Acoustic Articulatory Evidence for Quantal Vowel Categories:

The Features [low] and [back]

by

Youngsook Jung

B.S., Korea Advanced Institute of Science and Technology 2000
M.S., Korea Advanced Institute of Science and Technology 2002

Submitted to the Harvard-MIT Division of Health Sciences and Technology in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Speech and Hearing Bioscience and Technology

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2009

Author____
Harvard-MIT Division of Health Sciences and Technology
August 21, 2009

Certified by____
Kenneth N. Stevens
Clarence J. LeBel Professor of Electrical Engineering and Computer Science and
Professor of Health Sciences and Technology
Thesis Supervisor

Accepted by____
Ram Sasisekharan, Ph. D.
Edward Hood Taplin Professor of Health Sciences & Technology and Biological Engineering
Director, Harvard-MIT Division of Health Sciences and Technology

This page is intentionally left blank.

Acoustic Articulatory Evidence for Quantal Vowel Categories:

The Features [low] and [back]

By

Youngsook Jung

Submitted to the Harvard-MIT Division of Health Sciences and Technology

on August 31, 2009 in Partial Fulfillment of the

Requirements of the Degree of Doctor of Philosophy

ABSTRACT

In recent years, research in human speech communication suggested that the inventory of sound units that are observed in vowels across languages is strongly influenced by the acoustic properties of the human subglottal system. That is, there is a discrete set of possible vowel features that are constrained by the interaction of the acoustic/articulatory properties of the vowels and a small set of attributes that are observed in the subglottal region. This thesis tests the hypothesis that subglottal resonances govern vowel feature boundaries for three populations: adult speakers of English; adult speakers of Korean; and children learning English.

First, we explored the relations among *F*1 of vowels, the first subglottal resonances (*SubF*1) and the feature [low] in English. For the diphthong [ɑɪ], *F*1 peaks for vowels showed an acoustic irregularity near the speaker's *SubF*1. For monophthongs, analysis of *F*1 frequency distributions shows a boundary between [+low] and [-low] vowels at the speakers' *SubF*1. Second, we studied the relations among *F*2 of Korean vowels, *SubF*2 and the feature [back], to test whether the relation between subglottal resonances and the feature boundary, demonstrated earlier for English, also can be applied to other languages. Results show that the *F*2 boundary between [back] and [front] vowels was placed near *SubF*2 in Korean, as in English. Third, we explored the development of vowel formants in relation to subglottal resonances for 10 children in the age range of 2;6–3;9 years using the database of Imbrie (2005). Results show that at the earlier ages, formant values deviated from the expected relations, but during the six month period in which the measurements were made, there was considerable movement toward the expected values. The transition to the expected relations appeared to occur by the age of 3 years for most of these children, in a developmental pattern that was inconsistent with an account in terms of simple anatomical increase.

These three sets of observations provide evidence that subglottal resonances play a role in defining vowel feature boundaries, as predicted by Stevens' (1972) hypothesis that contrastive phonological features in human languages have arisen from quantal discontinuities in articulatory-acoustic space.

Thesis Supervisor: Kenneth N. Stevens
Title: Clarence J. LeBel Professor of Electrical Engineering and Professor of Health Sciences and Technology

This page is intentionally left blank.

Acknowledgements

I thank my advisor Ken Stevens who has been an ideal teacher, researcher and mentor, and also my spiritual father. He has been a role model to me. If I become a research advisor of someone in future, I want to be like him. I could not have had a more wonderful advisor. He truly has been blessings to my life. I thank Stefanie Shattuck-Hufnagel who is a woman of encouragement. Several times with her helps, I could make a big jump and breakthrough. She has provided me a guidance of how to combine ideas into one big picture. I thank a thesis committee member Adam Albright for his brilliant feedback on my research from different perspectives I never have thought. Also, I thank Jim Glass for sparing his precious time to serve as my oral exam committee member.

I thank faculty members of HST Speech and Hearing program for their support and patience, for example, by giving me a chance to take a written qualifying exam even when I was sick. Especially I thank my academic advisor John Guinan, Louis Braida, Bertrand Delgutte, Jennifer Melcher and Joseph Perkell. Without their helps, I would not have finished my PhD study.

I thank the Speech Communication group people – Arlene, Chi-youn, Jae Young, Nancy, Elisabeth, Yoko, Xuemin, Steven, Shanqing, Satrajit, Miwako and Sherry. They are simply the best. Especially I thank Chi-youn, Jae Young, Nancy and Steven for discussions and feedbacks on my research. Also, I thank friends of Speech and Hearing program – especially to Amanda, Courtenay, Lauryn, Steven and Chris.

I thank my church friends and members of "Before the Throne" for their prayers and friendship. I especially thank Jinsook, Boeun, Jiyoun, Eunkyoung, Sunghoon and Younkyung.

I thank my Dad, Mom, brother Seil, sister Seyoung and boyfriend Woojin for their unfailing love toward me. Through their prayers and encouragement, I could have survived at MIT.

Finally, I thank my Heavenly Father who has been my faithful friend and helper during my PhD period. Everything has been always enough with Him and in Him.

This research has been supported by an NIH Grant R01-DC00075 and Korean Government Scholarship for Future IT Leader.

This page is intentionally left blank.

Contents

List of Figures

List of Tables

This page is intentionally left blank.

Chapter 1. Introduction

1.1 Overview

Speech can be analyzed as a sequence of discrete sounds. Phonologically distinctive features were introduced to describe the attributes of discrete sound units and to categorize the sound segments (Jakobson et al., 1952; Chomsky and Halle, 1968). Binary features are assumed and assigned to the sound segment. Concerning binary distinctive features, efforts have been made for several decades to connect phonetics and phonology. Quantal theory, suggested by Stevens (1989), might provide good explanations for phonologically binary features drawn from non-linear relationships between articulatory parameters and acoustic parameters in the human speech production system.

Following ideas from quantal theory, this thesis provides evidence for binary feature aspects of speech sounds in acoustic production of vowels, in relation to subglottal resonances, in three different directions: (1) the quantal nature in $F1$ frequencies of vowels, relative to the feature [low], (2) the quantal nature in Korean, which has different vowel inventories from English, and (3) the quantal development in young children's vowels. The third investigation is the primary study of this thesis research. The backgrounds, objectives and significances of this thesis study are provided in Chapter 1. We explore the quantal nature of vowels in relation to the first subglottal resonance and the universal vowel feature [low] in Chapter 2. We explore the quantal nature of Korean vowels in relation to subglottal resonances and vowel features, as the first step toward across-language studies in Chapter 3. We study the quantal development of vowels in young children in Chapter 4. Conclusions and future directions are addressed in Chapter 5.

1.2 General Background

1.2.1 Quantal Theory and Distinctive Features

In quantal theory, there are two discrete stable states, where the acoustic output in the sound production system is relatively insensitive to articulatory parameter changes, as shown in Figure 1.1. In the transition state, the acoustic output greatly varies with the articulator movement, and the placement of the articulator is hypothesized to avoid this unstable region. Therefore, the sound output tends to be quantally represented. One of the quantal stable states is matched to the [+feature], and another state corresponds to the [–feature] (Halle and Stevens, 1991).

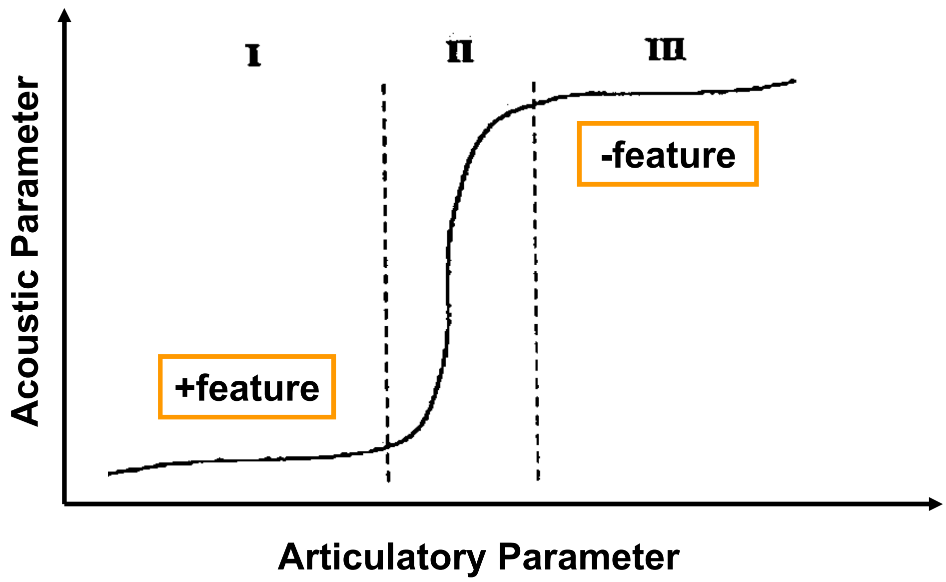


Figure 1.1 Non-linear relationship between articulatory parameters and acoustic parameters in quantal theory. Region I and III: stable states, Region II: transition state. One of the quantal stable states is matched to the [+ feature], and another side corresponds to the [–feature]. (Modified from Stevens, 1989.)

This theory implies that there may be disfavored places for the articulators where the acoustic result relatively sensitive to the articulatory configuration. The acoustic sound output may show discontinuities when the articulator moves from one stable state to another stable state when it crosses the unstable region (Halle and Stevens, 1991).

One of the examples of the evidence for quantal theory is found in the feature [anterior]. For phonation of a fricative consonant /s/ ([+anterior]), the vocal tract is excited primarily in the frequency ranges of $F4$ and $F5$, while it is excited primarily in the range of $F3$ for the consonant /š/ ([-anterior]). The tongue blade tends to be placed at a certain distance from the lips for the production of these consonants, because of acoustic coupling between the front of the vocal tract and back cavities which are formed by the constriction of the tongue blade in the vocal tract. This coupling effect may bring unstable regions at certain frequencies, and therefore, the tongue blade is placed to avoid the unstable regions (Stevens, 2003).

1.2.2 Vowel Features and Subglottal Coupling

Vowel Features

The vowel sounds can be modeled as the output from the vocal tract filter with the source due to the vocal fold vibration. The shape of the vocal tract varies with the placement of the articulators, which changes the characteristics of the filter. The specific position of the tongue body and lips in the mouth determines the shaping of the formants in vowels.

Linguists have divided vowels into several categories according to the features they have. The primary articulator for producing vowels is the tongue body. The articulator-bound features, [high], [low], and [back] are assigned for the tongue body. The features [high] and [low] are related to the height of the tongue body, whereas the vowel feature [back] is related to the backness of the tongue body in the mouth during the phonation. These feature values for [low] and [back] are the focuses of this thesis. The set of English vowels are listed in Table 1.1. Interestingly, the vowel categories are very similar across languages and the vowel features are hypothesized to be universal.

Table 1.1 Feature values for English vowels (from Stevens, 1998)

Feature	**i**	**ɪ**	**e**	**ɛ**	**æ**	**ɑ**	**ɔ**	**o**	**ʌ**	**ʊ**	**u**	**ɚ**
Low	–	–	–	–	+	+	+	–	–	–	–	–
Back	–	–	–	–	–	+	+	+	+	+	+	–
High	+	+	–	–	–	–	–	–	–	+	+	–

Subglottal Coupling

In the modeling of vowel production, only the vocal tract above the larynx is usually taken into consideration because the glottal impedance is assumed to be very large comparing to the impedance of the subglottal system (Hanson and Stevens, 1995; Hunt, 2009). The glottal impedance is infinite when the glottal opening is zero. However, in reality because the larynx does not close completely during phonation, there is acoustic coupling between the vocal tract and the subglottal system which consists of the trachea, bronchi and lungs. The resonances of the vocal tract, usually called formant frequencies, vary with the vocal tract shape. On the other hand, the resonances of the subglottal system are assumed to be constant for a speaker during the phonation of speech. The reported values of the first two subglottal resonances are about 600, 1550 Hz for male speakers and 700, 1650 Hz for female speakers, respectively (Fant et al., 1972; Ishizaka et al., 1976; Cranen and Boves, 1987).

The equivalent circuit model of the vocal tract coupled to the subglottal system is shown in Figure 1.2(a) (Hanson and Stevens, 1995). There is a volume velocity source at the glottis *Uo* during phonation of vowels. The same amount of airflow through the glottis assumes to pass through the vocal tract and goes to the vocal tract and to the trachea. *Um* is the volume velocity at the mouth. *Zsg* is the impedance into the subglottal system at the glottis and *Zvt* is the impedance into the vocal tract.

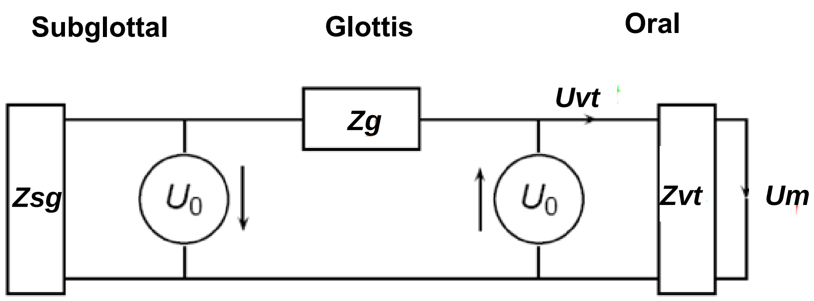


Figure 1.2(a) Equivalent circuit model of the vocal tract coupled to the subglottal system. *Zsg* is the impedance of the subglottal system, *Zg* is the glottal impedance, *Zvt* is the impedance of the vocal tract, *Uo* is the volume velocity source, *Uvt* is the volume velocity at the vocal tract, and *Um* is the volume velocity at the mouth (Modified from Hanson and Stevens, 1995).

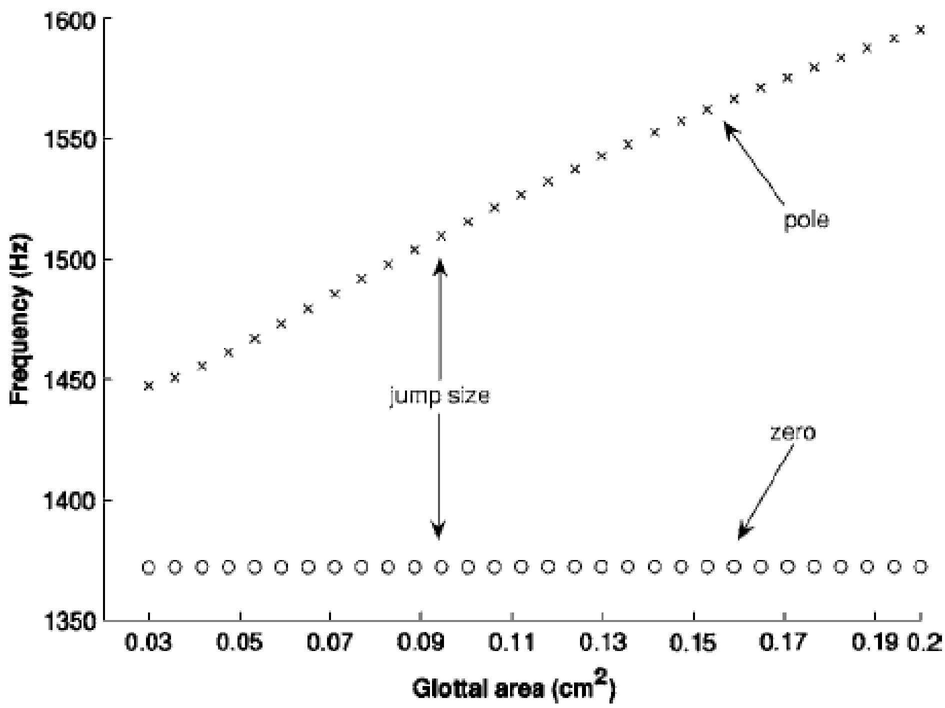


Figure 1.2(b) Simulated frequencies of the pole-zero pair due to subglottal coupling near the second subglottal resonance. Glottal area increased from 0.03 to 0.2 cm2, and a pole-zero pair was separated (Chi and Sonderegger, 2007).

When *Zg* is infinite (closed glottis), the transfer function of *Uvt/Uo* defined as the following equation is equal to one (no subglottal coupling).

$$\frac{U_{vt}}{U_o} = \frac{Z_g}{Z_g + Z_{vt} + Z_{sg}}$$, where *Uvt* is the volume velocity at the vocal tract.

If *Zg* is finite, the subglottal coupling is not negligible. For example, at the frequencies of the subglottal resonances, the vocal tract transfer function becomes zeros because *Zsg* becomes a maximum. The resonances of the subglottal system introduce additional pole-zero pairs to the vocal tract transfer function. When the glottis is closed, the pole-zero pairs are cancelled out. If the glottis is open, the locations of pole-zeros are separated (see Figure 1.2(b), Chi and Sonderegger, 2007). The poles which are the natural frequencies of the subglottal system remain at the same frequencies. On the other hand, the zeros are shifted to higher frequencies and the amount of the shift is depending on the size of glottal openings. The poles of the subglottal system appear as zeros in the transfer function of the coupled system (vocal tract and subglottal system together), whereas the zeros of the subglottal system appear as poles in the transfer function (Fant et al., 1972; Stevens, 1998). These pole-zero pairs bring additional prominences or attenuations into the formant spectra near the subglottal resonances. Therefore, acoustically irregular regions where formant amplitude attenuations or formant frequency jumps appear are formed near the subglottal resonances.

Quantal theory hypothesizes that the acoustically irregular region or unstable region arising from subglottal coupling may construct the vowel feature boundaries. Therefore, constrictions by the tongue body might be made to avoid the unstable region during the phonation of vowels.

