of vowels (/i/-/i/, /e/-/æ/,  $\epsilon/-\frac{1}{2}$ ,  $\frac{1}{2}$ ,  $\frac$ 

The results of this ANOVA test show that there was no difference between several vowel pairs: /i/-/I/, /e/-/æ/, /e/-/æ/, /9/-/0/, /9/-/0/, /0/-/U/, /0/-U/, and /U/-/U/.

# 4.3 Perception-production differences

# 4.3.0. Introduction

Table 17 shows the mean F1 and F2 values for perception and production for each vowel, separated by native and nonnative speaker group. The standard deviation for each mean is shown in parentheses below the mean.

		Native S	Speakers	Nonnative Speakers			
Vowel		Perception	Production	Perception	Production		
i	F1	346	290	335	390		
		(176)	(101)	(127)	(52)		
	F2	2320	2659	2382	2640		
		(459)	(266)	(467)	(377)		
Ι	F1	529	501	335	350		
		(137)	(96)	(103)	(59)		
	F2	2123	2185	2419	2617		
		(450)	(258)	(284)	(398)		
e	F1	632	460	676	689		
		(135)	(89)	(182)	(99)		
	F2	2220	2477	2160	2475		
		(480)	(257)	(516)	(418)		
3	<b>F</b> 1	652	728	713	689		
		(172)	(141)	(203)	(138)		
	F2	2156	1912	2174	2074		
		(405)	(243)	(484)	(346)		
æ	Fl	803	/00	(212)	692		
		(153)	(185)	(212)	(123)		
	F2	2138	2240	2128	20/9		
	<b>F</b> 1	(422)	(204)	(515)	(431)		
a	F1	(162)	8/3 (161)	(222)	(206)		
	ED	(103)	(101)	(232)	(200)		
	ΓΖ	(301)	(101)	(1244)	(200)		
	F1	(391)	(191)	7/3	(290)		
ð	11	(171)	(141)	(156)	(135)		
	F2	1331	1398	1378	1041		
	1 2	(390)	(156)	(393)	(156)		
2	F1	779	777	606	624		
Ŭ	• •	(150)	(97)	(185)	(216)		
	F2	1291	1222	1220	2188		
		(429)	(225)	(539)	(547)		
0	F1	613	611	615	696		
_		(178)	(104)	(193)	(120)		
	F2	909	1385	989	2123		
		(356)	(204)	(523)	(481)		
υ	F1	508	578	400	495		
		(175)	(91)	(164)	(147)		
	F2	1067	1183	1072	1401		
		(346)	(322)	(405)	(422)		
u	F1	409	407	361	399		
		(184)	(57)	(151)	(142)		
	F2	953	1333	1142	1239		
		(365)	(357)	(527)	(270)		

Table	17:	Mean	and	standard	deviation	values	for p	erception	n and	production,	native	and
					nonn	ative sp	peake	ers				

4.3.1 Intra-group perception versus production

Tables 18 and 19 show the results of ANOVA tests measuring the difference between the perception and production tasks. One prediction of the study was that both perception and production would be more closely aligned among native speakers than nonnative speakers.

	Ι	e	ε	æ	a	ə	Э	0	υ	u
i	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
I		0.111	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
e			0.000	0.000	0.000	0.000	0.000	0.651	0.379	0.000
ε				0.000	0.000	1.000	0.000	0.006	0.000	0.000
æ					0.008	0.000	1.000	0.000	0.000	0.000
a						0.000	0.060	0.000	0.000	0.000
ə							0.000	0.044	0.000	0.000
Э								0.000	0.000	0.000
0									0.001	0.000
υ										0.000
	I	e	3	æ	a	ə	Э	0	υ	u
i	0.000	0.105	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Ι		0.013	0.871	1.000	0.000	0.000	0.000	0.000	0.000	0.000
e			0.000	0.102	0.000	0.000	0.000	0.000	0.000	0.000
3				0.442	0.000	0.000	0.000	0.000	0.000	0.000
æ					0.000	0.000	0.000	0.000	0.000	0.000
a						1.000	0.698	0.000	0.000	0.000
ə							0.690	0.000	0.000	0.000
э								0.000	0.039	0.003
0										
									0.979	1.000
U U									0.979	1.000

Table 18: The results of ANOVA tests comparing perception and production among native speaker subjects for F1 (Top) and F2 (Bottom). Only the vowel pair /a/ - /ɔ/ shows no significant difference (p>.05) between pairs for both F1 and F2.

	Ι	e	ε	æ	а	ə	Э	0	υ	u
i	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.454	0.916
Ι		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.864	0.999
e			0.262	0.128	0.000	0.400	1.000	0.950	0.000	0.000
3				1.000	0.016	1.000	0.115	0.004	0.000	0.000
æ					0.044	1.000	0.048	0.001	0.000	0.000
a						0.007	0.000	0.000	0.000	0.000
ə							0.199	0.009	0.000	0.000
э								0.994	0.000	0.000
0									0.000	0.000
U										1.000
	I	e	3	æ	a	ə	Э	0	υ	u
i	1.000	0.222	0.005	0.001	0.000	0.000	0.000	0.000	0.000	0.000
Ι		0.236	0.006	0.001	0.000	0.000	0.000	0.000	0.000	0.000
e			0.977	0.783	0.000	0.000	0.000	0.000	0.000	0.000
ε				1.000	0.000	0.000	0.000	0.000	0.000	0.000
æ					0.000	0.000	0.000	0.000	0.000	0.000

					0.000	0.000	0.000	0.000	0.000	0.000	
а						1.000	1.000	0.296	0.567	0.432	
ə							1.000	0.749	0.935	0.866	
Э								0.740	0.931	0.859	
0									1.000	1.000	
υ										1.000	
Tab	Table 19: The results of ANOVA tests comparing perception and production among										

Table 19: The results of ANOVA tests comparing perception and production among nonnative speaker subjects for F1 (Top) and F2 (Bottom). The vowel pairs /i/-/1/, /e/-/ε/,/ε/-/æ/,/ε/-/æ/,/o/-/o/,and /υ/-/u/ show no significant difference (p>.05) between pairs for both F1 and F2.

consistent between perception and production than nonnative speakers would.

### 4.3.2. Cross-group perception and production

Table 20 gives the standard deviations of F1 and F2 for each vowel for native speakers (NS) and nonnative speakers (NNS). For comparison purposes, and where comparable vowels are available, the standard deviations from the study by Peterson and Barney (1952) are also given, labeled "PB." The category that has the greatest standard deviation, and thus the greatest amount of within-group variation, is marked in bold. There is a striking difference between F2 and F1. Nonnative speakers have the greater variation in F2, which measures frontness, than native speakers for almost all vowels. The two groups are almost evenly split in number of categories with the more variation for F1, which measures vowel height.

One hypothesis of the study was that the native speaker group in this study, which was composed of members from the same speech community, would exhibit less variation than nonnative speakers. While the data in Table 20 seems to support that hypothesis for vowel frontness, as measured by F2, there seems to be no trend either way in vowel height, measured by F1.

			F2	· · · · · · · · · · · · · · · · · · ·		F1				
Vowel		PB	NS	NNS		PB	NS	NNS		
i	Perception		459	467	[		176	127		
	Production	374	266	377		60	101	52		
I	Perception		450	284			137	103		
	Production	337	258	398		75	96	59		
e	Perception		480	516			135	182		
	Production	336	257	418		97	89	99		
ε	Perception		405	484			172	203		
	Production		243	346			141	138		
æ	Perception		422	515			153	212		
	Production	288	264	431		172	185	123		
a	Perception		391	446			163	232		
	Production	157	195	290		146	161	206		
ə	Perception		390	393	1 [		171	156		
	Production	190	156	156		113	141	135		
э	Perception		429	539	1 F	· · · · ·	150	185		
	Production		225	547			97	216		
0	Perception		356	523			178	193		
	Production	144	204	481		96	104	120		
U	Perception		346	405			175	164		
	Production	194	322	422		71	91	147		
u	Perception		365	527	[		184	151		
	Production	220	357	270		76	57	142		

Table 20: Standard deviations for F1 and F2 values of vowels in the perception and production tasks for native and nonnative speakers. For each vowel, the first line is the perception value, and the second line is the production value. The column "PB" is the values from Peterson and Barney (1952). The value for the group with the greatest variation is in boldface.

Another hypothesis of the study was that nonnative speaker perception would be more native-like than their production. To test this hypothesis, an ANOVA test was performed, comparing native and nonnative speakers' performance on both the perception and the production tasks. Table 21 summarizes the results of the pairwise comparison for perception, and Table 22 summarizes the results of the pairwise comparison for production.

	Ι	e	ε	æ	а	Э	э	0	U	u
i	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.297
Ι		0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.003
e			0.990	0.000	0.000	0.943	0.000	0.964	0.000	0.000
3				0.000	0.000	1.000	0.034	0.344	0.000	0.000
æ					1.000	0.000	0.320	0.000	0.000	0.000
a						0.000	0.081	0.000	0.000	0.000
ခ							0.085	0.180	0.000	0.000
э								0.000	0.000	0.000
0									0.000	0.000
U										0.004
r	r									
	I	e	3	æ	a	ə	Э	0	ប	u
i	0.256	0.421	0.064	0.015	0.000	0.000	0.000	0.000	0.000	0.000
I		1.000	1.000	0.996	0.000	0.000	0.000	0.000	0.000	0.000
e			0.999	0.974	0.000	0.000	0.000	0.000	0.000	0.000
3				1.000	0.000	0.000	0.000	0.000	0.000	0.000
æ					0.000	0.000	0.000	0.000	0.000	0.000
a						0.979	1.000	0.000	0.010	0.000
ə							0.982	0.000	0.000	0.000
э								0.000	0.009	0.000
0									0.277	0.977
U										0.968

Table 21: The results of ANOVA tests comparing perception between native and nonnative speaker subjects for F1 (Top) and F2 (Bottom). The vowel pairs  $\frac{}{-\frac{1}{2}}, \frac{\frac{1}{2}}{2}$ ,  $\frac{\frac{1}{2}}{2}, \frac{\frac{1}{2}}{2}$ , show no significant difference (p>.05) between pairs for both F1 and F2.

	i	Ι	e	ε	æ	a	ə	Э	0	υ	u
i		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
Ι	0.000		0.941	0.000	0.000	0.000	0.000	0.000	0.000	0.161	0.305
e	0.000	0.941		0.000	0.000	0.000	0.000	0.000	0.002	0.956	0.004
ε	0.000	0.000	0.000		0.828	0.000	0.801	1.000	0.000	0.000	0.000
æ	0.000	0.000	0.000	0.828		0.000	1.000	0.822	0.000	0.000	0.000
a	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000
ə	0.000	0.000	0.000	0.801	1.000	0.000		0.794	0.000	0.000	0.000
Э	0.000	0.000	0.000	1.000	0.822	0.000	0.794		0.000	0.000	0.000
0	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000		0.198	0.000
υ	0.000	0.161	0.956	0.000	0.000	0.000	0.000	0.000	0.198		0.000
u	0.001	0.305	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
<b></b>						r	,	r	r		
	i	Ι	e	3	æ	a	ə	Э	0	U	u
i		0.000	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
I	0.000		0.359	0.000	0.343	0.000	0.000	0.000	0.000	0.000	0.000
e	0.004	0.359		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000		0.004	0.000	0.000	0.000	0.000	0.000	0.000
æ	0.000	0.343	0.000	0.004		0.000	0.000	0.000	0.000	0.000	0.000
a	0.000	0.000	0.000	0.000	0.000		0.100	0.003	0.268	0.000	0.009
ə	0.000	0.000	0.000	0.000	0.000	0.100		0.994	1.000	0.872	1.000
Э	0.000	0.000	0.000	0.000	0.000	0.003	0.994		0.930	1.000	1.000
0	0.000	0.000	0.000	0.000	0.000	0.268	1.000	0.930		0.620	0.984
υ	0.000	0.000	0.000	0.000	0.000	0.000	0.872	1.000	0.620		0.999
u	0.000	0.000	0.000	0.000	0.000	0.009	1.000	1.000	0.984	0.999	

Table 22: The results of ANOVA tests comparing production between native and nonnative speaker subjects for F1 (Top) and F2 (Bottom). The vowel pairs /I/-/e/, /o/-/o/, /o/-/o/ show no significant difference (p>.05) between pairs for both F1 and F2.