1.3 Objectives and Significance

In recent years, research in human speech communication has suggested that the inventory of sound units that are observed in vowels across languages is strongly influenced by the acoustic properties of the human subglottal system (Stevens, 1998; Chi and Sonderegger, 2004, 2007; Lulich et al., 2007; Lulich 2009). That is, there is a discrete or quantal set of possible vowel features that are constrained by the interaction of the acoustic/articulatory properties of the vowels and a small set of attributes that are observed in the subglottal region below the vocal tract. In this thesis, the hypothesis that subglottal resonances govern feature boundaries was tested, by exploring the relations among subglottal resonances, vowel formants and features for three populations: adult speakers of American English (Aim #1) ; adult speakers of Korean (Aim #2) ; and children learning American English (Aim #3).

1.3.1 Aim #1

Stevens first proposed that vowels seem to be divided into [+low] vowels and [-low] vowels according to the first subglottal resonance, and into [+back] vowels and [-back] vowels according to the second subglottal resonance (Stevens, 1998). Recent research has suggested that the subglottal system plays significant roles in speech production and perception by constraining the frequency band of the speech output. Evidence of quantal relation among the second formant frequencies ($F2$), the second subglottal resonances ($SubF2$) and the feature [back] has been provided (Chi and Sonderegger, 2007; Lulich et al., 2007, Lulich, 2009).

Chi and Sonderegger predicted subglottal coupling effects such as frequency jumps and amplitude attenuations on $F2$ due to $SubF2$ by using a model with a simplified subglottal system (Chi and Sonderegger, 2007). They found acoustic irregularities such as frequency jumps or amplitude attenuations in the formant trajectories of $F2$ near $SubF2$ for adult English speakers. In addition, they found that for each speaker, the second subglottal resonance is located between the $F2$ frequencies of back vowels and those of front vowels (Chi and Sonderegger, 2004). Sonderegger collected $F2$ frequency data from various languages and showed that the boundary of $F2$ frequency distributions between [+back] vowels and [-back] vowels is close to the mean values of *SubF*2, which were obtained from experiments. Lulich and colleagues carried out perception experiments by changing the locations of amplitude attenuation, characteristics of *SubF*2, and the results showed that the perception boundary for the backness of vowels changes depending on the modified locations of *SubF*2 (Lulich et al., 2007).

The results from the studies about *SubF*2 and the feature [back] have raised a possibility that *SubF*1 may play a role in defining another feature in vowels – the feature [low], as suggested by Stevens (1998). The quantal hypothesis for the feature [low], however, has not been tested in acoustic data.

One motivation for this study is to verify that it is not a coincidence that a second subglottal resonance happens to lie near the frequency boundary between back vowels and front vowels. If these observations also apply to the first subglottal resonance and the feature [low], we may provide stronger evidence for the quantal hypothesis.

Objective

The objective of the first part in this thesis research is to explore the role of the first subglottal resonance (*SubF*1) in defining the feature contrast of [low]. The hypothesis of this study is that *SubF*1 plays a role in defining the feature contrast of [low], similar to the way *SubF*2 plays a role in defining the feature contrast of [back]. Predictions based on this hypothesis are (1) acoustic irregularities will be observed near speaker's *SubF*1 in *F*1 trajectories of vowels, (2) The boundary between *F*1 frequencies of [+low] vowels and those of [–low] vowels will be placed at about speaker's *SubF*1, and therefore, the vowels will follow the quantal relation between *F*1 and *SubF*1, as shown in Figure 1.3. It was hypothesized that the first subglottal resonance (*SubF*1) lies between [+low] and [-low] vowels, while the second subglottal resonance (*SubF*2) forms a boundary between [+back] and [-back] vowels. There are many other vowel features, which are not shown in this figure, such as [high] and [tense]. For example, the *F*1 boundary between [+high] vowels and [-high] vowels is expected to lie between /i, u/ and /e, o/ in the figure.

In this part, the following questions are addressed. (i) How are the acoustic characteristics of $F1$ formant frequency and amplitude near $SubF1$ for English vowels? (ii) How do $F1$ frequencies relate to the speaker's $SubF1$ for English? (iii) How are $F1$ distributions of [+low] vowels vs. [-low] vowels from various languages?

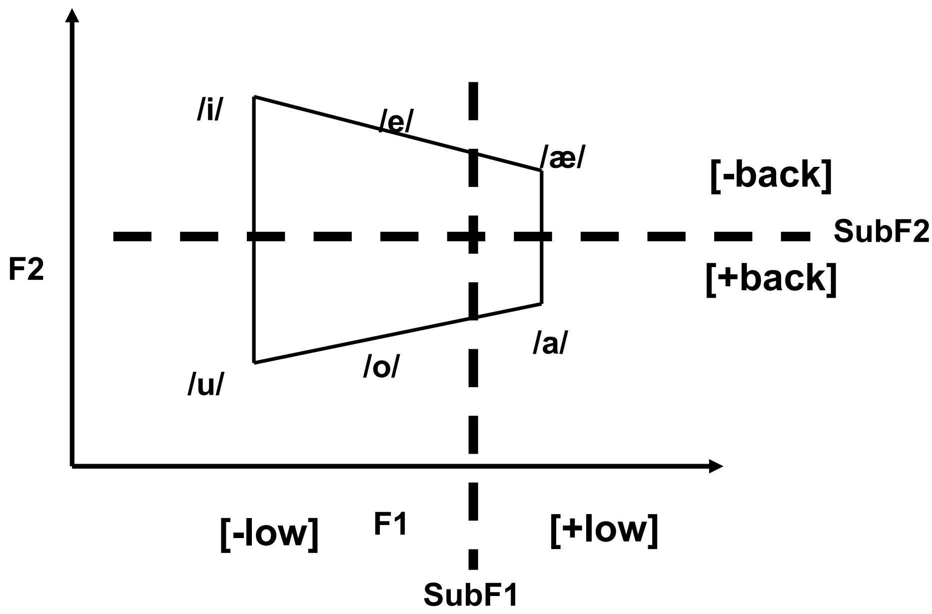


Figure 1.3 The quantal relations among vowel formants, subglottal resonances and vowel features. It was hypothesized that the first subglottal resonance (*SubF*1) lies between [+low] and [low] vowels, while the second subglottal resonance (*SubF*2) forms a boundary between [+back] and [-back] vowels. There are many other vowel features, which are not shown in this figure, such as [high] and [tense].

1.3.2 Aim #2

As discussed in the previous section, research has suggested that the second subglottal resonance plays a role in defining the vowel feature [back] for English. These subglottal quantal effects on speech signals are theoretically expected to be universal and independent of language since the subglottal coupling is related to physics of speech production system. However, there have been few acoustic studies of the relations between subglottal resonances and vowels in other languages apart from English. For several Spanish child speakers, frequency discontinuities near *SubF*2 were found in Spanish front/back diphthongs (Wang et al, 2008). Quantal relations between formants and subglottal resonances were examined in German and Swabian (Madsack et al, 2008). The acoustics of subglottal coupling effects on speech spectra have not been examined in detail.

Vowel inventories are different for different languages. There are languages that have only one low vowel, so that there is no contrast of [+back] and [-back] for the low vowel (e.g., German, Mandarin Chinese, Japanese and Korean). The results of the quantal hypothesis test on German show that *F*2 values for the low back vowel /a/ have large variability depending on the speaker in relation to *SubF*2 (Madsack et al, 2008). Except for /a/, other vowels generally follow quantal relations. Large variability in *F*2 of the vowel /a/ was reported depending on the following nasal consonants in Mandarin Chinese (Mou, 2006). If the vowel /a/ is followed by an alveolar nasal coda /n/, *F*2 of the vowel is fronted, whereas *F*2 is low if it is followed by a velar nasal coda /ŋ/.

The question arises whether in such languages the vowel /a/ is made in a way that avoids the second subglottal resonance. Possible views are: (1) the low vowel is always [+back], *SubF*2 is avoided for *F*2 of the low vowel, or (2) *SubF*2 is avoided but the low vowel is front or back depending on the adjacent consonant. We test which of these hypotheses is correct for the low vowels in Korean which has only one low vowel.

Korean Vowels and Consonants

Korean has nine monophthong vowels, /i, e, ɛ, ø, ɨ, a, ʌ, o, u/, and their [back] features are given in Table 1.2. The vowel /ɨ/ can be replaced by the back vowel /ɯ/. The Korean vowel /ɨ/ is categorized into a central or front vowel. The vowel /a/ is categorized into a central or back vowel (Ahn, 1998; Ahn, 2006). Thus, for these vowels, the feature of [back] is not specified in Table 1.2. The front /ø/ can be removed from this list when it is regarded as a diphthong, and /e/ and /ɛ/ can be merged into /e/ vowel (Yang 1996; Ahn, 1998; Ahn and Iverson, 2005). In this thesis research, however, we follow nine vowel systems.

Table 1.2 Korean vowel IPA symbols and [back] value. '*' represents 'not specified'.

Vowel	[back]
/i/	–
/e/	–
/ø/	–
/ɛ/	–
/ɨ/	*
/a/	*
/ʌ/	+
/o/	+
/u/	+

Korean consonants are broadly divided into three groups depending on the place of constrictions: labial (/m, p, p^h/), velar (/ŋ, k, k^h/), and alveolar (/n, t^h, t, s/) or post-alveolar (/tʃ, $tʃ^h$/). The primary articulator is the lips for labial consonants, the tongue body for velar consonants, and the tongue blade for alveolar or post-alveolar consonants. The feature values concerning primary articulators, place of constrictions, nasality for Korean consonants are summarized in Table 1.3.

Table 1.3 IPA symbols and feature values for Korean consonants.

Feature	m	p	p^h	ŋ	k	k^h	n	t^h	t	s	tʃ	tʃ^h
Lips	+	+	+									
Tongue blade							+	+	+	+	+	+
Tongue body				+	+	+						
Anterior								+	+	+		
High				+	+	+						
Low				–	–	–						
Back				+	+	+						
Nasal	+			+			+					

Objective

The objective of the second part in this thesis research is to explore the role of the second subglottal resonance in defining the feature contrast of [back] in Korean which has different vowel inventory from English. This is a preliminary cross language study of the role of *SubF*2 in defining quantal feature values in different vowel systems.

The hypothesis is that *SubF*2 plays a role in defining feature contrast of [back] in other languages as well as in English. Predictions based on this hypothesis are (1) acoustic irregularity will be observed near speaker's *SubF*2 in Korean, (2) *F*2 of front vowels will be placed above speakers' *SubF*2, while *F*2 of back vowels will be placed below *SubF*2 in Korean.

In this part, the following questions are addressed. (i) How are the acoustic characteristics of *F*2 formant frequency and amplitude near *SubF*2 for Korean vowels? (ii) How do *F*2 frequencies relate to the speaker's *SubF*2 for Korean? (iii) How do *F*2 frequency shift of the vowel /a/ due to adjacent consonants relate to *SubF*2 for Korean?

1.3.3 Aim #3

As discussed in the previous section, for adult speakers, the quantal relation between *F*2 and *SubF*2 was observed (Chi and Sonderegger, 2004; Madsack et al, 2008). For several children under 16 year olds (only one child younger than 6 years), the quantal relation between *F*2 and *SubF*2 was observed (Lulich, 2009). Previous observations that older children and adults already show the quantal relations raises the question when younger children begin to follow the quantal relation.

Concerning the development of vowel production in children, much research has been performed based on the analysis of changes in the vowel formant frequencies with age. (Peterson and Barney, 1952; Eguchi and Hirsh, 1969; Kent and Murray, 1982; Pentz and Gilbert, 1983; Childers and Wu, 1991, Lee et al.,1999) Studies on the development of children's speech have been mainly focused on three areas: (1) anatomical development, (2) articulation skill development and (3) phonological development. Previous studies based on the acoustic data of vowel formant frequencies have shed light on how the anatomic development of articulatory organs relate to formant changes with age, and how children's articulatory abilities develop with age. More direct studies of anatomical development of children's articulation have been performed using technologies such as MRI. The dimensions of several articulation organs (e.g., vocal tract, tongue etc.) have been examined from birth to adulthood (Fitch and Giedd, 1999; Vorperian et al., 2005). The results show that the vocal tract growth rate in the first 2 years is faster than other periods of life. Beyond 2 years old, the vocal tract uniformly grows with relatively slow speed. In addition, there is acoustical evidence that vowel categories are largely in place until 24 months (Fitch and Giedd, 1999; Ishizuka et al., 2007). Ishizuka and colleagues concluded that the formant frequencies of vowels might be dominantly determined by the effect of the anatomical development of articulatory organs in the first 24 months; after 24 months, fine tuning remains, and the patterns of changes in the frequencies are different for each speaker. However, there have been few detailed acoustic studies on how fine tuning is developed.

Vorperian and Kent (2007) summarized the significant results of various studies which have been done by several researchers concerning acoustics of children's vowels, especially vowel space. The results were summarized in relating to the anatomical changes in three parts: supralaryngeal, laryngeal and velopharyngeal systems. However, there has been very little work on the characteristics of anatomy and acoustics of subglottal system of children.

Objective

The objective of the third part of this thesis research is to explore the roles of subglottal resonances in defining the vowel features for young children in ages 2-3 years. The hypothesis is that adult-like quantal relations between subglottal resonances and formant frequencies in the vowels will be developed in this age range.

In this part, the following questions are addressed. (i) How do the first and second subglottal resonances change in the age range of 2-3 years? (ii) How do $F1$ and $F2$ change relative to the subglottal resonances in this age range for each child? (iii) Do quantal relations better predict $F1$ and $F2$ changes than vocal tract length?

1.3.4 Significance

As far as we know, the first part of this thesis research is the first acoustic study about subglottal coupling effects on $F1$ formant peaks and a quantal relation between $F1$ and $SubF1$. The second part is the first detailed acoustic study about subglottal coupling effects of $SubF2$ on $F2$ formant beside English. The third part is the first longitudinal study on the vowel production in relation to the subglottal resonances in the ages of 2-3 year olds. Preliminary results related to this thesis research were presented at the conferences of the Acoustical Society of America (Jung and Stevens, 2007; Jung et al., 2009, Jung, 2009).

The finding of these studies can be applied to clinical areas such as speech testing for maturity and/or disorders, and speech therapy. Second, the finding of this study can give insights into why natural vowel categories exist in the world and into modeling of sound acquisition in young children. Third, the finding can give us more understanding of the subglottal system, because little acoustic data is currently available for the subglottal system. Last, subglottal resonances can be used in automatic speech recognition for speaker normalization. The detailed applications are explained in Chapter 5.

Chapter 2. First Subglottal Resonance and the Feature [low]

2.1 Introduction

In this chapter, we explore the role of the first subglottal resonance (*SubF*1) in defining the feature contrast of [low]. Previous research suggested that the second subglottal resonance (*SubF*2) defines the vowel feature contrast of [back] for English (Chi and Sonderegger, 2007; Lulich et al. 2007, Lulich 2009). The results from study on *SubF*2 and the feature [back] raise the possibility that *SubF*1 may play a role in defining another feature contrast in vowels.

The hypothesis of this study is that *SubF*1 plays a role in defining the feature contrast of [low], similar to the role played by *SubF*2 in defining the feature contrast of [back]. Predictions based on this hypothesis are (1) acoustic irregularities will be observed near a speaker's *SubF*1 in *F*1 trajectories of vowels, (2) The boundary between *F*1 frequencies of [+low] vowels and those of [–low] vowels will be placed near the speaker's *SubF*1.

Testing this hypothesis is achieved by two experiments. In the first experiment, using Chi's database (Chi and Sonderegger, 2007), we measured *SubF*1 and *F*1 for several adult English speakers. We examined whether acoustic irregularities such as amplitude attenuations and frequency discontinuities are observed near speaker's *SubF*1 in *F*1 trajectories of vowels, and whether the boundary between *F*1 frequencies of [+low] vowels and those of [–low] vowels agree with speaker's *SubF*1. In the second experiment, *F*1 frequencies from several languages were collected. And we examined how the boundary between *F*1 frequencies of [+low] vowels and those of [–low] vowels across languages, relates to *SubF*1 values obtained from the first experiment and reported in the literature.

2.2 Experiment #1: English

2.2.1 Method

Database

The database we used for this study was collected and reported by Chi and Sonderegger (2004, 2007). The recordings of speech and subglottal signals were made from nine adult (three males and six females) speakers of English, who were students or affiliates at MIT. They are denoted as M1–M3 for male speakers and F1–F6 for female speakers. All subjects except for one subject (a male speaker M1, who is a speaker of Canadian English) are American English speakers. To measure the subglottal resonances of a speaker, an accelerometer was attached to the neck of the speaker, and subglottal signals were obtained using the accelerometer, following the method demonstrated by Cheyne (2002). The sentence, "hVd, say hVd, again." was produced, where V is an English monophthong or a diphthong vowel. The vowel list with the feature values for [low], and carrier words is shown in Table 2.1.

Acoustic Measurements

The xkl speech analysis tool was used for acoustic measurements for this experiment.

A. First Subglottal Resonance

From the subglottal signals which were captured from an accelerometer, DFT spectra using a one-pitch period length of Hamming window were obtained in the steady regions of monophthong vowels for each speaker. From the spectra, the first formant peaks were manually measured for the first subglottal resonances.

B. Monophthong Analysis

From the speech signals, the first formant peaks were measured using the 12^{th} order LPC with a 25.6 ms Hamming window in the middle of monophthong vowels for each speaker.

Table 2.1 IPA symbols of English vowels, feature values for [low] and carrier words.

IPA	Feature [low]	Word
/æ/	+	“had”
/ɔ/	+	“hawed”
/ɛ/	−	“head”
/ɚ/	−	“heard”
/i/	−	“heed”
/ɪ/	−	“hid”
/ɑ/	+	“hodd”
/o/	−	“hoed”
/ʊ/	−	“hood”
/ʌ/	−	“hud”
/u/	−	“who’d”
/e/	−	“hade”
/ɑɪ/	+ / −	“hide”
/oɪ/	− / −	“hoid”
/ɑʊ/	+ / −	“how’d”
/ju/	− / −	“hued”

C. Diphthong analysis

The time courses of the frequencies ($F1$) and the amplitudes (A1) for the first formant peak were measured from diphthong vowels. A DFT spectrum was obtained using a window with one pitch period length. The analysis window was shifted by one pitch period. From the spectra, the first formant frequencies and amplitudes were manually measured.

2. 2. 2 Results

First Subglottal Resonances

For nine speakers (three male and six female), the first subglottal resonance values were obtained. The mean values and the standard deviations of the frequencies of *SubF*1 are summarized in Table 2.2. The range of the means of *SubF*1 frequencies was 530–550 Hz for male speakers and 610–690 Hz for female speakers. The ranges of *SubF*1 are similar but slightly lower than directly measured *SubF*1 values which were about 600 Hz for adult male speakers and about 700 Hz for adult female speakers (Ishizaka et al., 1976; Cranen and Boves, 1987).

Table 2.2 The mean and standard deviation (in parenthesis) values of *SubF*1. The measurements were obtained from 11 tokens (11 monophthongs) per speaker.

Male		**Female**	
Speaker	***SubF*1 (Hz)**	**Speaker**	***SubF*1 (Hz)**
M1	547(38)	**F1**	648(50)
M2	535(48)	**F2**	664(57)
M3	540(44)	**F3**	645(61)
		F4	610(60)
		F5	650(52)
		F6	690(59)
Mean	541	**Mean**	651

Monophthong Analysis

A. Distribution of *F*1 in Relation of *SubF*1

For eight speakers (M1, M3, F1–F6) of English, the first formant frequencies were measured and compared to the speaker's *SubF*1. Figure 2.1 shows *F*1 frequency plots along with their *SubF*1 values for female speakers F5 and F6, showing conforming to predictions. Each measurement point was obtained by averaging *F*1 values across four tokens for each vowel. Horizontal lines represent speakers' mean *SubF*1 values from Table 2.2. For these two speakers, for English [+low] vowels, their *F*1 frequencies were above speakers' *SubF*1, while *F*1 frequencies were below speaker's *SubF*1 for [–low] vowels. This relation between *F*1 and *SubF*1 was almost always observed for other speakers. Some variation of *F*1 relative to *SubF*1 was found in the vowel /ɛ/ depending on speakers (For other speakers, see Appendix A). Quantitative analysis on this variation is given in the following section.

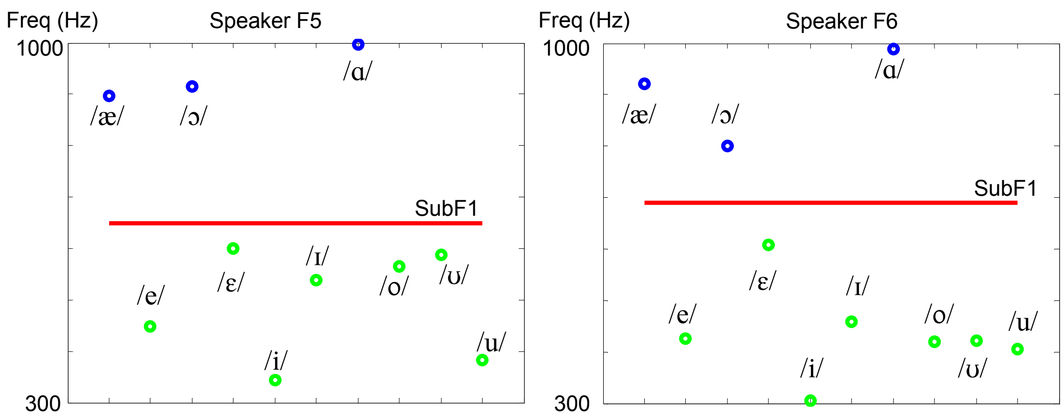


Figure 2.1 *F*1 frequencies of English monophthong vowels for female speakers F5 (left) and F6 (right). Each point represents the averaged *F*1 value from four tokens. Horizontal lines represent *SubF*1 values from Table 2.2, which were 650 Hz for the speaker F5 and 690 Hz for the speaker F6. Blue dots: [+low] vowels; Green dots: [-low] vowels.

The distributions of ($F1$ – $SubF1$) across [+low] vowels and across [-low] vowels are given in Figure 2.2 for each speaker. The mean values of ($F1$ – $SubF1$) across [+low] vowels were always positive ($p < 0.05$), whereas those of ($F1$ – $SubF1$) across [–low] vowels were always negative for each individuals ($p < 0.05$). This result implies that the distribution of $F1$ frequencies is quantally represented according to speakers' $SubF1$.

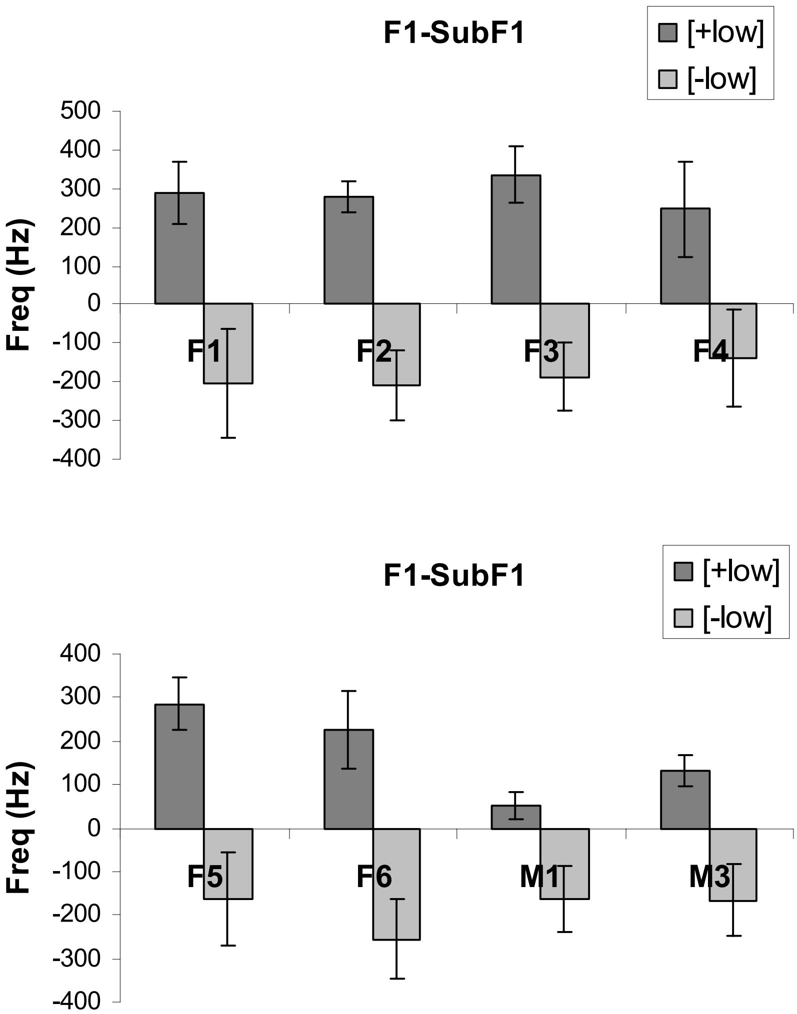


Figure 2.2 Distributions of ($F1$ – $SubF1$) values across [+low] vowels and across **[**–low] vowels for eight speakers of English (Upper panel: subjects F1–F4, lower panel: subjects F5, F6, M1, M3).

B. Percentage Following Quantal Relation

The percentage of tokens following the quantal relation between *SubF*1 and *F*1 was calculated for eight speakers (M1, M3, F1–F6) of English, and the result is summarized in Table 2.3. For English [+low] vowels, all of /ɑ/ and /æ/ tokens and most of the /ɔ/ (expect only one case) tokens, their *F*1 frequencies were above speakers' *SubF*1. In contrast, for [–low] vowels /i, ɪ, o, ʊ, u, e/ except for /ɛ/, *F*1 frequencies were always below speakers' *SubF*1. *F*1 frequencies of the vowel /ɛ/ were often higher than speakers' *SubF*1 depending on speakers: half of the speakers (four out of eight speakers) made *F*1 of /ɛ/ vowels above *SubF*1.

Table 2.3 The percentage of tokens following the quantal relation between *SubF*1 and *F*1 for eight adult speakers of English. The total number of tokens was 292.

[+low]		[–low]	
	*F*1 – *SubF*1 > 0		*F*1 – *SubF*1 < 0
/ɑ , æ, ɔ /	99.8%	/i, ɪ, o, ʊ, u, e/	100%
		/ɛ/	65.2%
		Total (/ɛ/ included)	96.1%
	99.8%	Total (/ɛ/ excluded)	100%

Diphthong Analysis

For six speakers of English (M1–M3, F1–F3), the time courses of the frequencies (*F*1) and amplitude (A1) for the first formant peak were obtained from the diphthong vowel /ɑɪ/. In this diphthong, the feature for [low] changes from [+low] to [–low]. The example of the time course of *F*1 and A1 is given in Figure 2.3. During the phonation of /ɑɪ/, *F*1 frequency decreases as the tongue body moves upward. At the time of about 1260 ms, *F*1 passes through the speaker's *SubF*1. At the same time, two events occur simultaneously:

(1) there is a frequency jump in $F1$ formant frequency and (2) $F1$ amplitude becomes a local minimum. Similar observations were reported related to the second subglottal resonance (Chi and Sonderegger, 2004, 2007). In the diphthongs, where the feature changes from [+back] to [–back], frequency discontinuities or amplitude attenuations in $F2$ trajectories were observed when $F2$ passed through the speaker's $SubF2$. These results show that the acoustic irregularities such as frequency jumps or amplitude dips in the first two formants are probably due to the subglottal coupling effect.

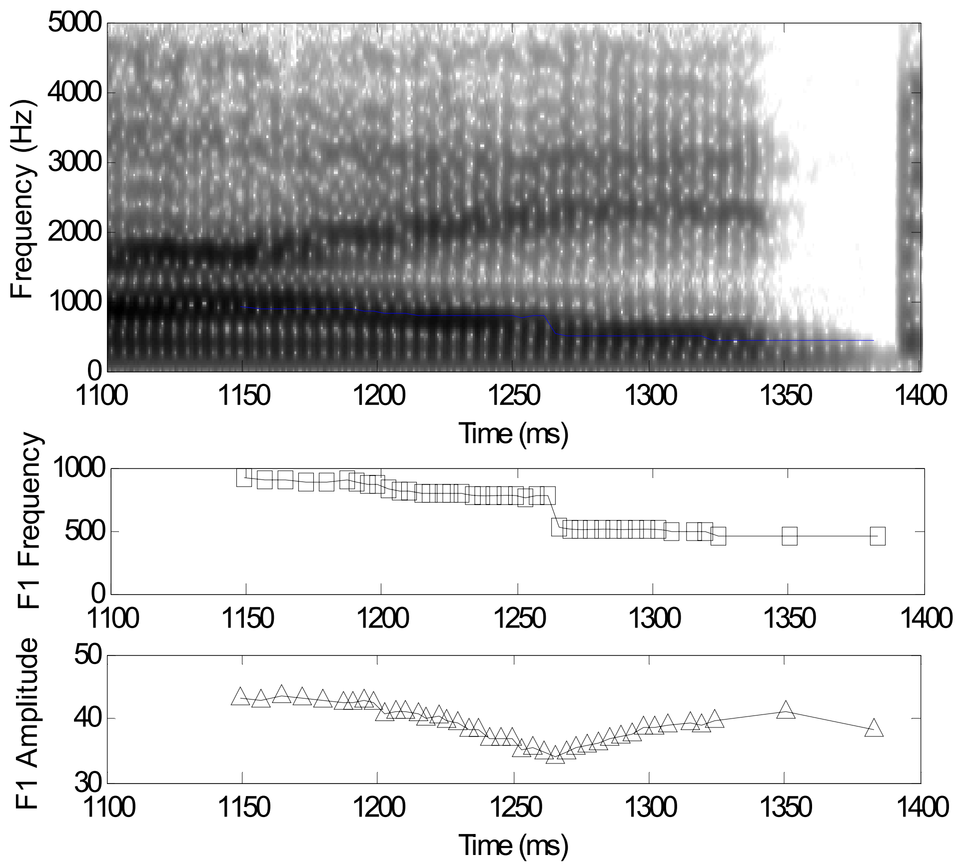
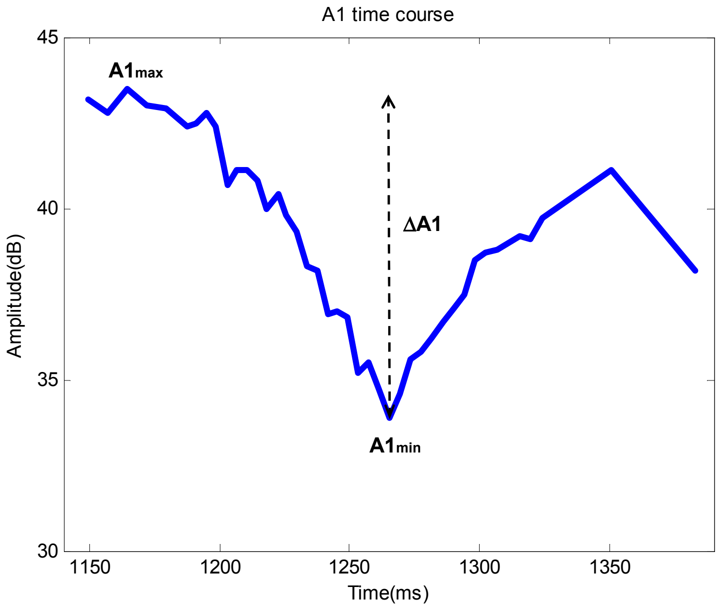
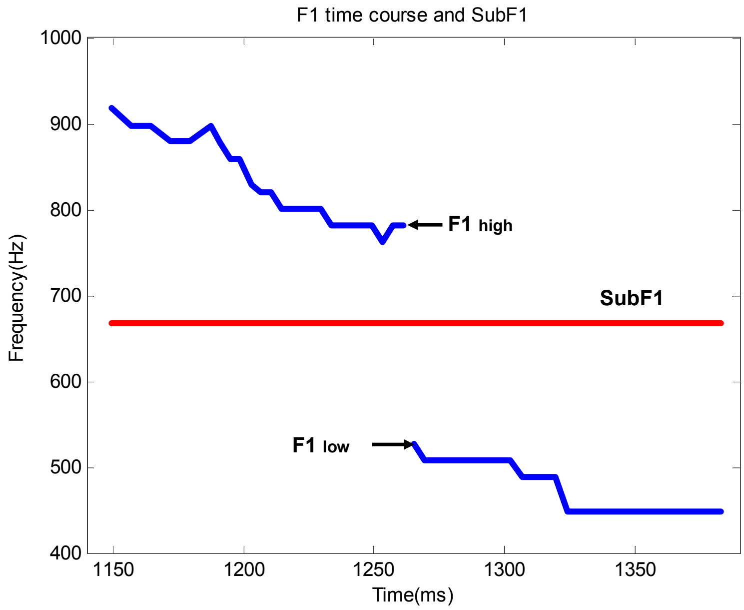


Figure 2.3 An example of the trajectories of F1 frequency and F1 amplitude for a female speaker F1, from the diphthong /ɑɪ/ in "hide" (Upper: Spectrogram; Middle: $F1$ formant time course; Lower: $F1$ amplitude time course). When $F1$ passes through the speaker's $SubF1$, there is a frequency jump in $F1$ formant frequency, and simultaneously, $F1$ amplitude becomes a minimum.

To quantify the subglottal effects arising from the first subglottal resonances on speech spectra, the amount of $F1$ amplitude attenuation (ΔA1) defined as in Figure 2.4(a), was measured from the time course of A1 in the diphthong. In addition, values of $F1_{high}$ and $F2_{low}$ were measured at the frequency jump in $F1$ trajectory, as indicated in Figure 2.4(b). The measurements of ΔA1, $F1_{high}$, $F1_{low}$ and $F1_{middle}$ (the average of $F1_{high}$ and $F1_{low}$) for each speaker are summarized in Table 2.4. For comparisons, *SubF*1 values obtained from subglottal signals are also given from Table 2.2.



(a)



(b)

Figure 2.4 The same utterance as in Figure 2.3. (a) $F1$ amplitude trajectory and $F1$ amplitude attenuation ($\Delta A1 = A1_{max} - A1_{min}$). $A1_{max}$ is a local maximum of the amplitude of $F1$ around the place of $A1_{min}$. $A1_{min}$ is a local minimum of the amplitude of $F1$ which occurs near $SubF1$. (b) $F1$ frequency trajectory, $F1_{high}$ (the frequency on the higher side at a frequency jump), $F1_{low}$ (the frequency on the lower side at a frequency jump) and the speaker's $SubF1$.

Table 2.4 *F*1 amplitude attenuation (ΔA1), high frequency side of *F*1 frequency jump (*F*1high), low frequency side of *F*1 frequency jump (*F*1low), average of *F*1high and *F*1low (*F*1middle), and *SubF*1. Jump frequency represents how often *F*1 jump was found in the diphthongs.