The data in Tables 21 and 22 show the source of the variation between subject groups in the two tasks. In the perception task, there was no significant difference between 24 out of 110 pairs of vowels in the F1 formant, and 30 out of 110 for the F2. For the production task, there were no significant differences between 22 out of 110 pairs in the F1 formant, and 28 out of 110 for the F2 formant. Three vowel pairs (/e/-/ $\epsilon$ /, /a/-/o/,

### **Chapter 5: Discussion and Conclusions**

## 5.0. Introduction

In this section, I review the research questions and hypotheses of the study, and evaluate them in light of the results. Following that, I discuss some of the new findings of the study, and the implications of the findings for the fields of linguistics and second language acquisition. Finally, I discuss the directions that future studies could take to further the exploration along the lines of this study.

5.1. Review of research questions and hypotheses

The research questions of this study were first given in Chapter 1, section 1.5, and are repeated here:

- 1. To what extent can we determine the acoustic properties of the ideal vowel phoneme of each monothongal vowel in English based on native speakers' intuitions?
- 2. To what extent do nonnative speakers of English agree on the ideal vowel phoneme of each monothongal vowel in English?
- 3. Are the intuitions of nonnative learners of English similar to the intuitions of native speakers with regard to the vowel phonemes of English?

4. How similar is production of vowels in English by both native speakers of English and nonnative learners of English to their respective identification of an ideal vowel?

This study sought to address the research questions. The hypotheses were that nonnative speakers' first language would influence their performance on the production task, and their perception would be more native-like than their production. In Chapter 1, the following hypotheses were made:

- Given a perception task in which participants identify an English vowel from a continuum of F1-F2 combinations (the "perception task"), there will be less variation among native speakers than among nonnative speakers.
- 2. In the comparison of the performance of native speakers and nonnative speakers on the perception task and a "production task" in which participants produce words containing the English vowels, there will be less variation between the two groups' performance on the perception task than there will be on the production task.
- 3. In a perception and a production task, there will be less variation between the two tasks among native speakers than among nonnative speakers.

5.2.1. Hypothesis 1: Native speakers will show less variation than nonnative speakers in perception

This hypothesis was motivated by the assumption that members of a language community have a common linguistic competence. If all the native speaker participants of the study were from the same language community, then they should show similar responses on the perception task. The data largely supports this hypothesis. There is more variation among nonnative speakers than among native speakers.

The range in values of responses by native speakers was unexpected, given the assumption of uniformity that underpinned Hypothesis 1. Other studies have shown a similar phenomenon, however, suggesting that the variation shown by native speakers in this study may actually be the norm for native speakers. A wide range in judgments on a perception task was noted in Frieda et al (1999). The Peterson and Barney (1952) study showed a wide range of responses in production as well. In light of the results of previous studies, the wide range of values found in the perception task in this study is not unusual.

The variation of responses within each group of participants of this study was measured by calculating the standard deviation for each vowel. The means and standard deviation values for the F1 and F2 value for each of the eleven monothongal English vowels that were used in this study are shown in Table 23 below. The table shows the data by native speaker and nonnative speaker participant groups. The hypothesis was that

the nonnative speaker group would show more variation, which could be evidenced by a higher standard deviation from the mean.

Vowel		Native S	Speakers	Nonnative Speakers			
		Mean	Standard	Mean	Standard		
			Deviation		Deviation		
i	F1	346	176	335	127		
	F2	2320	459	2382	467		
I	F1	529	137	335	103		
	F2	2123	450	2419	282		
e	F1	632	135	676	182		
	F2	2220	480	2160	516		
ε	F1	652	172	713	203		
	F2	2156	405	2174	484		
æ	F1	803	153	759	212		
	F2	2138	422	2128	515		
a	F1	815	163	774	232		
	F2	1283	391	1244	446		
ə	F1	652	171	743	156		
	F2	1331	390	1378	393		
Э	F1	779	150	606	185		
	F2	1291	429	1220	539		
0	F1	613	178	615	193		
	F2	909	356	989	523		
υ	F1	508	175	400	164		
	F2	1067	346	1072	405		
u	F1	409	184	361	151		
	F2	953	365	1142	527		

Table 23: Mean and standard variation values on the perception task for 11 simple vowel phonemes of English, native and nonnative speakers. The higher value of native and nonnative speakers is shown in boldface.

Table 23 shows in boldface type the higher standard deviation value, whether of native or nonnative speaker participant group. In only 6 out of 22 of the formant values there was greater variation among native speaker participants than nonnative speakers. This data supports Hypothesis 1. As reflected in standard deviation, nonnative speaker participants showed more variation than native speakers in the majority of vowels.

Although greater variation was seen among nonnative speakers in 16 out of 22 of the formant values, there were some items in which the data from native speakers had greater within-group variation. Native speakers had greater variability in the F1 formant for /i/. The tendency among some native speakers of English to prefer a hyperarticulated /i/ was documented by Johnson et al (1993) and Johnson (2000). It is possible that the greater variation of /i/ exhibited by native speakers does not reflect their actual competence, but is influenced by what Johnson et al termed the "hyperspace effect," in which there is a preference for the /i/ phoneme that is beyond the scope of the speaker's normal range for the vowel. However, the greater variation was limited to one formant, F1, which reflects the height of the vowel. Native speakers show more consensus for the dimension of frontness, which is measured by the F2 formant value.

Native speakers also showed more variety for /1/ than nonnatives both in F1 and F2. This may be a result of the different category distinction between groups. As noted in Chapter 4, a statistical analysis shows that nonnative speakers seem not to have separate phonemes for /i/ and /1/. The nonnative group's /1/ is statistically indistinguishable from /i/. It seems that the nonnative speaker participants have not differentiated the two vowels in English, and are quite uniform in their judgments that the two English vowels are the

same. If that is the case, then a comparison between nonnative /I and native /I may not be meaningful.

The other items in which native speakers exhibited more variety were the F1 values in the vowels / $\varphi$   $\upsilon$  u/.

5.2.2. Hypothesis 2: There will be a less variation between the two groups' performance on the perception task than there will be on the production task.

Hypothesis 2 was that there would be less variation between native and nonnative speakers in perception than in production. This hypothesis goes beyond the observation that nonnative speakers have foreign accents, and posited that in the course of second language acquisition, perception could be more native-like than production. This hypothesis was that foreign accents would affect the nonnative speakers' production, but would not necessarily reflect their perceptual competence. The hypothesis was tested by comparing the two groups' performance on the two tasks.

The ANOVA test that analyzed the differences between native and nonnative speakers was discussed in section 4.3.2 of Chapter 4. The results of that test showed the sources of differences between subject groups. As was mentioned in section 4.3.2, there was no clear pattern of differences between the comparisons of perception and production between native and nonnative speakers.

5.2.3 Hypothesis 3: There will be a less variation between the performance on the perception task and the production task by native speakers than by nonnative speakers.

As discussed in section 4.3.1 of Chapter 4, Tables 18 and 19 showed ANOVA comparisons of perception and production within participant groups. The native speakers showed much more agreement between perception and production than the nonnative speakers did.

### 5.2.4 Summary

Hypothesis 1 was that on the perception task, there would be less variation among native speakers than among nonnative speakers. Although there was more variation among the native speakers than was expected on the perception task, there was more variation among nonnative speakers than native speakers. Hypothesis 1 was confirmed.

Hypothesis 2 was that there would be a closer correlation between the two groups' performance on the perception task than on the production task. The data that was gathered could neither support nor refute Hypothesis 2.

Hypothesis 3 was that there would be a closer correlation between the performance on the perception task and the production task by native speakers than by nonnative speakers. Hypothesis 3 was confirmed.



#### 5.3 Additional findings of the study

#### 5.3.1 Production less variable than perception

An unexpected finding of the study was that there was greater variation in the perception task than in the production task. Table 20 above showed the standard deviation value for each vowel, separated by task and participant group. On 18 of the 22 items, participants of both groups showed more internal consistency in the production task than in the perception task.

While the general clustering of each token was in the correct general area for each vowel on the perception task, there was much greater variation among participants than was anticipated. The data for the perception task showed a larger number of outliers for each vowel than was found in the production task. It could be that participants could not make an accurate identification using the synthesized vowels. It could also be the case that perception is more variable among individuals within a speech community than this study assumed. More research that focuses on variation in perception is needed.

5.3.2 Conflation of /a-ɔ/ in native speakers' perception

A dialectical distinction in areas of the United States is the conflation of the vowels /a/ and /ɔ/. Research done by Labov et al (1997) indicates that Michigan is an area in which the vowels typically show a distinction. The native speaker participants were all from Michigan, and so would be predicted to maintain the distinction between the vowels. Indeed, in the production task, native speakers produced statistically distinct vowels for

each vowel. Data from the production task is consistent with Labov et al's study. Section4.1 of Chapter 4 gave the results of statistical analyses of differences.

Labov (1994: 363) refers to inconsistency within a participant on vowel merging as the "Bill Peters effect," named for a participant who merged vowels on a formal task, but distinguished them in casual speech. Labov attributes the difference in performance to the level of formality while speaking. Labov's study involved reading word lists. However, a similarity can be drawn to the perception task of this study. Another of his participants ("Mrs. V," p. 363) hesitated before reading the *cot/caught* pair, stating uncertainty as to the distinction between the two words before she read them aloud. On the *Don/Dawn* pair, the participant evaluated the vowel sounds as "slightly different," but produced an almost indistinguishable vowel sound when producing them. Although not the focus of the study, Labov's work documents a perception-production distinction on the /a-o/ vowels that this study also found.

## **5.4 Conclusions**

This study examined the phonological systems of Korean learners of English as a second language, and compared their systems to those of native speakers. While the data supported some of the hypotheses, some unexpected results give reason to reconsider some assumptions about language learners and phonological competence. I conclude with addressing the research questions of Chapter 1.

5.4.1. Research question 1: To what extent can we determine the acoustic properties of the ideal vowel phoneme of each monothongal vowel in English based on native speakers' intuitions?

Hypothesis 1 and Hypothesis 2 were predicated on the assumption of the uniform judgments of an idealized speech community composed of homogenous hearer-speakers. Although each group was made up of speakers from similar geographic regions and age groups, the data from the perception task in this study did not reflect such uniformity. The high level of variation suggests that with respect to vowel categories, even within a speech community it is difficult to define something even as basic as the phonological categories of the language. The scope of this study did not explore variation in depth, partly because it was assumed that the amount of variation would be negligible, and so this finding should be regarded as tentative.

In spite of the unexpected level of within-group variation, on average, native speakers of English displayed less variation than the nonnative speakers did. When asked to identify a given English vowel sound, native speakers were more uniform in their choice than nonnative speakers were. It seems that it is more accurate to view the degree of variation, not its absence or presence, in examining a vowel system. An answer to Research question 1 could be that we can give a probabilistic prediction of the acoustic properties of ideal vowels in English, and presumably, other languages as well.

5.4.2 Research question 2. To what extent do nonnative speakers of English agree on the ideal vowel phoneme of each monothongal vowel in English?

In one aspect, the data from nonnative speakers was closer to the assumptions of the study than the data from native speakers. The responses of the perception and production tasks for nonnative speakers showed more consistency. There was far less variance between nonnative speakers' choices of a vowel on the perception task, and their production of the same vowel.

Another aspect of the nonnative speaker group, however, did not display nativelike tendencies. The greater within-group variation on the perception task among nonnative speakers did not indicate agreement within the group about the English vowel categories. On this basis, it is difficult to determine the value of an ideal vowel for nonnative learners of English. With regards to a match between perception and production, the nonnative speakers showed less evidence of being a "speech community" as described in Chomsky (1965:3) and Saussure (1916:19) than native speakers did.

Based on the performance of the nonnative speaker group in this study, the answer to Research question 2 is that nonnative speakers show even less agreement for an ideal vowel segment than the loose agreement that native speakers show.

5.4.3 Research question 3: Are the intuitions of nonnative learners of English similar to the intuitions of native speakers with regard to the vowel phonemes of English?

The perception task was designed to measure participants' intuitions about part of the phonological inventory of English. What they perceived as the correct vowel among the choices available to them is assumed to be their intuition about that English vowel. The results of the perception task show that there is a significant correlation between the intuitions of native and nonnative speakers. This study indicates that the interlanguage phonological inventories of the participants are native-like in many respects.