Speaker	ΔA1(dB)	$F1_{high}$(Hz)	$F1_{low}$(Hz)	$F1_{middle}$(Hz)	Jump Frequency	*SubF*1(Hz)
M1	2.9	-	-	-	0/8	547
M2	5.8	-	-	-	0/8	535
M3	6.9	684	469	577	1/5	540
F1	4.4	797	494	646	10/12	648
F2	5.4	778	530	654	7/8	664
F3	5.0	733	498	616	1/4	645

Table 2.4 shows that the amplitude attenuation in the first formant peak near the first subglottal resonance was observed for all speakers. Even when there was no *F*1 frequency jump, the amplitude attenuation was found. This implies that amplitude attenuation is more robust cue of subglottal coupling effects. *F*1 frequency jumps were more often found for female speakers than for male speaker. The values of $F1_{middle}$ which were taken from the frequency discontinuities in *F*1 trajectories were within 40 Hz from *SubF*1 values obtained from an accelerometer.

In sum, previous research about *SubF*2 and the feature [back] raised the possibility that *SubF*1 may play a role in defining the feature [low] contrast, similar to the role of *SubF*2 in defining feature contrast of [back]. Testing this hypothesis for the feature [low], we measured *SubF*1 and *F*1 from English monophthong vowels for several adult speakers of English, using Chi's data (Chi and Sonderegger, 2004; 2007). Time courses of *F*1 frequency and *F*1 amplitude were obtained from English diphthong vowels. Results show that acoustically irregular region was found near speaker's *SubF*1. The boundary between *F*1 frequencies of [+low] vowels and those of [–low] vowels agreed with the speaker's *SubF*1.

2.3 Experiment #2: Across Languages

The results from study of the first subglottal resonance and the vowel feature [low] for English in the previous section raised possibilities that *SubF*1 may play a role in defining the boundary of the universal vowel feature [low] for different languages. The acoustic coupling between the oral cavity and the subglottal system are theoretically expected, independent of languages, since the coupling arises from the physical characteristics of the human speech production system. We hypothesized that *SubF*1 plays a role in defining the vowel feature contrast of [low], independent of language in this part. To test this hypothesis, we examined whether *F*1 frequencies of [+low] vowels and [–low] vowels from the various languages show the quantal relation between *F*1 and Sub*F*1. To test the quantal hypothesis by using individual data of *F*1 and *SubF*1 frequencies is ideal, as in the experiment for English and Korean in this chapter and Chapter 3, respectively. Due to the limitation of data collection, the hypothesis was indirectly tested in this section, for various languages in the world.

2.3.1 Method

From a variety of reports in the literature (see Section 2.3.2.), the first formant frequencies of vowels were collected in various languages for adult male speakers and for adult female speakers, separately. The collected data of *F*1 were divided into two groups: a [+low] vowel group and a [–low] vowel group. We determined the *F*1 boundary between two groups using the method of Support Vector Machine (SVM) and compared the boundary with the value of *SubF*1 obtained from the first experiment and the average value of *SubF*1 reported in the literature.

2.3.2 Results

The results presented in this thesis research are preliminary data from ten languages – Spanish, American English, Hindi, Portuguese, Bengali, Japanese, German, Dutch, Korean and French in the 15 most widely spoken languages in the world, except for Dutch (Cervera et al., 2001; Yang, 1996; Khan et al., 1994; Escudero et al., 2009; Ray and Ghoshal, 1996; Tohkura et al., 1992; Heid et al.,1995; Pols et al., 1973; Gendrrot and Adda-Decker, 2005) The more detailed information of number of subjects and averaged values of the first two formant frequencies of standard vowels for each language is given in Appendix B.

The first two formant frequencies of vowels from the languages as listed above were collected. The $F1$-$F2$ plots of the collected formant frequencies are shown in Figure 2.5 for male speakers, and in Figure 2.6 for female speakers, respectively. In Figure 2.5, there is an avoided region in $F1$ frequencies between 535 Hz and 650 Hz (center around 590 Hz), and by the frequency, $F1$ formant frequencies of low vowels and those of non-low vowels were well separated. The mean value of first subglottal resonances measured using an accelerometer from Experiment #1 for male speakers falls into this frequency gap between low vowels and non-low vowels. The boundary between low vowels and non-low vowels determined by Support Vector Machine (SVM) was $F1$ = 592.5 Hz. This frequency boundary agrees with the average $SubF1$ of 603 Hz reported in the literature (Cranen and Boves, 1987; Cheyne, 2002; Harper et al, 2001; Habib et al., 1994; Chi, 2005; Stevens, 1998).

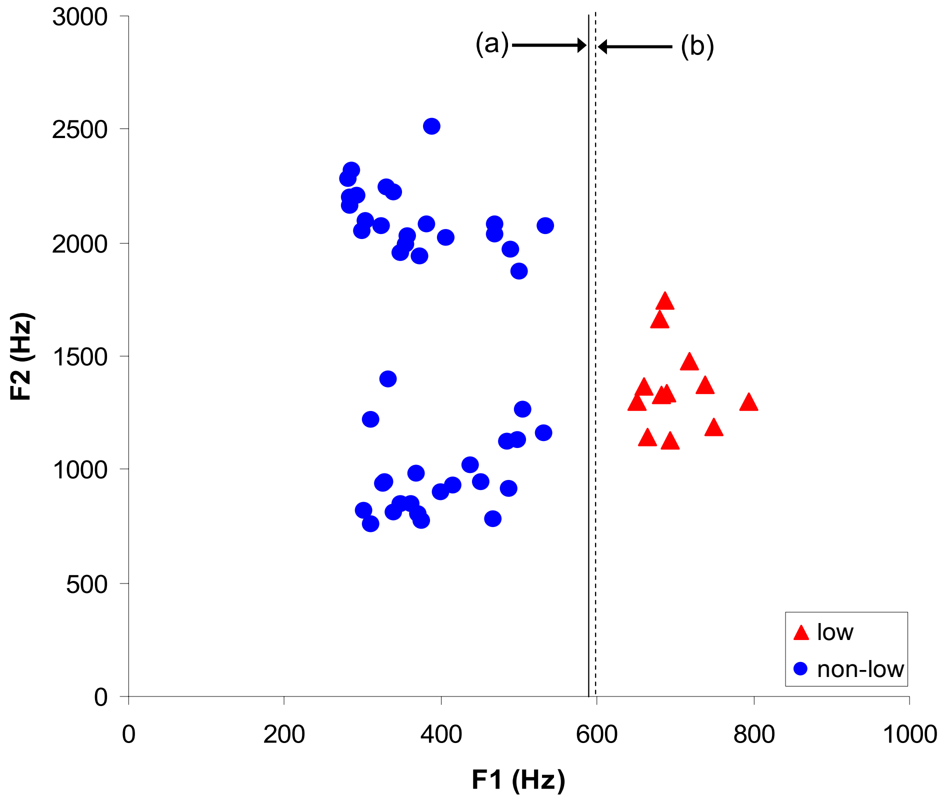


Figure 2.5 F1-F2 plot of vowels produced by male speakers from ten languages in the world. Each points represent averaged formant values of a certain vowel for each language. Red circle: non-low vowels /e, i, o, u/, blue triangle: low vowels /a, ɑ, æ/. (a) Solid line: a boundary between low vowels and non-low vowels of collected data at 592.5Hz, (b) dashed line: average of *SubF*1 of males from the literature by direct measurements at 603 Hz (Cheyne, 2002; Harper et al, 2001; Habib et al., 1994; Chi, 2005**)**.

Similar observations were made in Figure 2.6 for female speakers. There was an avoided region in *F*1 frequencies between 650 Hz and 750 Hz, and *F*1 formant frequencies of low vowels and those of non-low vowels were well separated by the avoided region. For low vowels, the *F*1 frequencies were higher than the avoided region, whereas the *F*1 frequencies were lower than the avoided region for non-low vowels.

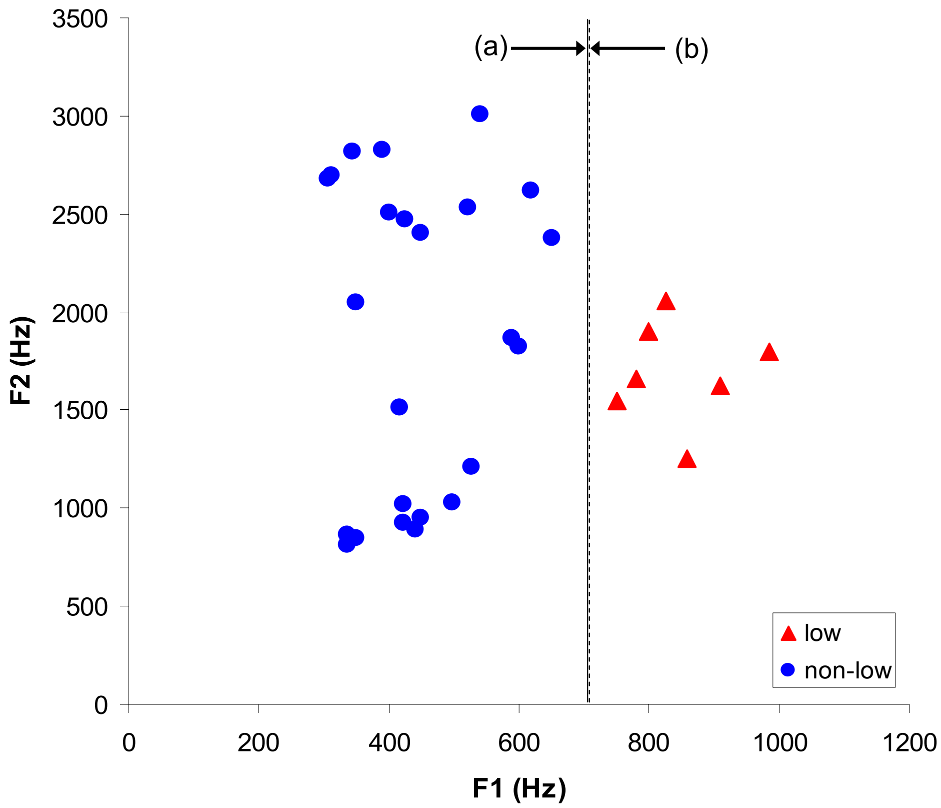


Figure 2.6 F1-F2 plot of vowels produced by female speakers from five languages in the world. Each points represent averaged formant values of a certain vowel for each language. Red circle: non-low vowels /e, i, o, u/, blue triangle: low vowels /a, ɑ, æ/. (a) Solid line: a boundary between low vowels and non-low vowels of collected data at 700 Hz, (b) dashed line: average of *SubF*1 of males from the literature by direct measurements at 703 Hz **(**Cheyne, 2002; Habib et al., 1994; Chi, 2005).

The mean value of the first subglottal resonances measured using an accelerometer from Experiment #1 for female speakers falls into this frequency gap between low vowels and non-low vowels. Also, this frequency gap agrees with the reported value of 703 Hz of females (Cheyne, 2002; Habib et al., 1994; Chi, 2005**;** Stevens, 1998). The boundary between low vowels and non-low vowels determined by SVM was *F*1 = 700 Hz.

Figure 2.7 and Figure 2.8 show the distributions of $F1$ frequencies across vowels from various languages for male speakers and for female speakers, respectively. In Figure 2.7, there was a dip in $F1$ distributions at around 600 Hz. This frequency agrees with the average value 600 Hz of the first subglottal resonances of males as reported in the literature (Stevens, 1998; Cheyne, 2002; Harper et al, 2001; Habib et al., 1994; Chi, 2005). Similarly, in Figure 2.8, there was a dip in $F1$ distributions at around 700 Hz. This frequency agrees with the average value 700 Hz of the first subglottal resonances of females as reported in the literature (Stevens, 1998; Cheyne, 2002; Habib et al., 1994; Chi, 2005). It is noteworthy that the overall pattern of $F1$ distributions across languages for male speakers was similar with the pattern for female speakers.

In sum, we indirectly tested a quantal hypothesis that $SubF1$ plays a role in defining the vowel feature contrast of [low], independent of language in this part. From the literature, the first formant frequencies of vowels were collected from various languages for adult male speakers and for adult female speakers, separately. The collected data of $F1$ were divided into two groups: a [+low] vowel group and a [–low] vowel group. The result shows that the boundary between these two groups agreed with the average value of $SubF1$ from the laboratory study with English, and the reported values in the literature. In addition, there was a disfavored frequency region in the $F1$ distributions near the average value of $SubF1$.

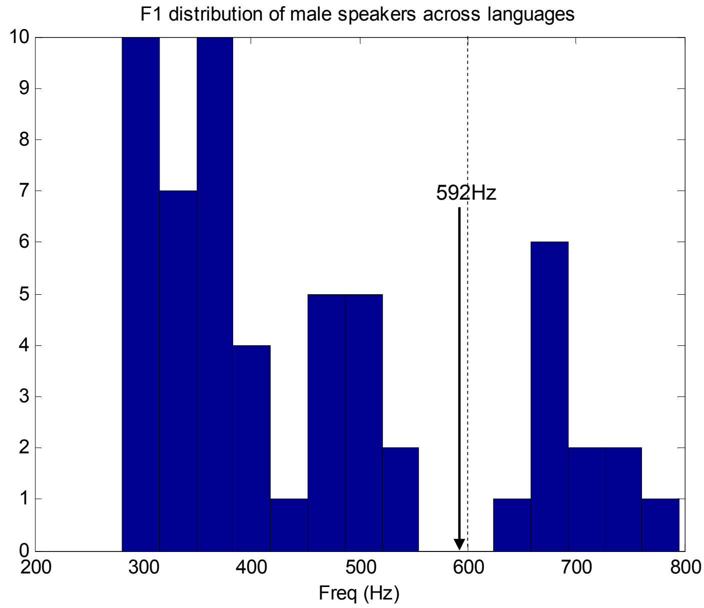


Figure 2.7 Distribution of *F*1 frequencies across vowels from ten languages for male speakers. There is an empty zone at around 590 Hz. The dashed vertical line represents the averaged *SubF*1 of males from literature.

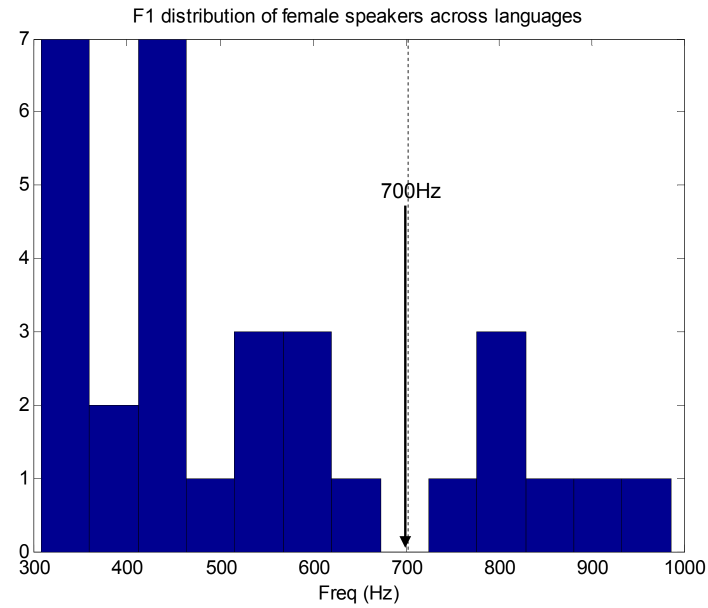


Figure 2.8 Distribution of *F*1 frequencies across vowels from five languages for female speakers. There is an empty zone at around 700 Hz. The dashed vertical line represents the averaged *SubF*1 of females from literature.

2.4 Discussion and Conclusions

The objective of this part in this thesis research was to explore the role of the first subglottal resonance (*SubF*1) in defining the feature contrast of [low]. Previous research suggested that the second subglottal resonance (*SubF*2) defines the vowel feature contrast of [back] for English (Stevens, 1989; Chi and Sonderegger, 2004, 2007; Lulich et al., 2007; Lulich, 2009). The results from study on *SubF*2 and the feature [back] raised the possibility that *SubF*1 may play a role in defining another feature contrast in vowels. It was hypothesized that *SubF*1 plays a role in defining the feature contrast of [low], as similar as to *SubF*2 plays a role in defining the feature contrast of [back]. The following questions were addressed. (i) How are the acoustic characteristics of *F*1 formant frequency and amplitude near *SubF*1 for English vowels? (ii) How do *F*1 frequencies relate to the speaker's *SubF*1 for English? (iii) How are *F*1 distributions of vowels from various languages relative to the average of *SubF*1?

In the first experiment of this chapter, we determined whether acoustic coupling between the first subglottal resonance (*SubF*1) and the *F*1 frequency for vowels creates a region near *SubF*1 in which the *F*1 prominence shows an irregularity. The time course of *F*1 in relation to *SubF*1 was examined for diphthongs produced by a number of speakers of English using the database of Chi and Sonderegger (2004, 2007). A discontinuity in *F*1 or a dip in amplitude of the *F*1 prominence was observed as it passed through *SubF*1. These findings are similar to previous observations that a discontinuity in *F*2 or a dip in amplitude of the *F*2 prominence was observed as it passed through *SubF*2 (Chi and Sonderegger, 2004; Chi and Sonderegger, 2007; Lulich, 2009). This result implies that a frequency discontinuity or amplitude attenuation of the *F*1 prominence near *SubF*1 may be caused by the subglottal coupling.

Amplitude attenuations were a more robust cue than frequency discontinuities. Amplitude attenuations were always found near speaker's *SubF*1 for each speaker, while frequency discontinuities were not always found. These observations agree with the results for *F*2 and *SubF*2 (Chi and Sonderegger, 2004; 2007). The range of amplitude attenuations in $F1$ near $SubF1$ was 3–7 dB. The range of frequency jumps was 220–300 Hz from this study. On the other hand, the results for $F2$ and $SubF2$ show that the range of amplitude attenuations in $F2$ near $SubF2$ was 5–12 dB. The range of frequency jumps was 50–300 Hz, which is similar to that of current study of $SubF1$ and $F1$. However, frequency jumps occurred more often in $F2$ near $SubF2$ than in $F1$ near $SubF1$. In addition, the frequency jump near $SubF2$ was more visible in the spectrogram.

These observations may imply that the subglottal coupling effects of $SubF2$ on $F2$ are more prominent than the subglottal coupling effects of $SubF1$ on $F1$. This result agrees with the theoretical expectations that the coupling effect arising from the first subglottal resonances is less prominent because the acoustic energy loss occurs due to the glottal opening (Stevens, 1998). Variations of the amount of the frequency jump or amplitude attenuation were found depending on speakers, which may arise from the individual differences in the glottal impedance. Sometimes though the amplitude attenuation occurred, no frequency jump was found. Both acoustic observations of frequency jumps and amplitude attenuations arise from the subglottal coupling effects. However, the origins of these two acoustic results are slightly different. Amplitude attenuations of formants near subglottal resonances can occur due to the zeros introduced by the subglottal poles. At these frequencies, the sound energy in the oral cavity is expected to be absorbed by the subglottal system, causing a reduction of the amplitude of formants. On the other hand, frequency discontinuities or frequency jumps of formants near subglottal resonances may arise from the pole-zero separation due to the coupling between the oral cavity and the subglottal system. The amount of pole-zero separations is assumed to depend on the glottal opening size. Therefore, for some speakers and for a certain condition of phonation, it can be expected that the first formant peak is broader, and a pole and a zero are not well separated (Stevens, 1998; Chi and Sonderegger, 2007). $F1$ frequency jumps are more often found for female speakers. In theory, larger subglottal coupling effects would occur as the size of the glottal opening is larger. Usually female speakers have larger and longer glottal openings (Hanson, 1997). Thus, it might be expected that for females' speech the subglottal coupling effects are more prominent. However, it is not clear whether this is related to gender differences due to limited data.

The data of the subglottal signal analysis show that the first subglottal resonance was in the range of 530–690 Hz from nine speakers. The ranges of *SubF*1 agree with invasively measured *SubF*1 values which were about 600 Hz for adult male speakers and about 700 Hz for adult female speakers (Ishizaka et al., 1976; Cranen and Boves, 1987). Individual differences in the first subglottal resonance depending on speakers might arise from the differences in the shape and size of the subglottal system. For example, Ishizaka reported the average of 700 Hz for *SubF*1 from Japanese females, while the current data were obtained from American females and the average of *SubF*1 was 651 Hz. This was slightly lower than that from Ishizaka. This difference may arise from the difference in the body size.

The values of $F1_{\text{middle}}$ which were taken from the frequency discontinuities in *F*1 trajectories were within 40 Hz from *SubF*1 values obtained from an accelerometer. This result suggests that $F1_{\text{middle}}$ values can give us a good estimate of *SubF*1 when direct measurement data of *SubF*1 are not available. This also agrees with the previous findings for *SubF*2 estimation from acoustic speech signals. Lulich (2009) estimated the value of *SubF*2 from the locations of *F*2 discontinuities in the transitions of vowels and consonants for one male subject. Recent studies (Wang et al., 2008a,b) reported that the estimation of *SubF*2 based on the locations of *F*2 discontinuities is accurate, compared to the *SubF*2 obtained by using an accelerometer for child speakers. Thus, this method of estimating subglottal resonances from speech signals is used for the experiment on children's speech in Imbrie's corpus (Imbrie, 2005) in Chapter 4.

The data of monophthong analysis for English show that *F*1 frequencies for 99% of low vowel tokens were higher than *SubF*1 of speakers, while *F*1 for 96% of non-low vowel tokens were lower than *SubF*1. This observation supports our hypothesis that *SubF*1 of speakers may divide *F*1 space into [+low] and [–low]. Some variations were found in *F*1 frequencies of the vowels /ɛ/ and /ɔ/ depending on speakers. This result implies that the *F*1 boundary is located as a frequency band near *SubF*1 of the speakers rather than forming a strict line. These vowels /ɛ/ and /ɔ/ are classified as lax vowels: the articulatory configuration for the vowel /ɛ/ lies between the tense vowels /æ/ and /e/; the articulatory configuration for the vowel /ɔ/ lies between the tense vowels /ɑ/ and /o/. On the other hand, for the tense vowels /ɑ, æ, i, e, o, u/, the expected relation between $F1$ and $SubF1$ was always observed. This observation is reasonable because it is known that for lax vowels, large variations in formant frequencies are observed depending on speakers. For examples, in many dialects, the vowel /ɛ/ is considered to be switched with /æ/. That is, a lowed vowel is produced for /ɛ/, while /æ/ becomes diphthongized to a mid vowel. It was observed that $F1$ of /ɔ/ tokens was below the speaker's $SubF1$ for one subject M1 (one out of eight speakers) who is a Canadian speaker. This observation also implies that $F1$ values of lax vowels differ for individual speakers depending on dialects. Since lax vowels can be regarded as rounded configuration in articulation from tense vowels, large variations are expected in the articulatory configuration for these lax vowels. Also, this might occur because the open quotient (OQ) is smaller during the phonation of lax vowels, causing less subglottal effects on speech spectra. The current data of the relation between $SubF1$ and $F1$ are compatible with the previous work for $SubF2$ and $F2$ (Sonderegger, 2004). The results from English monophthong vowels show that $F2$ was almost always above $SubF2$ of a speaker ($p<0.05$) for front vowels, while $F2$ was almost always below $SubF2$ ($p<0.05$) for back vowels. Thus, both of the current study and the previous study by Sonderegger may support that the general hypothesis that the subglottal resonances may play a role in defining the vowel feature boundaries in English.

The data of $F1$ distributions from various languages show that there were disfavored $F1$ regions near the average values of $SubF1$ for male speakers and for female speakers, respectively. As addressed in Section 2.1, it is ideal to test the hypothesis using individual data of $F1$ and $SubF1$ frequencies, as in the experiments for English and Korean in the first experiment of this chapter and in Chapter 3. Due to the limitation of data collection for subglottal resonances, the hypothesis was indirectly tested in this chapter, for various languages in the world. Though the current data reflect the average $F1$ and $SubF1$ across speakers and across languages, the boundary between $F1$ frequencies of low vowels and those of non-low vowels was surprisingly close to the average value of *SubF*1. The difference was about 10 Hz. Similar observations were made in the data of *F*2 distributions from various languages (Sonderegger, 2004). The *F*2 boundary between front vowels and back vowels agreed with the average values of *SubF*2. Though these experiments used averaged data rather than individual data, both results may support the hypothesis that the subglottal resonances may play a role in defining the boundaries of vowel features [low] and [back], independent of language.

In conclusion, previous research of *SubF*2 and the feature [back] has raised the possibility that *SubF*1 may play a role in defining the feature [low] contrast, similar to the role of *SubF*2 in defining feature contrast of [back]. Testing this hypothesis for the feature [low], we measured *SubF*1 and *F*1 from English vowels for several adult speakers of English, using Chi's data (Chi and Sonderegger, 2004; 2007). The results from English monophthong data show that the boundary between *F*1 frequencies of [+low] vowels and those of [–low] vowels agreed with the speaker's *SubF*1. The result from English diphthong vowels shows that acoustically irregular region was found near speaker's *SubF*1. In addition, the hypothesis was indirectly tested whether *SubF*1 plays a role in defining the feature contrast of [low], independent of language. From the literature, *F*1 frequencies of vowels were collected from various languages for adult male speakers and for adult female speakers, separately. The collected data of *F*1 were divided into two groups: a [+low] vowel group and a [–low] vowel group. The result shows that the boundary between these two groups agreed with the average value of *SubF*1 reported in the literature. There was a disfavored frequency region in the *F*1 distributions near the average value of *SubF*1. These observations support the hypothesis that *SubF*1 may play a role in defining the boundary of the feature [low]. In addition, these observations provide stronger evidence for the general hypothesis that subglottal resonances play a role in defining the vowel feature boundaries, verifying that it is not a coincidence that the second subglottal resonance happens to lie near the frequency boundary between back vowels and front vowels.

Chapter 3. Second Subglottal Resonance and the Feature [back] in Korean

3.1 Introduction

In this chapter, we explore the role of the second subglottal resonance in defining the feature contrast of [back] in Korean. This is preliminary study toward across languages study of the quantal nature in vowels.

As discussed in Chapter 1, previous research suggested that the second subglottal resonance defines the vowel feature [back] for English (Chi and Sonderegger, 2004; Chi and Sonderegger, 2007; Lulich et al., 2007; Lulich 2009). Acoustic irregularities such as frequency jumps or amplitude attenuations in the formant frequencies of $F2$ have often been observed near $SubF2$ for adult English speakers. For each speaker, the second subglottal resonance is located between the $F2$ frequencies of back vowels and those of front vowels (Chi and Sonderegger, 2004). In addition, there has been evidence that listeners' perceptions of the backness of the vowel sound are affected by the locations of $SubF2$ for English adult speakers (Lulich et al., 2007). These subglottal quantal effects on speech signals are theoretically universal and independent of language since the subglottal coupling is related to the physics of human speech production system. The hypothesis is that $SubF2$ plays a role in defining the feature contrast of [back] in other languages as well as in English. Predictions based on this hypothesis are (1) acoustic irregularity will be observed near speaker's $SubF2$ in Korean, (2) $F2$ of front vowels will be placed above speakers' $SubF2$, while $F2$ of back vowels will be placed below $SubF2$ in Korean.

In particular, we tested which of these hypotheses is correct for the low vowel in Korean, since there is only one low vowel with no contrast of [+back] and [-back]: (1) the low vowel is always [+back] or (2) $SubF2$ is avoided but the low vowel is front or back depending on adjacent consonants.

This study consists of three experiments. In the first experiment, we quantitatively measure whether Korean vowels follow the quantal relation between *SubF*2 and *F*2. In the second experiment we determine whether acoustic irregularity is observed near *SubF*2 of speakers in Korean. In the third experiment, *F*2 frequencies of the vowel /a/ are examined in relation to *SubF*2 when the vowel is preceded and followed by consonants.

3.2 Experiment #1: Korean Monophthong Vowels

In this part, we determine whether Korean vowel follows the quantal relation between *SubF*2 and *F*2. The measurements of *F*2 and *SubF*2 were obtained from Korean monophthongs in the context /hVdɨ/, where V is a monophthong, and the vowel /ɨ/ was used. As discussed in Chapter 1, Korean has nine monophthong vowels, /i, e, ɛ, ø, ɨ, a, ʌ, o, u/**,** where vowels /i, e, ɛ, ø/ fall into [-back] vowels**,** and /ʌ, o, u/ fall into [+back] vowels**.** The phonological categories of Korean vowels /ɨ/ and /a/ for the feature [back] are controversial. The vowel /ɨ/ (sometimes denoted as /ɯ/) is categorized as a central, front or back vowel. The vowel /a/ is categorized as a central or back vowel (Yang 1996; Ahn, 1998; Ahn and Iverson, 2005). Thus, for these vowels /ɨ/ and /a/, the feature [back] is not specified in this study.

3. 2. 1 Method

Subject and Materials

We made recordings of speech and subglottal signals simultaneously from eight adult native Korean speakers (four male and four female) who were students or affiliates at MIT or Harvard University. To measure the subglottal resonances of a speaker, an accelerometer was attached to the neck of the speaker, and subglottal signals were obtained using the accelerometer. Recordings were made in the sound attenuated room in the Speech Communication Group at MIT or a quiet room at First Korean Church in Cambridge. Korean monophthong vowels, /i, e, ɛ, ø, ɨ, a, ʌ, o, u/**,** were produced in the context of /hVdɨ/. Most of the speakers made three repetitions for each word, and some speakers produced the utterances once or twice. The sampling rate was 16 kHz.

Acoustic Measurements

For eight speakers (four male and four female), *F*2 frequencies were manually measured in the middle of monophthong vowels**.** The total number of tokens for *F*2 frequency analysis was 162 (nine vowels * eight speakers * one to three repetitions). *F*2 formant frequencies were obtained using the 12th order LPC with a 32 ms Hamming window. The second subglottal resonances were also measured from the subglottal signals at the locations in the middle of the monophthongs for the same speakers and for the same utterances. The total number of tokens for *SubF*2 frequency for each speaker was 72. (nine vowels * eight speakers). As secondary interest, the height of each subject was measured so that a comparison could be made between a speaker's height and subglottal resonances.

3. 2. 2 Results

Second Subglottal Resonances

The mean values and the standard deviations of the frequencies of *SubF*2 are summarized in Table 3.1. The range of the means of *SubF*2 frequencies was 1200–1370 Hz for male speakers and 1450–1690 Hz for female speakers. The ranges of *SubF*2 agree with those obtained for adult English speakers reported by Chi and Sonderegger (2007) which were 1280–1370 Hz for adult male speakers, and 1450–1620 Hz for adult female speakers. The relation between speakers' mean *SubF*2 values and speakers' height was analyzed. The *SubF*2 values across speakers are plotted in Figure 3.1. The strong negative correlation between speakers' height and speakers' *SubF*2 was found ($r = -0.769$, $p = 0.0257$). The linear line between speakers' height and *SubF*2 was estimated from the data as follows.

$$y = -13.4968\,x + 3716 \text{, where } y \text{ is } SubF2 \text{ in Hz, and } x \text{ is height in Cm.} \quad (3.1)$$

Table 3.1 The mean and standard deviation (in parenthesis) values of *SubF*2, and height for each speaker. The measurements were obtained from 9 tokens (9 monophthongs) per speaker.

Male			**Female**		
Speaker	***SubF*2 (Hz)**	**Height (Cm)**	**Speaker**	***SubF*2 (Hz)**	**Height (Cm)**
M1	1367 (44)	178	**F1**	1687 (77)	160
M2	1260 (29)	176	**F2**	1451 (58)	160
M3	1298 (89)	180	**F3**	1507 (76)	169
M4	1364 (40)	166	**F4**	1533 (69)	164
Mean	1322		**Mean**	1545	

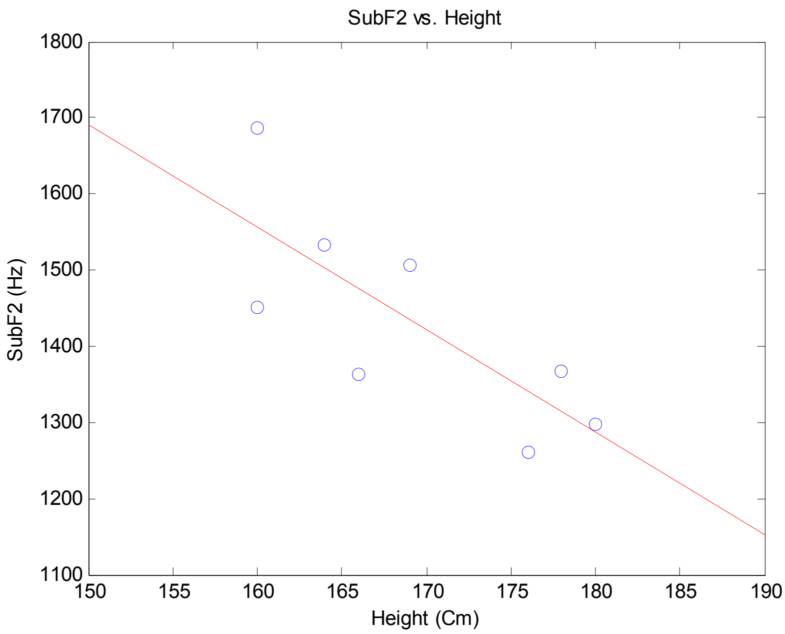


Figure 3.1 *SubF*2 vs. height across eight adult Korean speakers.

Percentage Following Quantal Relation

According to the quantal relation between *F*2 and Sub*F*2 in the vowels, *F*2 frequencies of the back vowels are expected to be lower than *SubF*2 of the speaker, while *F*2 frequencies of the front vowels are expected to be higher than *SubF*2. The percentage of tokens following quantal rules were calculated for the eight Korean speakers as shown in Table 3.2. For Korean front vowels, /i, e, ɛ, ø/, their *F*2 was always above speakers' *SubF*2, while *F*2 was always below speaker's *SubF*2 for Korean back vowels, /ʌ, o, u/. The *F*2 values of the vowel /ɨ/ turn out to be front because it was always above *SubF*2. Thus, this vowel is matched to the front vowel /ɨ/, rather than the back vowel /ɯ/ in the context of /hVdɨ/.

Table 3.2 The percentage of tokens following quantal rules for eight Korean speakers. The total number of tokens is 162.

Front		Back	
	$F2 - SubF2 > 0$		$F2 - SubF2 < 0$
/i, e, ɛ, ø/	100%	/ʌ, o, u/	100%
/ɨ/	100%	/a/	22%
Total (/ɨ/ included)	100%	Total (/a/ included)	81%
Total (/ɨ/ excluded)	100%	Total (/a/ excluded)	100%

Some variability was observed in the *F*2 frequencies for the vowel /a/ in relation to *SubF*2 depending on the speaker, as in the study on German and Swabian (Madsack et al, 2008). The vowel /a/ is phonologically expected to be back. In this study, the feature of [back] was not specified at the beginning because it is sometimes classified as a central vowel and is not paired with the front vowel in the same frequency range of *F*1. The result of *F*2 analysis in this study shows that in many tokens of this vowel, *F*2 was often fronted and higher than *SubF*2 (78%). It might be because the tongue moves forward due to the following alveolar consonant /d/. Therefore, *F*2 of the vowel /a/ (and possibly the vowel /ɨ/) may increase. It was observed that *F*2 was placed close to *SubF*2 (within about 100 Hz for most speakers) for this vowel /a/, no matter whether it is front or back.