The role of the native language could be a factor in this finding. Although the acoustic properties of vowel categories vary across languages, there is overlap in many vowels. An individual's intuition about a nonnative vowel category could be influenced by the intuition about the corresponding vowel category in the native language. Nonnative speakers must develop categories for vowels in the second language that do not exist in the first language. For Korean learners of English, these new vowels are /1,  $\alpha$ ,  $\sigma$ ,  $\sigma$ ,  $\sigma$ . The results of this study show that nonnative speakers do not show evidence of having formed native-like categories for these new vowels. The data in Table 13 in section 4.2.1 of Chapter 4 indicates that the intuitions of nonnative speakers about these new vowels of English are not distinct from vowels that are similar to Korean vowels. Comparisons of intuitions about new vowel categories in the second language need to take into account the role of the first language.

5.4.4 Research question 4: How similar is production of vowels in English by both native speakers of English and nonnative learners of English to their respective identification of an ideal vowel?

Although native and nonnative speaker groups agreed on F1 and F2 values of most of the vowels on the perception task, their performance varied on the production task. A surprising result of this study was the extent to which the nonnative speakers were internally consistent in the two tasks, and the extent to which native speakers were not internally consistent.

As mentioned in the discussion of Hypothesis 3 in section 5.2.3 above, the only vowels for which native speakers' pronunciation did not differ significantly from their selection of ideal vowels were the vowels /a/ and /ɔ/. There were statistically significant differences between their perception and production in at least one of the F1 or F2 formants in each of the other vowels in this study. The relationship between the perception and pronunciation of English vowels among native speakers in this study was closer than the same relationship among nonnative speakers.

# 5.5 Areas for future research

This study was a preliminary exploration of the relationship between perception and production in native and nonnative speakers. The study used synthesized speech samples for stimuli, which allowed a great deal of control over the acoustic properties of the stimuli. However, we can't necessarily assume that the participants responded to the synthetic stimuli in the same way that they would respond to natural speech. One way to

verify that would be to compare participants' responses to natural speech to their responses to synthetic speech with the same acoustic values.

This study forced participants to decide on a single, best exemplar of a vowel. However, the variation found in this study suggests that categories are gradient rather than categorical. A comparison of the vowel category boundaries in native and nonnative speakers' vowel systems, similar to the studies of Scholes (1967, 1968) could shed some light on the degrees of variation found in both groups. References:

Al-Banyan, Ahmed Abdullah M. 1996. The accessibility of universal grammar in language acquisition: a cross-linguistic perspective. Unpublished doctoral dissertation: Michigan State University.

Assman P.F. et al. 1982. Vowel identification: orthographic, perceptual, and acoustic aspects *Journal of the Acoustical Society of America* 71, 975-982.

Birdsong, David. 1992. Ultimate attainment in second language acquisition. *Language* 68, 4, 706-755.

Birdsong, David. 1999. Introduction: Whys and why nots of the critical period hypothesis for second language acquisition. In Birdsong (ed) *Second language acquisition and the critical period hypothesis* (pp 1-22). Mahwah, NJ: Lawrence Erlbaum.

Bley-Vroman, Robert. 1989. What is the logical problem of foreign language learning? in Gass, Susan and Jacqulyn Schacter (eds). *Linguistic perspectives on second language acquisition* (pp 41-68). Cambridge: Cambridge University Press.

Boersma, Paul & David Weenink (1999–2000). *Praat, a system for doing phonetics by computer* [Computer program]. Web site: www.praat.org.

Berko, J. and R. Brown. 1960. "Psycholinguistic Research Methods." in Paul M. Mussen, (ed). *Handbook of Research methods in Child Development* (pp 517-557). New York: John Wiley.

Bohn, Ocke-Schwen and James Emil Flege. 1997. Perception and production of a new vowel category by adults second language learners. In Leather, Jonathan and Allan James (ed). *Second-Language Speech: Structure and Process* (pp. 53-74).Berlin: Mouton de Gruyter.

Bosch, Laura and Núria Sebastian-Gallés. Native-language recognition abilities in 4month-old infants from monolingual and bilingual environments. *Cognition* 65, 33-69.

Bradlow, Ann R. 1995. A Comparative Acoustic Study of English and Spanish Vowels. *Journal of the Acoustical Society of America* 97,3, 1916-1924.

Broselow, Ellen. 1984. An investigation of transfer in second language phonology. *International Review of Applied Linguistics* 22, 253-269.

Broselow, Ellen. 1988. An investigation of transfer in second language phonology. *Studies in Descriptive Linguistics* 17, 77-93.

Broselow, Ellen. 1992. Transfer and universals in second language acquisition. In Gass, Susan and Larry Selinker (eds) *Language transfer in language learning* (pp 71-86). Amsterdam: John Benjamins.

Broselow, Ellen and Daniel Finer. 1991. Parameter setting in second language phonology and syntax. Second Language Research 7, 35-59.

Chomsky, Noam. 1964. Current issues in linguistic theory. The Hague: Mouton.

Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge, M.I.T. Press.

Chomsky, Noam. 1986. Knowledge of language: its nature, origin, and use. Westport, CT: Praeger.

Chomsky, Noam. 1995. The minimalist program. Cambridge, MA: The MIT Press.

Chomsky, Noam and Miller, 1963. Introduction to the formal analysis of natural languages in Luce R. Duncan, Robert R. Bush and Eugene Galanter (eds). *Handbook of mathematical psychology* (pp 269-321).New York: Wiley.

Clahsen, Harold. & P. Muysken (1986), The availability of universal grammar to adult and child learners: A study of the acquisition of German word order, *Second Language Research* 2, 93-119.

Clahsen, Harold. & P. Muysken (1989), The UG paradox in L2 acquisition. Second Language Research 5, 1-29.

Coppieters, Ren\_. 1987. Competence differences between native and nonnative speakers. *Language* 63, 544-573.

Delattre, P.C., A.M Liberman, F.S. Cooper. 1955. Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America* 27, 769-773.

Diller, Karl Conrad. 1978. The language teaching controversy. Rowley, MA: Newbury House.

Doucherty, Gerard and Paul Foulkes. 2000. Speaker, speech, and knowledge of sounds in Burton-Roberts, Noel, Philip Carr, and Gerard Docherty (eds) *Phonological knowledge:* conceptual and empirical issues (pp 105-130). Oxford: Oxford University Press.

Eckman, Fred. 1977. Markedness and the contrastive analysis hypothesis. *Language Learning* 27, 315-330.

Eckman, Fred R. 1987. The reduction of word-final consonant clusters in interlanguage. In James, Allen and Jonathan Leather (eds). *Sound patterns in second language acquisition* (pp 143-162). Dordrecht: Foris. Eckman, Fred R. and Gregory K. Iverson. 1994. Pronunciation difficulties in ESL: Coda consonants in English interlanguage. In M. Yavas (ed). *First and second language phonology* (pp 251-265).San Diego: Singular Press.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., and Vigorito, J. (1971). Speech perception in infants. *Science* 171, 303-306.

Ellis, Rod. 1994. *The study of second language acquisition*. Oxford: Oxford University Press.

Fant, Gunthar. 1960. Acoustic theory of speech production. The Hague: Mouton.

Flege, James Emil. 1987. The Production of "New" and "Similar" Phones in a Foreign Language: Evidence for the Effect of Equivalence Classification. *Journal of Phonetics* 15 (1), 47-65.

Flege, James Emil. 1991. Perception and production: the relevance of phonetic input to L2 phonological learning. In T. Huebner and C.A. Ferguson (eds.) *Crosscurrents in Second Language Acquisition* (pp 249-289). Amsterdam: Benjamins.

Flege, James Emil. 1995 Second language speech learning: Theory, findings and problems. in Strange, Winifred. (ed). Speech perception and linguistic experience: Theoretical and methodological issues (pp 233-272). Timonium, MD: York Press.

Flege, James Emil. 1997. English vowel production by Dutch talkers: more evidence for the similar vs new distinction. In Leather, Jonathan and Allan James (ed). *Second-Language Speech: Structure and Process* (pp 11-52). Berlin: Mouton de Gruyter.

Flege, James Emil. 1999. Age of learning and second-language speech. In Birdsong, David P. (ed), Second language acquisition and the critical period hypothesis (101-132). Mahwah, NJ: Erlbaum.

Flege, James Emil and Robert Hillenbrand. 1984. Limits on pronunciation accuracy in adult foreign language speech production. *Journal of the acoustical society of America* 76, 708-721.

Flege, James Emil; Bohn, Ocke-Schwen; Jang, Sunyoung. 1997. Effects of Experience on Non-Native Speakers' Production and Perception of English Vowels. *Journal of Phonetics* 25, 437-470.

Fodor, Jerry A. 1983. *The Modularity of Mind.* Bradford Books. MIT Press, Cambridge, MA.

Foster-Cohen, Susan. 1999. An introduction to child language development. London: Longman.

Frieda, Elaina M, Amanda C. Walley, James E. Flege, Michael E. Sloane. 1999. Adults' Perception of Native and Nonnative Vowels: Implications for the Perceptual Magnet Effect. *Perception and Psychophysics* 61 (3), 561-577.

Fromkin, Victoria and Robert Rodman. 1993. An introduction to language. Fort Worth: Harcourt Brace College Publishers.

Fry, D.B., Arthur S. Abramson, Peter D. Eimas, and Alvin M. Liberman. 1962. The identification and discrimination of synthetic sounds. *Language and Speech* 5, 171-189.

Gass, Susan. 1984. Development of speech perception and speech production abilities in adult second language learners. *Applied Psycholinguistics* 5, 51-74.

Gass, Susan. 1996. Second language acquisition and linguistic theory: the role of language transfer. In Ritchie, William C. and Tej K. Bhatia (eds). *Handbook of second language acquisition* (pp 317-345). San Diego: Academic Press.

Gass, Susan and Larry Selinker. 1994. Second language acquisition: an introductory course. Hillsdale, NJ: Lawrence Erlbaum.

Gerstman, L.H. 1968 Classification of self-normalized vowels. Institute of Electrical and Electronic Engineers, *Transactions on Audio Electroacoustics* 16, 78-80.

Gordon, Peter C., Lisa Keyes and Yiu-fai Yung. 2001. Ability in perceiving nonnative contrasts: Performance on natural and synthetic speech stimuli. *Perception and Psychophysics* 63, 746-758.

Gregg, Kevin 1989. Second language acquisition theory: the case for a generative perspective. In Gass, Susan and Jacqulyn Schacter (eds). *Linguistic perspectives on second language acquisition* (pp 15-40). Cambridge: Cambridge University Press.

Harnad, Stevan. 1987. Psychophysical and cognitive aspects of categorical perception: A critical overview. In Harnad, Stevan (ed.) *Categorical Perception: The Groundwork of Cognition* (pp 1-52). New York: Cambridge University Press.

Herschensohn, Julia Rogers. 2000. The second time around: minimalism and L2 acquisition. Philadelphia, PA: J. Benjamins.

Hillenbrand, James M., Getty, L.A., Clark, M.J., and Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97, 3099-3111.

Hindle, D. 1978. Approaches to Vowel Normalization in the Study of Natural Speech. In: D. Sankoff (ed). *Linguistic Variation: Models and Methods* (pp 161-172). New York. Academic Press.

Hockett, 1955. A manual of phonology. Baltimore, Waverly Press.

Howell, David C. 1995. Fundamental statistics for the behavioral sciences, 3rd edition. Belmont, CA: Duxbury Press.

Ingram, David. 1989. First language acquisition: Method, description, and explanation. Cambridge: Cambridge University Press.

Ingram, David. 1999. Phonological acquisition. In Barrett, Martyn (ed). *The development of language*. Hove, East Sussex, UK: Psychology Press.

Ingram, John C.L. and See-Gyoon Park. 1996. Inter-language vowel perception and production by Korean and Japanese listeners. *Proceedings of the Fourth International Conference on Spoken Language Processing*, Philadelphia, PA.

Iverson, Peter, Patricial K. Kuhl, Reiko Akahane-Yamada, Eugen Diesch, Yoh-ichi Tohkura, Andreas Kettermann, and Claudio Siebert. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87, B47-B57.

Jackendoff, Ray. 1994. Patterns in the mind: Language and human nature. New York: BasicBooks.

Jackobson, Roman. 1968 [1941]. Child Language, Aphasia, and Phonological Universals. The Hague, Paris: Mouton.

James, Allan R. 1988. The acquisition of a second language phonology: a linguistic theory of developing sound structures. Tubingen: G. Narr.

Johnson, J. & Newport, Elizabeth. 1989. Critical period effects in second language learning: the influence of maturational state on the acquisition of ESL. *Cognitive Psychology* 12, 60-99.