In sum, for Korean speakers, the ranges of the means of *SubF*2 frequencies were 1200–1370 Hz for male speakers and 1450–1690 Hz for female speakers, respectively. These ranges of *SubF*2 were not different from those obtained for English speakers. For Korean front vowels, their *F*2 was above speakers' *SubF*2, while *F*2 was below speaker's *SubF*2 for Korean back vowels. Some variability was observed in the vowel /a/. *F*2 is often front in the context of /hadɨ/. Thus, we explored in more detail this variation of *F*2 in /a/ in the third experiment.

3.3 Experiment #2: Korean Diphthong Vowels

In this part, we determined whether acoustic irregularities in *F*2 such as amplitude attenuation or frequency discontinuities are observed near speaker's *SubF*2 in Korean. For the same speakers as in the previous experiment, *F*2 trajectories and spectrograms were obtained.

3. 3. 1 Method

Subject and Materials

We made recordings of speech and subglottal signals simultaneously from eight adult native Korean speakers who were the same speakers as in the previous experiments. To measure the subglottal resonances of a speaker, an accelerometer was attached to the neck of the speaker, and subglottal signals were obtained using the accelerometer. Recordings were made in the sound attenuated room in the Speech Communication Group at MIT or a quiet room at First Korean Church in Cambridge. Korean words containing at least two consecutive different vowels or glides were produced. The target word list is given in Table 3.3. Most of the speakers made three repetitions for each word, and some speakers produced once or twice. The sampling rate was 16 kHz.

Acoustic Measurements

As preliminary research for three speakers (two male and one female), *F*2 trajectories and spectrograms were obtained in /ju/ or /uj/ parts of the target words in Table 3.3**.** This sequence was chosen because in /ju/ or /uj/, the vowel feature for *F*1 is not changed ([+high]). When the tongue height changes abruptly, another resonance from interdental cavities can be introduced (Honda et al., in press). The coupling effects of the dental spaces will be discussed in more detail in Section 3.5. Thus, this selection of sequences allows us to effectively examine the subglottal coupling effects on the feature [back] contrast. When frequency discontinuities or amplitude attenuations in the *F*2 trajectories were observed near *SubF*2, *F*2 frequencies on the high frequency side and on the low frequency side of the discontinuity were measured with a 20 ms Hamming window.

Table 3.3 Target word list with their IPA symbols, backness, and meaning. Each word includes at least two consecutive different vowels or glides. (+: [+back] vowel; -: [-back] vowel;*: not specified.) Each target word is a two-syllable word. In the IPA symbols, distinctions between syllable are made with '-'. For example, in the word /a-i/, the first syllable is /a/, and the second syllable is /i/.

IPA	[back]	Meaning	IPA	[back]	Meaning
/a-i/	* / –	child	/ja-ju/	– / * / – / +	evening amusement
/o-i/	+ / –	cucumber	/ka-jo/	* / – / +	popular song
/u-ɨi/	+ / * / –	rain suit	/a-je/	* / – / –	from the very beginning
/a-u/	* / +	younger brother	/jʌ-ju/	– / + / – / +	room
/i-ju/	– / – / +	reason	/mu-je/	+ / – / –	military arts
/je-ɨi/	– / – / * / –	Polity	/ku-i/	+ / –	roast
/u-ɛ/	+ / –	brother's love	/so-a/	+ / *	children
/u-ju/	+ / – / +	Milk	/kʌ-ul/	+ / +	mirror
/ju-je/	– / + / – / +	extension	/iʌ-ɨi/	– / + / * / –	female doctor

3. 3. 2 Results

Frequency Discontinuities in *F*2

As a preliminary analysis, for three speakers (M1, M3, F1), the subglottal coupling effects on *F*2 trajectories were examined in the /ju/ or /uj/ sequence. The frequency discontinuities in *F*2 near *SubF*2 were often observed in the spectrograms of the utterances for every speaker. Examples of frequency discontinuities in the spectrograms are shown in Figure 3.2**.** The amplitude attenuations in *F*2 are clearly visible when *F*2 crosses through speaker's *SubF*2, which are labeled with circles in the spectrograms. In the upper and lower panels of Figure 3.2, the produced word is /uju/, where the feature changes from [+back] to [-back] in /uj/, and then changes from [-back] to [+back] in /ju/. The frequency discontinuity occurs simultaneously with these feature changes. Similar patterns are found in the middle panel of Figure 3.2. In the word /jʌju/, the first circle corresponds to the point where the feature changes from [-back] to [+back] in /jʌ/ and the second circle corresponds to the point where the feature changes from [+back] to [-back] in /ʌj/.

Regard to the locations of frequency discontinuities, the measurements of *F*2 frequencies on the high frequency side and on the low frequency side of the discontinuity were made and summarized in Table 3.4**.** The measurements of averaged values of $F2_{high}$ and $F2_{low}$ are close to the values of *SubF*2 obtained from subglottal signals using an accelerometer. The difference between them is within 50 Hz as shown in Table 3.4. Frequency discontinuities near *SubF*2 are found in the same frequency range for a given subject.

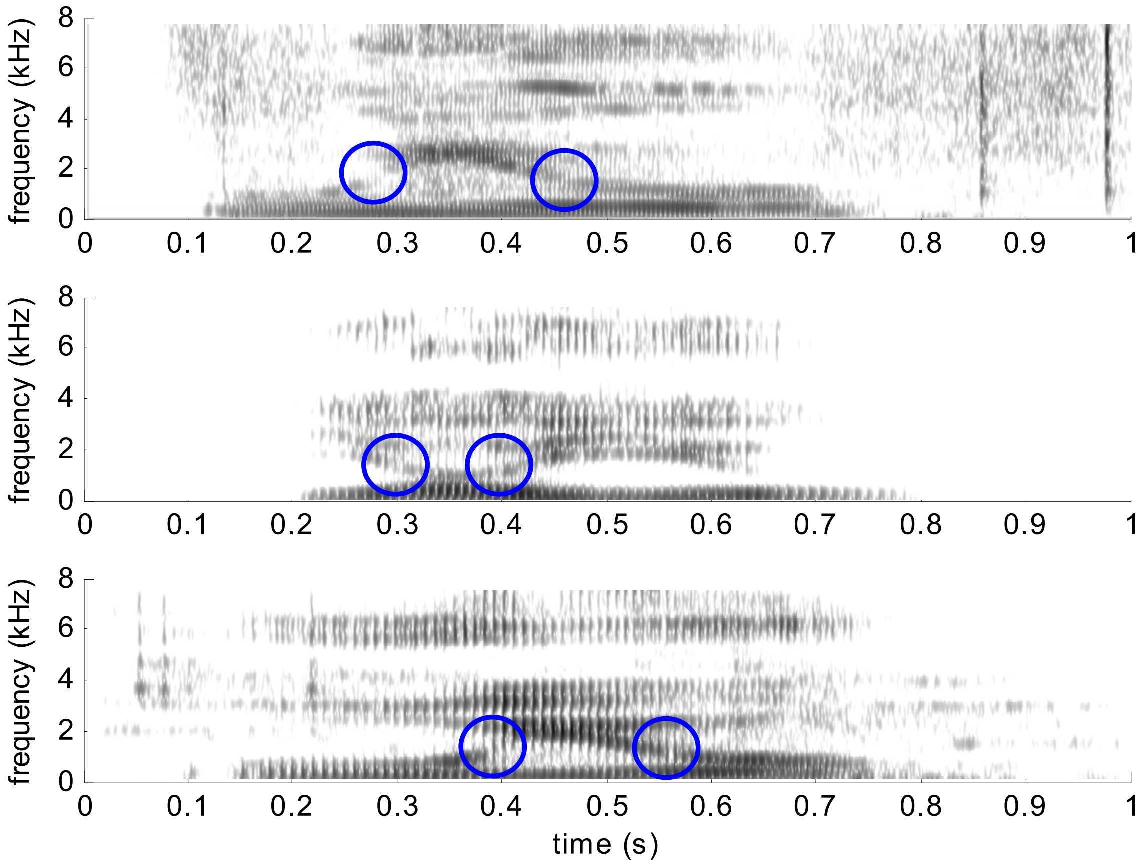


Figure 3.2 Examples of discontinuities in *F*2 as it passed through *SubF*2 in the spectrograms for speaker F1(top), M1(middle), and M3(bottom). The produced words are /uju/(top), /jʌju/(middle), and /uju/(bottom).

Table 3.4 Mean values of $F2$ frequencies on the high frequency side ($F2_{high}$) and on the low frequency side ($F2_{low}$) of frequency discontinuity near *SubF*2, averaged values of $F2_{high}$ and $F2_{low}$ ($F2_{middle}$), and *SubF*2 measured from subglottal signals for comparisons. Measurements were obtained from five tokens for each speaker. The ranges of data are given in parenthesis.

		Speech signal		Subglottal signal
Speaker	**$F2_{high}$ (Hz)**	**$F2_{low}$ (Hz)**	**$F2_{middle}$ (Hz) = (F2 high + F2 low) / 2**	***SubF*2 (Hz)**
F1	1817 (1719–1950)	1525 (1438–1637)	1671 (1579–1767)	1687
M1	1455 (1403–1494)	1198 (1131–1316)	1326 (1307–1360)	1367
M3	1372 (1306–1471)	1139 (1108–1205)	1256 (1219–1294)	1298

In sum, we found acoustic irregularities such as frequency discontinuities or amplitude attenuations in $F2$ near *SubF*2 in Korean vowels, as in English vowels.

3.4 Experiment #3: *F*2 Shift in /a/ and *SubF*2

In this experiment, the *F*2 shift in the vowel /a/ is examined in relation to *SubF*2 when the vowel is preceded and followed by consonants. The results of the first experiment that some variability was found in *F*2 of the vowel /a/ imply that *F*2 in /a/ can be changed due to adjacent consonants and suggest more detailed analysis of *F*2 shift in the various consonant environments in relation to speakers' *SubF*2. Possible views are: (1) the low vowel is always [+back], (2) *SubF*2 is not avoided for *F*2 of /a/, or (3) *SubF*2 is avoided but the low vowel is front or back depending on the adjacent consonant. To test which of these hypotheses is correct for the Korean low vowel /a/, the *F*2 formant frequencies of the vowel /a/ were measured in the context of /CaC/.

3. 4. 1 Method

Subject and Materials

Three adult native Korean speakers (two male and one female), who were students or affiliates at MIT or Harvard University, participated in this experiment. Recordings of speech and subglottal signals were made simultaneously. Subglottal signals were obtained from the accelerometer as in the first and second experiments. The recordings were made in the sound attenuated room in the Speech Communication Group at MIT with the sampling rate of 16 kHz. The Korean low vowel /a/ was produced in the form of /CaC/, where C is a consonant. C is the same in the initial position and final position. However, there is no burst of the stop consonants in the word final position in Korean. Consonants are / k, k^h / (major constrictions by the tongue body), /n, t, t^h, s, tʃ, $tʃ^h$/ (tongue tip), /m, p, p^h / (lips). The velar nasal /ŋ/ cannot be placed in the initial position. Thus, there are total of 11 consonant environments.

Acoustic Measurements

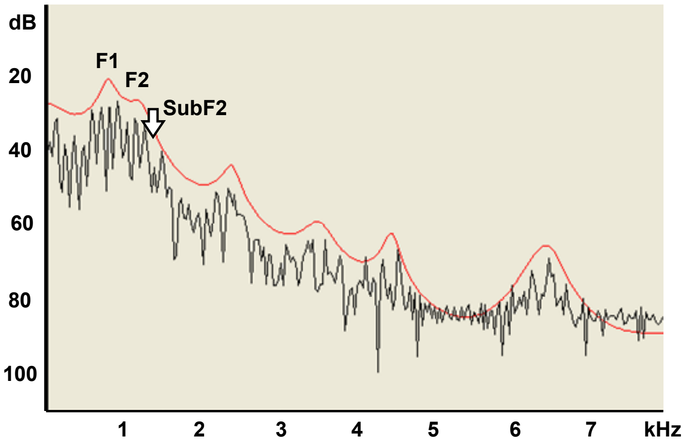
For the speakers, *F*2 frequencies were manually measured in the middle of the vowel /a/. *F*2 formant frequencies were obtained using the 12^{th} order LPC with a 32 ms Hamming window. The second subglottal resonances were also measured from the subglottal signals in the middle of the vowels for the same speakers and for the same utterances.

3. 4. 2 Results

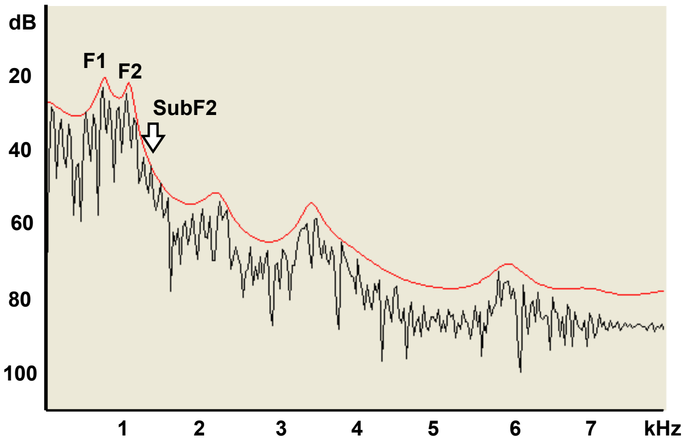
Spectra

The effects of the second subglottal resonance on the spectra of the vowel /a/ were analyzed. Examples of DFT spectra, taken in the middle of the vowel /a/, produced by a male speaker (M1), with a 32 ms Hamming window, are shown in Figure 3.3. The 20^{th} order of LPC spectra are superimposed. The speaker's second subglottal resonance is pointed by arrows at 1370 Hz. In Figure 3.3, the spectra of the vowel /a/ are compared depending on adjacent consonants. The first panel shows the spectrum of the vowel /a/ in /kak/. The consonant /k/ is velar. The second panel shows the spectrum of the vowel /a/ in /mam/. The consonant /m/ is labial. The third panel shows the spectrum of the vowel /a/ in /nan/. The consonant /n/ is alveolar. The reason why /kak/ is shown instead of /ŋaŋ/ is because /ŋ/ cannot appear in the initial position in Korean.

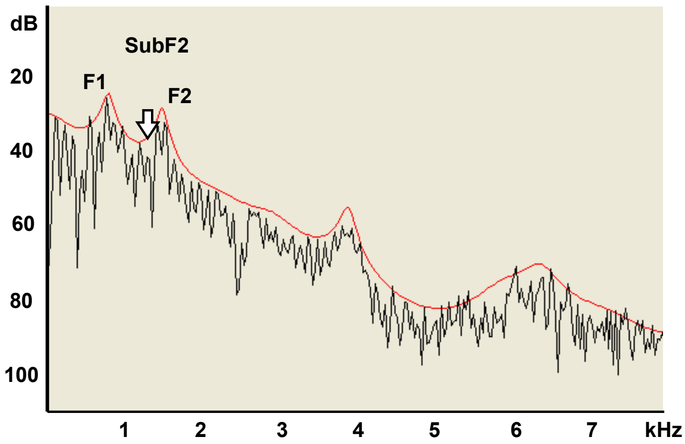
In Figure 3.3(a), the amplitude dip is clearly visible at around 1370 Hz, the speaker's *SubF*2. The second prominence of the vowel is placed slightly lower than the amplitude dip. In Figure 3.3(b), in /mam/, the second prominence of the vowel is placed slightly lower than *SubF*2 as in /kak/. On the other hand, in Figure 3.3(c), the second prominence of the vowel is placed slightly higher than *SubF*2. The amplitude dip is clearly visible at around his *SubF*2.



(a)



(b)



(c)

Figure 3.3 Spectra taken from the middle of /a/ produced by a male speaker (M1). His *SubF*2 is 1370 Hz. (a) in /kak/, *F*2 = 1285 Hz, (b) in /mam/, *F*2 = 1196 Hz, (c) in /nan/, *F*2 = 1481 Hz.

*F*1-*F*2 Plot

The first two formant frequencies of the vowel /a/, when the vowel is surrounded by Korean consonants for which the constrictions are made in different places in the mouth using the tongue blade. Figure 3.4 shows *F*1-*F*2 plots of /a/ tokens produced by the male speaker M5 and the female speaker F1. Horizontal lines depict speakers' *SubF*2 values which were taken from accelerometer signals. Consonants are / k, k^h / (velar), /n, t, t^h, s, tʃ, tʃ^h/ (alveolar or post-alveolar), /m, p, p^h / (labial).

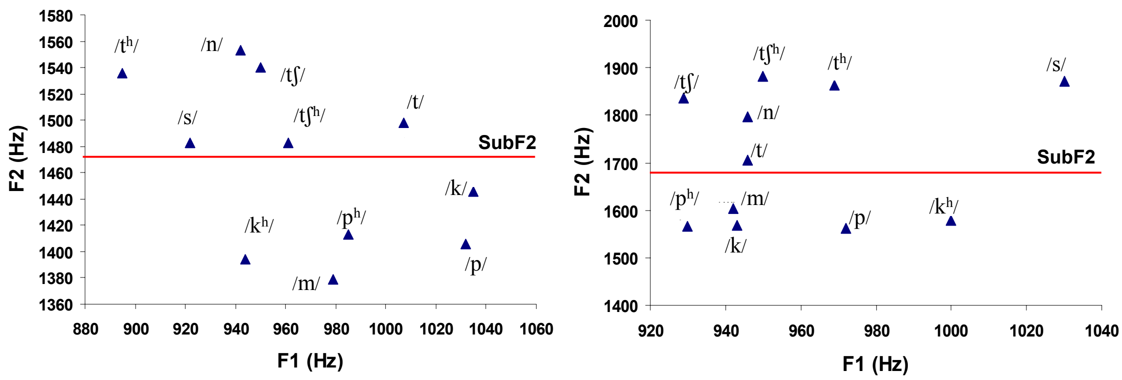


Figure 3.4 *F*1-*F*2 plot of /a/ tokens produced by the male speaker M5 (left) and the female speaker F1 (right). Horizontal lines depict speakers' *SubF*2 values which were taken from accelerometer signals.