Johnson, J. & Newport, Elizabeth. 1991. Critical period effects on universal properties of language: the status of subjacency on in the acquisition of a second language. *Cognition* 39, 215-258.

Johnson, Keith. 2000. Adaptive dispersion in vowel perception. Phonetica 57, 181-188.

Johnson, Keith, E. Flemming, and R. Wright. 1993. The hyperspace effect: Phonetic targets are hyperarticulated. *Language* 69: 505-528.

Joos, Martin. 1948. Acoustic Phonetics. *Language Monographs* 23. Baltimore: Linguistic Society of America.

Jusczyk, Peter. 1986. Speech perception. In Handbook of perception and human performance. New York: John Wiley and Sons.

Katamba, Francis. 1989. An introduction to phonology. London: Longman.

Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Cambridge, MA: Blackwell.

Kenstowicz, Michael. 2003. Salience and Similarity in Loanword Adaptation: a Case Study from Fijian. *ROA-609-0803, Rutgers Optimality Archive*, http://roa.rutgers.edu.

Kluender, K. R., Diehl, R. L., & Killeen, P. R. 1987. Japanese quail can learn phonetic categories. *Science* 237, 1195-1197.

Kluender, K. R., Lotto, A. J. 1994. Effects of first formant onset frequency on [- voice ] judgments result from auditory processes not specific to humans. *Journal of the Acoustical Society of America* 95, 1044-1052.

Kuhl, Patricia K. 1980. Perceptual constancy for speech sound categories in early infancy in Yeni-Komshian, Grace H., James F. Kavanagh, Charles A. Ferguson (eds) *Child Phonology* New York: Academic Press.

Kuhl, Patricia K. 1981. Auditory category formation and development of speech perception. In Stark, Rachel (ed), *Language behavior in infancy and early childhood*. New York: Elsevier North-Holland.

Kuhl, Patricia K. 1987. The special-mechanisms debate in speech research: Categorization tests on animals and infants. in Stevan Harnad, (ed). *Categorical perception: The groundwork of cognition* (pp 355-386). New York: Cambridge University Press.

Kuhl, Patricia K. 1991. Human adults and human infants show a Perceptual magnet effect for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics* 50, 93-107.

Kuhl, Patricia K. and James D. Miller. 1975. Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. *Science* 190, 4209, 69-72.

Kuhl, Patricia K. Karen A. Williams, Francisco Lacerda, Kenneth N. Stevens, and Bjorn Lindblom. 1992. Linguistic Experience Alters Phonetic Perception in Infants by 6 Months of Age. *Science* 255, 606-608.

Kuhl, Patricia K. and Peter Iverson. 1995. Linguistic experience and the Perceptual magnet effect. in Strange, Winnifred (ed). Speech perception and linguistic experience: issues in cross-language research (pp 121-154). Baltimore: York Press.

Labov, William. 1994. Principals of linguistic change: Internal factors. Cambridge, MA: Blackwell.

Labov, William, Malcah Yaeger, and Richard Steiner. 1973. *The quantitative study of sound change in progress*. Philadelphia: U.S. Regional Survey.

Labov, William, Sharon Ash and Charles Boberg. 1997. A National Map of the Regional Dialects of American English. The Linguistics Laboratory, Department of Linguistics, University of Pennsylvania. Retreived November 5, 2003 from http://www.ling.upenn.edu/phono\_atlas/NationalMap/NationalMap.html

Ladefoged, Peter. 1993. A course in phonetics. Orlando: Harcourt Brace & Company.

Ladefoged, Peter. 2001. Vowels and consonants: an introduction to the sounds of languages. Malden, Mass.: Blackwell.

Ladefoged, Peter and Broadbent, Donald. 1957. Information conveyed by vowels. *Journal of the Acoustical Society of America* 29, 98-104.

Lado, Robert. 1957. *Linguistics across cultures*. Ann Arbor: University of Michigan Press.

Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology*, 20, 215-225.

Leather, Jonathan and Allan James. 1996. Second language speech. In Ritchie, William C. and Tej K. Bhatia (eds). *Handbook of second language acquisition* (pp 269-316). San Diego: Academic Press.

Leather, Jonathan. 1997. Interrelation of perceptual and productive learning in the initial acquisition of second-language tone. In Leather, Jonathan and Allan James (ed). *Second-Language Speech: Structure and Process* (pp 75-101). Berlin: Mouton de Gruyter.

Lehiste, I. And D. Meltzer. 1973. Vowel and speaker identification in natural and synthetic speech. *Language and Speech* 16, 356-364.

Lenneberg, E. 1967. Biological foundations of language. New York: Wiley.

Liberman, Alvin. 1970. Some characteristics of perception in the speech mode. *Perception and its Disorders* 48, 238-254.

Liberman, Alvin.M., K. Harris, H. Hoffman, and B. Griffith. 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54, 358-368.

Lin, Yuh-Huey. 2001. Syllable Simplification Strategies: A Stylistic Perspective. *Language Learning* 51, 681-718.

Lisker, L. and A. Abramson. 1964. A cross language study of voicing in initial stops: acoustical measurements. *Word* 20, 384-422.

Lisker, L. and A. Abramson. 1967. Some effects of context on voice onset time in English stops. *Language and Speech* 10, 1-28.

Liu, Dilin and Johanna L. Gleason. 2002. Acquisition of the article the by nonnative speakers of English. *Studies in Second Language Acquisition* 24, 1-33.

Long, Michael. 1990: Maturational constraints on language development. *Studies in Second Language Acquisition* 12, 251-285.

Major, Roy. 1997. A model for interlanguage phonology. In Ioup, Georgette and Steven H. Weinberger (eds) *Interlanguage phonology* (pp 101-124). Cambridge, MA: Newbury.

Major, Roy. 2001. Foreign accent: the ontogeny and phylogeny of second language phonology. Mahwah, NJ: Laurence Erlbaum associates.

Major, Roy and Faudree, Michael C. 1996. Markedness Universals and the Acquisition of Voicing Contrasts by Korean Speakers of English. *Studies in Second Language Acquisition*, 1996, 18, 1, Mar, 69-90

Miller, J.D. 1989. Auditory perceptual interpretation of the vowel. *Journal of the Acoustical Society of America* 85, 2114-2134.

Nearey, T.M. 1989. Static, dynamic, and relational properties in vowel perception. *Journal of the Acoustical Society of America* 85, 2088-2113.

Odlin, Terence. 1989. Language transfer: Cross-linguistics influence on language learning. Cambridge: Cambridge University Press.

Patkowski, Mark. 1990. Age and Accent in a Second Language: A Reply to James Emil Flege. *Applied Linguistics* 11, 173-89.

Peterson, G.E. 1951. The phonetic value of vowels. Language 27, 541

Peterson, G.E. and H.L. Barney. 1952. Control methods used in a study of the vowels Journal of the Acoustical Society of America 24, 175-184.

Pinker, Steven. 1994. The language instinct: How the mind creates language. New York: W. Morrow and Co.

Polka, Linda and Janet F. Werker. 1994. Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20, 2, 421-435.

Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. RuCCS Technical Report 2. Piscataway, NJ: Rutgers Center for Cognitive Science, Rutgers University, and Boulder, CO: Department of Computer Science, University of Colorado.

Repp, B.H. 1984. Categorical perception: Issues, methods, findings. In Lass, N.J., ed. *Speech and language: Advances in basic research and practice vol 10.* New York: Academic Press.

Riney, Timothy and Naoyuki Takagi. 1999. Global foreign accent and Voice onset time among Japanese EFL speakers. *Language Learning* 42, 275-302.

Saussure, Ferdinand de. 1916. A course in general linguistics. Edited by Charles Bally and Albert Reidlinger. Translated from the French by Wade Baskin. New York, Philosophical Library.

Schachter, Jacquelin. 1989. Testing a proposed universal. In Gass, Susan and Jacqueline Schachter (eds). *Linguistic perspectives on second language acquisition* (pp 73-88). Cambridge: Cambridge University Press.

Schachter, Jacquelin. 1996. Maturation and the issue of universal grammar in second language acquisition. In Ritchie, William C. and Tej K. Bhatia (eds). *Handbook of second language acquisition* (pp 159-193). San Diego: Academic Press.

Schwartz, Jean-Luc, Louis-Jean Boe, Nathalie Vallde and Christian Abry. 1997. Major trends in vowel system inventories. *Journal of Phonetics* 25, 233-253.

Scholes, Robert J. 1967a. Phoneme categorization of synthetic vocalic stimuli by speakers of Japanese, Spanish, Persian and American English. *Language and Speech* 10, 1, 46-68.

Scholes, Robert J. 1967b. Categorical responses to synthetic vocalic stimuli by speakers of various languages. *Language and Speech* 10, 252-282.

Scoles, Robert J. 1968. Phonemic interference as a perceptual phenomenon. *Language and Speech* 11, 86-103.

Selinker, Larry. 1972. Interlanguage. International Review of Applied Linguistics 10, 209-231.

Sheldon, Amy and Winifred Strange. 1982. The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede perception. *Applied Psycholinguistics* 3, 243-261.

Silverman, Daniel. 1992. Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology* 9.2:289-328.

Sorace, Antonella. 1996. The use of acceptability judgments in second language acquisition research in Ritchie, William C. and Tej K. Bhatia (eds). *Handbook of second language acquisition* (pp 375-409). San Diego: Academic Press.

Stevens, Kenneth N. 1998. Acoustic phonetics. Cambridge, The MIT Press.

Stockwell, Robert P. and J. Donald Bowen. 1965. The Sounds of English and Spanish. Chicago: Chicago University Press.

Strozer, J.R. 1994. Language acquisition after puberty. Washington, DC: Georgetown University Press.

Takagi, Naoyuki. 2002. The Limits of Training Japanese Listeners to Identify English /r/ and /l/: Eight Case Studies. *The Journal of the Acoustical Society of America* 111, 2887-2896.

Trubetzkoy, Nikoli. S. 1969. *Grundzüge der Phonologie*. Translated by Christiane A. M. Baltaxe. Berkeley: University of California Press.

Vihman, Mary May. 1996. Phonological development: the origins of language in the child. Cambridge, MA: Blackwell.

Wayland, Ratree. 1997. Non-native production of Thai: acoustic measurements and accentedness ratings. *Applied Linguistics*, 18, 3, 345-373.

Weinberger, Steven. 2003. *Foreign accent archive*. Retrieved June 26, 2003 from http:// classweb.gmu.edu/accent/

Werker, Janet F. 1994. Cross-language speech perception: Developmental change does not involve loss. In Goodman, Judith C. and Howard C. Nusbaum (eds). *The development of speech perception: The transition from speech sounds to spoken words* (93-120). Cambridge, MA: MIT Press.

Werker, Janet F. and R.C. Tees. 1984. Cross-language speech perception: Evidence for a perceptual reorganization during the first year of life. *Infant behavior and development* 7, 49-64.

Werker, Janet E. and Chris E. Lalonde. 1988. Cross-Language Speech Perception: Initial Capabilities and Developmental Change. *Developmental Psychology* 24-5, 672-683.
Wexler, Ken. & Manzini, M. 1987. "Parameters and Learnability in Binding Theory", in T. Roeper & E. Williams (eds.), *Parameters and Linguistic Theory* (pp 41-76). Dordrecht, The Netherlands: Reidel.

White, Lydia. 1989. Universal grammar and second language acquisition. Amsterdam: John Benjamins.

White, Lydia. (2000). Second Language Acquisition: From Initial to Final State. in Archibald, J. (ed.). Second Language Acquisition and Linguistic Theory (pp 130-155). Blackwell Publishers: Malden, Massachusetts.

White, Lydia. 2003. Second language acquisition and universal grammar. Cambridge: Cambridge University Press.

White, Lydia and Fred Genesee. 1996. How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research* 12, 233-265.

Willerman, Raquel and Patricia K. Kuhl. 1996. Cross-language speech perception: Swedish, English, and Spanish speakers' perception of front rounded vowels. In H. Timothy Bunnell & William Idsardi (eds) *Proceedings of ICSLP 96* 1: 442-445.

Yamada, Reiko, Yoh-ichi Tohkura and Noriko Kobayashi. 1997. Effect on word familiarity on non-native phoneme perception: identification of English /r/, /l/ and /w/ by native speakers of Japanese. In Leather, Jonathan and Allan James (ed). Second-Language Speech: Structure and Process (pp 103-117). Berlin: Mouton de Gruyter.

Yang, Byonggon. 1996. A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics* 24, 245-261.