Again, the results show that the values of *F*2 of the vowel /a/ is higher than the speaker's *SubF*2 if the adjacent consonants are alveolar (/n, t, t^h, s, tʃ, tʃ^h/), while they are lower than *SubF*2 if the consonants are velar (/ k, k^h /) or labial (/m, p, p^h/). The locations of *SubF*2 are avoided for *F*2, and *SubF*2 forms a dividing line between alveolar and velar/labial. In contrast, for *F*1, evidence of relations between *F*1 formant and the place of constrictions for adjacent consonants is not found. Discussion for this finding will be presented in Section 3.5.

*SubF*2 and *F*2 Distributions

Figure 3.5 shows the mean frequencies of *F*2 of the vowel /a/ depending on the place of constrictions for adjacent consonants. The horizontal lines represent speakers' *SubF*2. The mean of *F*2 frequencies of the vowel /a/ surrounded by labial or velar consonants is significantly lower than the mean of *F*2 frequencies of the vowel /a/ surrounded by alveolar or post-alveolar consonants for all speakers ($p<0.001$). Moreover, speakers' *SubF*2 lie between them. The result indicates that the distribution of *F*2 of the vowel /a/ is quantally represented depending on the place of the tongue for the adjacent consonants, with the boundary of speakers' *SubF*2.

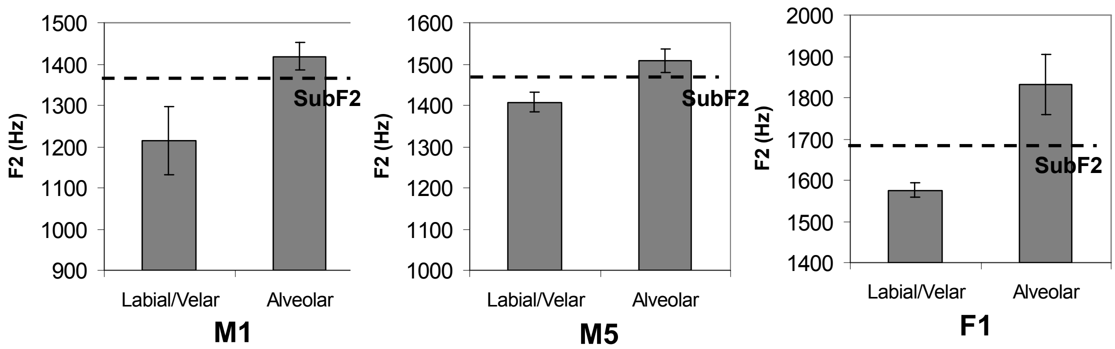


Figure 3.5 The mean frequencies of *F*2 of the vowel /a/ when adjacent consonants are labial and velar vs. when the consonants are alveolar/post-alveolar for the speakers M1 (left), M5 (middle), and F1 (right). The horizontal lines represent speakers' *SubF*2.

The degree of *F*2 shift affected by adjacent consonants is different for each individual. For the speaker M5, the mean *F*2 of /a/ in the environment of alveolar consonants is about 100 Hz higher than in the environment of labial or velar consonants. In contrast, for the speaker *F*1, the mean *F*2 of /a/ in the environment of alveolar consonants is about 260 Hz higher than in the environment of labial or velar consonants.

In sum, we explored the *F*2 shift of the low vowel /a/ in relation to speakers' *SubF*2 when it is preceded or followed by consonants. We examined whether the frequency of *SubF*2 is avoided for producing the vowel and how the *F*2 frequencies are related to *SubF*2. We measured *F*2 frequencies of /a/ tokens in the context of /CaC/ for three Korean adult speakers, where C is a labial, velar or alveolar consonant. The results show that (1) *F*2 frequencies of the vowel are always avoided and (2) *F*2 frequencies of the vowel are changed depending on the adjacent consonants: For all speakers, the values of *F*2 of the vowel /a/ is higher than the speaker's *SubF*2 if the adjacent consonants are alveolar, while they are lower than *SubF*2 if the consonants are velar or labial. The degree of *F*2 shift affected by adjacent consonants is different for each individual.

3.5 Discussion and Conclusions

The objective of this part was to explore the role of the second subglottal resonance in defining the feature contrast of [back] in Korean, as the first step toward across language study on quantal natures of vowels in relation to the subglottal resonances. The hypothesis is that *SubF*2 plays a role in defining feature contrast of [back] in other languages as well as in English since subglottal coupling effects on speech sound are theoretically the same, independent of language.

In the first experiment, we asked if *SubF*2 divides *F*2 vowel space into [+back] and [-back] in Korean. For Korean speakers, the ranges of the means of *SubF*2 frequencies are 1200–1370 Hz for male speakers and 1450–1690 Hz for female speakers. These ranges of *SubF*2 are not different from those for English speakers reported by Chi and Sonderegger (2004). The *SubF*2 values are correlated with the speaker's height. The subglottal resonances are presumed to be related to the size of the subglottal system of a speaker. It is expected that the subglottal resonances decrease as the size of subglottal system increase. The subglottal system consists of trachea, bronchi and lungs, and every air way is divided into two small airways. When only 0 - 13^{th} generations are taken into consideration, the total number of airways is 16383. Despite of the complexity in the subglottal system, however, the observation of linear correlation between the speaker's height and speaker's *SubF*2 suggests that the resonances of the subglottal system might be predictable using a linear equation, as was given in Equation 3.1.

The result in the first experiment shows *F*2 vowel space is divided into [+back] and [-back] in the boundary of speakers' *SubF*2 in Korean as similar as in English. For Korean front vowels, their *F*2 are above speakers' *SubF*2, while *F*2 are below speaker's *SubF*2 for Korean back vowels, except for the vowel /a/. Some variability is observed in the vowel /a/. *F*2 is often higher than *SubF*2. Similar observations are found in German and Swabian vowels (Madsack et al., 2008): *F*2 for front vowels are higher than *SubF*2, while *F*2 for back vowels are lower than *SubF*2 except for the low vowel /a/ in German and Swabian. *F*2 for the vowel /a/ is above *SubF*2 depending on speakers. Both in this study of Korean vowels and in the study of German and Swabian vowels, vowels were produced in the environment of /hVd/. The tongue body might be forwarded during the phonation of vowels, to make a constriction for the following alveolar consonant /d/. In contrast, $F2$ for the low back vowel /ɑ/ in English does not go above $SubF2$ when the alveolar consonant is followed, in the context of /hɑd/. These observations might be due to differences in the vowel inventory in the language. Korean, German and Swabian have only one low vowel /a/ and no pair of low vowels which have [+back]/[-back] contrast. On the other hand, English has a pair of low vowels which are the back vowel /ɑ/ and the front vowel /æ/. The possible relations between $F2$ shift in the vowel /a/ and $SubF2$ will be discussed in more detail in the next.

The results in the second experiment show that acoustic irregularities such as amplitude attenuations or frequency jumps in $F2$ are observed due to subglottal coupling effects in Korean speech signals, as the irregularities are found near subglottal resonances in English. Frequency discontinuities near $SubF2$ are often clearly visible in the spectrograms. The locations of frequency discontinuities are in the same frequency range for a given subject, and the locations are within 50 Hz from the ground truth $SubF2$, which are obtained from accelerometer signals. These observations provide preliminary evidence that the second subglottal resonance has the same effects on the speech signals, independent of languages.

In the second experiment, we examined /ju/ or /uj/ sequence, not /aj/ sequence, which was studied for English (Chi and Sonderegger, 2004, 2007). In the sequence /ju/ or /uj/, the vowel feature for $F1$ is not changed ([+high]). Only the feature for $F2$ is changed from [-back] to [+back] or [+back] to [-back]. On the other hand, the features for $F1$ and $F2$ change together in /aj/. When the tongue height changes abruptly, another resonance from the interdental cavities can be introduced (Honda et al, in press). Honda and colleagues suggest that the resonances of interdental cavities, which are above 1700 Hz for adult Japanese speakers, play a role in distinction between /a/ and /i/ in Japanese. These resonant values are 300-400 Hz higher than $SubF2$. Our results show that the frequency discontinuities of the $F2$ trajectories are related to the frequency of speakers'

*SubF*2 rather than the resonances from interdental spaces in the /ju/ or /uj/ sequence, where vowel feature changes only in the feature [back]. In addition, the *F*2 analysis for the Korean vowel pair /ɛ/ ([-back]) and /ʌ/ ([+back]) and the pair /e/ ([-back]) and /o/ ([+back]) suggests that the *SubF*2 is located in the middle of /ɛ/ and /ʌ/, as well as in the middle of /e/ and /o/ in *F*2 frequencies, whereas the interdental resonance is not located either in the middle of /ɛ/ and /ʌ/, or in the middle of /e/ and /o/ in *F*2 frequencies.

In the third experiment, we examined whether the frequency of *SubF*2 is avoided for producing the vowel and how the *F*2 frequencies are related to *SubF*2. The results show that *F*2 frequencies of the vowel are always avoided. For the three speakers, the values of *F*2 of the vowel /a/ is higher than the speaker's *SubF*2 if the adjacent consonants are alveolar, while they are lower than *SubF*2 if the consonants are velar or labial. The degree of *F*2 shift affected by adjacent consonants is different for each individual. The tongue body might be moved forward during the phonation of vowels, to make a constriction for adjacent alveolar or post-alveolar consonant sounds. On the other hand, for velar consonants, since the constrictions are made by raising the tongue body, the positions of the tongue during the phonation of the vowel are probably more backward than for alveolar consonants. For labial adjacent consonants, the primary constrictions are made by lips, and the positions of the tongue body are probably rather free. Therefore, during the phonation of the vowel, the constrictions for the tongue body would be maintained backward. It is noteworthy in our findings that the boundary of *F*2 shift depending on the constrictions for adjacent consonants might be related to *SubF*2.

Similar *F*2 shift of the vowel in the contexts is often found in other languages. For example, *F*2 of the Mandarin Chinese vowel /a/ is shifted higher when it is followed by the alveolar nasal coda /n/, while *F*2 is lower before the velar nasal coda /ŋ/ (Mou, 2006). In the study of Standard Chinese, *F*2 of /ban/ is about 500 Hz higher than *F*2 of /baŋ/ (*F*2 of /ban/ is about 1600 Hz, and *F*2 of /baŋ/ is about 1100 Hz for a male speaker.) Similar observations are made for the reduced English vowel /ə/, where the feature [back] is not specified. In the sentence "pass a dip", where the reduced vowel is preceded and followed by alveolar consonants, *F*2 of /ə/ is 1491 Hz. When the vowel is preceded and followed by labial consonants as in "rub a book", *F*2 of /ə/ is 912 Hz for a male English speaker (Stevens, 1998). Referring to the mean *SubF*2 for adult male Korean and English speakers is in the range of 1200-1370 Hz; this frequency range falls between *F*2 of /ban/ (about 1600 Hz) and *F*2 of /baŋ/ (about 1100 Hz) for a adult male speaker in Mandarin Chinese. In addition, the frequency range of *SubF*2 falls between *F*2 of "pass a dip" (1491 Hz) and *F*2 of "rub a book" (912 Hz) in English.

These observations suggest the following general implications. For the vowels where the feature [back] not has been specified in a certain language (e.g., /a/ in Korean, Standard Chinese, Japanese, or /ə/ in English), *F*2 shift depending on the place of adjacent consonants is often observed. In addition, the *F*2 shift direction is related to *SubF*2: If the adjacent consonants are alveolar, $F2 - SubF2 > 0$, whereas if the adjacent consonants are velar or labial, $F2 - SubF2 < 0$. In contrast, for the vowels which have the feature [back] contrast (e.g., /ɑ/ in English), although *F*2 of the vowel can be shifted due to the effect of adjacent consonants, the relative position to *SubF*2 ($F2 - SubF2 < 0$) would be maintained. In other words, for the vowels having contrast of the feature [back], the *F*2 information relative to *SubF*2 should be preserved, independent of adjacent consonants, because it may give the listeners information about the backness of the vowel, which needs to be distinctive from another vowel (e.g., /ɑ/ vs. /æ/ in English). For the vowels having no contrast of [back], however, information of *F*2 for the vowel itself is not critical for the perception of the backness of the vowel. Instead, *F*2 can be more freely shifted depending on the adjacent consonants, which might bring additional information of the place of adjacent consonants in relation to *SubF*2.

The finding that *F*2 shift in Korean /a/ vowel depending on the place of adjacent consonants can be related to the finding of the relationships between the *F*2 onset frequency in the consonants and *SubF*2 in English (Lulich, 2009). In the study, the boundary between $F2$ onset of alveolar consonants and $F2$ onset of velar consonants in the low vowel context is about the speaker's $SubF2$. In addition, the boundary between $F2$ onset of alveolar consonants and $F2$ onset of labial consonants in the low vowel context is about the speaker's $SubF2$. This observation also suggests that $F2$ frequencies may carry information related to backness of the tongue body, based on the relation between $F2$ and $SubF2$, depending on whether $F2 - SubF2 > 0$ or $F2 - SubF2 < 0$.

In conclusion, previous research suggested that the second subglottal resonance defines the vowel feature [back] for English. In testing this hypothesis, we explored whether $SubF2$ plays a role in the feature contrast in vowels independent of language. As preliminary research, we made recordings of speech and subglottal signals simultaneously for several adult Korean speakers. We found acoustic irregularities in $F2$ near $SubF2$ in Korean vowels. The boundary between the boundary between [+back] and [-back] agrees with speakers' SubF2. In addition, we tested which of these hypotheses is correct for a low vowel in Korean, which has only one low vowel with no contrast of [+back] and [-back]: (1) the low vowel is always [+back], (2) $SubF2$ is not avoided for $F2$ of the low vowel, or (3) $SubF2$ is avoided but the vowel is front or back depending on adjacent consonants. The measurements of $F2$ and $SubF2$ were obtained in the context /CaC/, where C is a consonant. We found that $SubF2$ was always avoided for the low vowel. If the adjacent consonants were labial or velar, $F2$ of the low vowel was below $SubF2$, whereas if the consonants were alveolar, $F2$ of the vowel was above $SubF2$. These findings suggest that $SubF2$ may have a role in assimilation of the feature [back] in Korean, as opposed to leading to a contrast for the feature in English.

Chapter 4. Development of Vowel production in Young Children

4.1 Introduction

In this chapter, we explore the role of subglottal resonances in defining the vowel features for young children ages 2;6–3;9 years. As we discussed in Chapter 1, previous studies have tested the quantal hypothesis in speech production and perception, and have provided evidence of quantal relations between formant frequencies and subglottal resonances – quantal relation between $F1$ and $SubF1$, and quantal relation between $F2$ and $SubF2$ (Chi and Sonderegger, 2004, 2007; Lulich et al., 2007; Jung and Stevens, 2007). For several children under 16 years old (only one child younger than 6 years), quantal relation between $F2$ and $SubF2$ was observed (Lulich, 2009). Previous observations that older children and adults already show quantal relations raise the question as to when younger children begin to follow quantal relations.

To answer the question for children ages 2;6–3;9 years, the characteristics of vowel formant frequencies were analyzed to examine how the changes in formant frequencies during a 6 month period relate to their subglottal resonances and when the quantal relations appear. The hypothesis is that the quantal vowel boundaries are determined by subglottal resonances of children above a particular age. Predictions based on this hypothesis are that quantal relations between formants and subglottal resonances will be observed above the age. The following questions are addressed. (i) How do the first and second subglottal resonances change in the age range of 2–3 years? (ii) How do $F1$ and $F2$ change in relation to the subglottal resonances in this age range for each child? (iii) Do quantal relations better predict $F1$ and $F2$ changes than vocal tract growth?

In the first experiment, we answer the question about how subglottal resonances change in this age range. Based on Imbrie's database (Imbrie, 2005), the first and second subglottal resonances were estimated. In the second experiment, we answer the question about how the first two formant frequencies change in relation to the subglottal resonances of the children. The first and second formant frequencies were measured in this age range. In addition, when and how adult-like relations between formant frequencies and subglottal resonances develop were examined. In the third experiment, two opposing hypotheses, the quantal hypothesis and the vocal tract growth hypothesis was tested as accounts of how children's production change over time.

4.2 Experiment #1: Subglottal Resonances and *F*3

In this part, we examine how the subglottal resonances change over a period of 6 months for ten children ages 2;6–3;9 years. The first and second subglottal resonances were obtained from the locations of discontinuities in the *F*1 and *F*2 trajectories in diphthongs and from consonant-vowel transitions. In addition, *F*3 frequencies were obtained from the middle of vowels for comparison.

4.2.1 Method

Database

The database was collected and reported by Imbrie (2005). There were ten children with ages 2;6–3;3 years (2 years and 6 months to 3 years and 3 months) at the beginning of the recording session. The utterances were recorded once a month over a period of 6 months. For this research, data from the first recording session (at Time 1) and the last session (at Time 2) were analyzed. Therefore, Time 1 and Time 2 are six months apart, and the age range is from 2;6 to 3;9 years across children. The age distributions for each subject are shown in Figure 4.1. All subjects are monolingual American English speakers. The words were elicited by an experimenter using a story-telling mode (for more detailed description, see Imbrie, 2005). Each target word was repeated 5-15 times. Table 4.1 shows the list of target words. Each word includes English monophthong vowels /i, ɪ, e, æ, ɑ, ɔ, ʌ, ʊ, u/ in the forms of /CVC/ (e.g., "bug"), /CVCV/ (e.g., "baby") and /CV/ (e.g., "bee") or English diphthongs /ɑɪ, ɑʊ, ɔɪ/ in /CV/ (e.g., "boy"). The target word list is the same for each child and each time. The utterances were recorded in a sound attenuated room in the Speech Communication Group at MIT or a quiet room using a digital audio tape (DAT) recorder, with the sampling rate of 48 kHz. The data were down-sampled to 16 kHz for analysis and were low-pass filtered at 8 kHz.

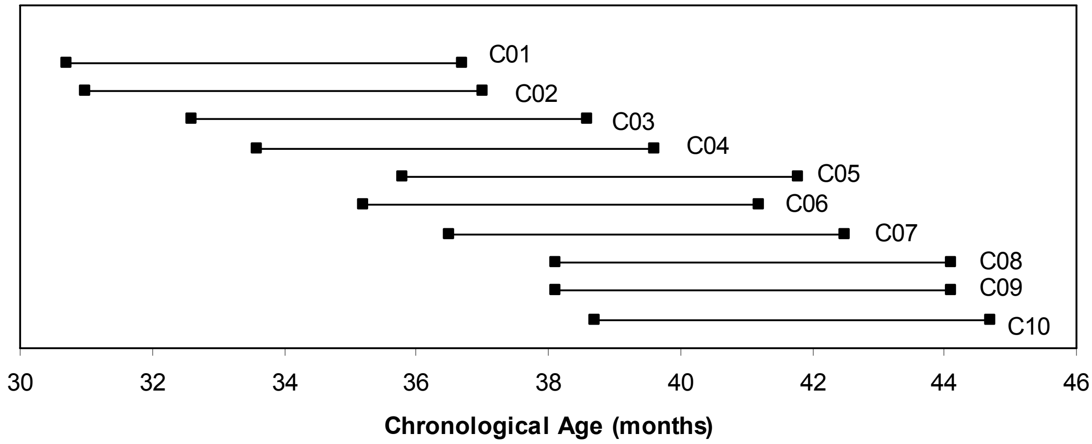


Figure 4.1 The age distributions of children. For each child, the starting points represent Time 1, and the end points represent Time 2. Time 1 and Time 2 are 6 month apart from each other for each child. The age at Time 1 is different for each child. C01: Child 1, C02: Child 2 and so on. Except for C05, the age of a child speaker increases, as the number increases. We follow the notation of child subjects in Imbrie's (2005).

Table 4.1 American English vowel symbols, vowel feature values and target words in Imbrie's database (Imbrie, 2005).

Vowel	[low]	[back]	Target Word
/i/	-	-	"bee"
/ɪ/	-	-	"puppy", "baby", "Kitty", "daddy", "cookie", "Maggie"
/e/	-	-	"baby"
/æ/	+	-	"cat", "daddy", "Maggie", "bath"
/ɑ/	+	+	"bottle"
/ʌ/	-	+	"cup", "tub", "mud", "duck", "bug", "puppy", "tub"
/ɔ/	+	+	"dog"
/ʊ/	-	+	"cookie"
/u/	-	+	"goose"
/ɑɪ/	+/-	+/-	"pie"
/ɑʊ/	+/-	+/+	"cow"
/ɔɪ/	+/-	+/-	"boy"

Acoustic Measurements

Subglottal Resonances

The measurements of *SubF*1 and *SubF*2 were obtained from the locations of discontinuities in the *F*1 and *F*2 trajectories in diphthongs and from consonant-vowel transitions. Previous studies have reported that frequency discontinuities or amplitude attenuations in the *F*1 and *F*2 trajectories are observed near subglottal resonances for adults, especially in diphthongs and consonant-vowel transitions (Chi and Sonderegger, 2007; Lulich, in 2006; Jung and Stevens, 2007). Thus, from the locations of the frequency discontinuities, the frequencies of subglottal resonances can be estimated. The simple method of obtaining the subglottal values from the locations of frequency discontinuities in consonant-vowel transitions was suggested by Lulich (2006). In the locations of frequency discontinuities, *F*2 frequencies on the high frequency side and on the low frequency side of the discontinuity were measured. *SubF*2 was estimated by averaging them. An automatic subglottal resonance detection method from acoustic data was proposed and tested on children's speech data by Wang and colleagues (Wang et al., 2008 a,b). They also estimated *SubF*2 values from the locations of *F*2 discontinuities. The locations of *F*2 jumps were automatically detected by smoothed frequency differences when the frequency differences between time frames are larger than a threshold value, which is optimized by training. Their experimental results showed that measured values of subglottal resonances based on the locations of frequency discontinuities in the formants are very close to the values directly measured from the subglottal signals. The measurement variations showed that COV (standard variation/mean frequency) is in the range of 0.01–0.05, and the frequency difference between frequencies measured obtained using an accelerometer and estimated values was within 7%. (Lulich, 2006; Wang et al., 2008b). These frequency discontinuities due to subglottal coupling can be found in speech signals of young children as young as 2 years old. Thus, in the same way, *SubF*1 and *SubF*2 were measured from the locations of discontinuities in the *F*1 and *F*2 trajectories in diphthongs and consonant-vowel transitions for this study.

An example of how to measure *SubF*2 from acoustic data is illustrated in Figure 4.2. Part (a) in Figure 4.2 is the spectrogram of an utterance, "and Y and" (/ænd yɑɪ ænd/) produced by a 2 year old child for whom subglottal measurements were obtained. The horizontal line represents *SubF*2 directly measured by using an accelerometer from the same speaker. In Figure 4.2 (a), there is a frequency discontinuity of *F*2 between the /ɑ/ sound and /ɪ/ sound in the diphthong /ɑɪ/, when the second prominence of the vowel (*F*2) crosses *SubF*2 due to subglottal coupling. $F2_{high}$ and $F2_{low}$ of the discontinuity were measured, as shown in Figure 4b. *SubF*2 was estimated by averaging $F2_{high}$ and $F2_{low}$. The measured *SubF*2 value agrees well with the *SubF*2 value obtained using an accelerometer.

Frequency discontinuities were found in the same frequency range for a given speaker, and the values of *SubF*1 and *SubF*2 from each token were averaged for each child. Figure 4.3 shows the examples of *F*2 formant tracks from several utterances for one subject from Imbrie's database. The formant track was generated using LPC. Frequency discontinuities were observed at the same frequency range around 2400 Hz. Part (a) in Figure 4.3 is the *F*2 formant track in the diphthong /ɑɪ/. The estimated *SubF*2 value is 2466Hz. Part (b) is the *F*2 formant track in the diphthong /ɔɪ /. The estimated *SubF*2 value is 2398Hz. Part (c) is the *F*2 formant track from the transition between the consonant /d/ and the back vowel /ɔ/ in "dog". The estimated *SubF*2 value is 2412 Hz. Thus, frequency discontinuities were observed at the same frequency range around 2400 Hz.

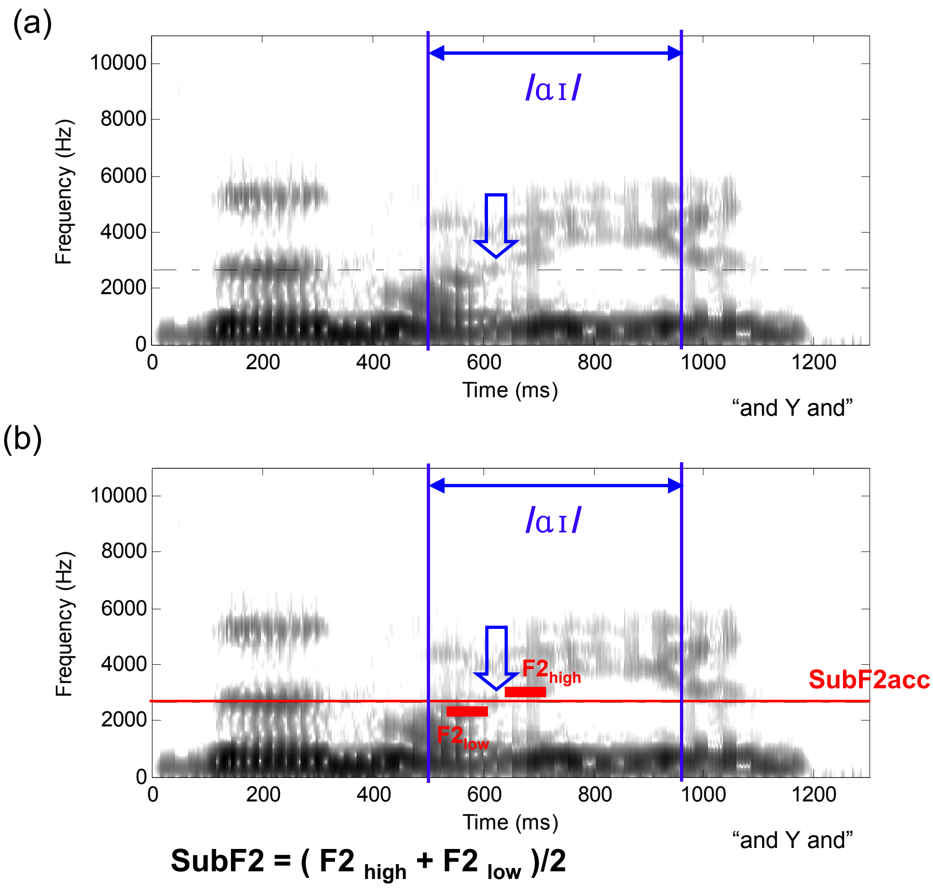


Figure 4.2 An example of subglottal resonance measurements. (a) The spectrogram of the utterance, "and Y and", produced by a 2-year old child. The horizontal line represents *SubF*2 (SubF2acc) which were obtained by using an accelerometer. There is a frequency discontinuity in *F2* near SubF2 due to the subglottal coupling at around 620 ms (pointed with an arrow). (b) The F2 frequencies on the high frequency side ($F2_{high}$) and on the low frequency side ($F2_{low}$) of the discontinuity were measured. The averaged value of $F2_{high}$ and $F2_{low}$ is within 100 Hz of the *SubF*2acc.

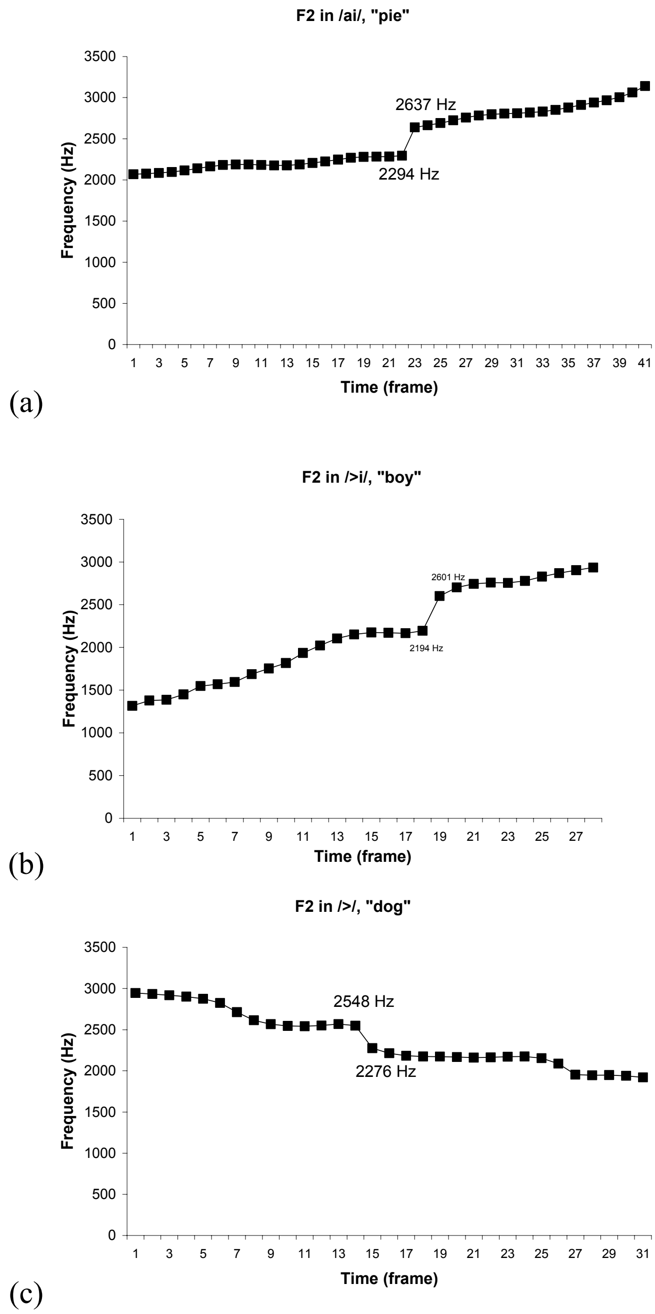


Figure 4.3 Examples of frequency discontinuities due to subglottal coupling for one subject from Imbrie's database. (a) *F*2 in the diphthong /ɑɪ/, *SubF*2~ (2294+2637)/2 = 2466 Hz, (b) *F*2 in the

diphthong /ɔɪ/, *SubF*2 ~ (2194+2601)/2 = 2398 Hz, and (c) *F*2 in the transition from the consonant /d/ to the back vowel /ɔ/, *SubF*2 ~ (2276+2548)/2 = 2412 Hz

*F*3 Frequencies

For ten children, *F*3 frequencies were manually measured in the middle of the vowel /ʌ/ (e.g., in "cup") at Time 1 and Time 2, to compare the effects of the vocal tract growth on the formant frequency changes with the effects of subglottal resonances. The vowel /ʌ/ was chosen because during the phonation of this vowel, the shape of the vocal tract can be assumed to be a uniform tube. This selection allows us to approximately examine the developmental properties of the vocal tract growth with age. The total number of tokens for *F*3 frequency analysis was 200 (10 children * 10 utterances * 2 sessions at Time 1 and Time 2). *F*3 frequencies were obtained using the 12^{th} order LPC with a 32 ms Hamming window.

Statistical Analysis

*SubF*1, *SubF*2, *F*3 Changes over a Six Month Period (Linear Regression)

To see the overall trend of the changes in the frequencies of *SubF*1, *SubF*2 and *F*3 across children in the age span of 2;6-3;9 years, the mean values of *SubF*1, *SubF*2 and *F*3 were collected from each subject, for Time 1 and Time 2. The data were fitted into a linear regression model in the form of $y = ax + b$, where y is the mean value of *SubF*1, *SubF*2 or *F*3 in frequency, x is the value of age in month, a and b are the estimated parameters (slopes and intercepts, respectively).

4.2.2 Results

The mean values and the standard deviations of the frequencies of *F*3, *SubF*1 and *SubF*2 are summarized in Table 4.2. For these children ages 2;6-3;9 years, the range of the means of *F*3 frequencies was 3800 - 4500 Hz. The mean values of *SubF*1 were in the range of 900 – 1100 Hz, while those of *SubF*2 were between 2200 – 2700 Hz. The ranges of *SubF*1 and *SubF*2 were similar but slightly lower than those obtained from the data using an accelerometer from a 2-year-old child, who was slightly younger than children in this thesis research. The values were 1135 Hz for *SubF*1, and 2750 Hz for *SubF*2 (From data of experiments performed in the Speech Communication Group. For this 2-year-old speaker, the data were partially reported Lulich, 2009). Except for the subject C08, the values of *F*3, *SubF*1 and *SubF*2 decreased over a 6-month period for all children. Mean values of *F*3, *SubF*1 and *SubF*2 across ten children at Time 1 and Time 2 are given in Figure 4.4, which also show that these values for children significantly decrease ($p<0.05$) during a 6 month period. The mean values of *F*3, *SubF*1 and *SubF*2 across ten children at Time 1 were 4207, 1047 and 2421 Hz, while those at Time 2 were 4053, 969 and 2329 Hz.

Table 4.2 Mean *F*3, *SubF*1 and *SubF*2 frequencies and standard deviations in parentheses for ten children. (*) For the cases that the mean values of the first subglottal resonances were obtained from less than three measurements, the standard deviations for them were not given in this table.

		C01	C02	C03	C04	C05	C06	C07	C08	C09	C10
***F*3**	**Time 1**	4361 (73)	4023 (177)	4124 (178)	4136 (173)	3995 (175)	4452 (118)	3988 (122)	4545 (59)	4024 (112)	4429 (116)
	Time 2	4084 (95)	3963 (107)	3950 (84)	3892 (174)	3811 (230)	4304 (262)	3875 (260)	4527 (192)	3889 (427)	4238 (242)
***SubF*1**	**Time 1**	1113 (14)	1123 (*)	1053 (60)	1015 (70)	1037 (61)	1086 (42)	1021 (55)	1027 (80)	957 (126)	1042 (106)
	Time 2	1004 (50)	1019 (51)	899 (33)	966 (103)	924 (76)	962 (43)	980 (66)	998 (38)	956 (100)	983 (69)
***SubF*2**	**Time 1**	2409 (69)	2421 (43)	2436 (73)	2393 (53)	2414 (73)	2672 (91)	2263 (131)	2406 (56)	2419 (86)	2390 (139)
	Time 2	2309 (62)	2359 (52)	2354 (75)	2305 (71)	2362 (78)	2589 (96)	2147 (93)	2388 (47)	2233 (73)	2257 (55)

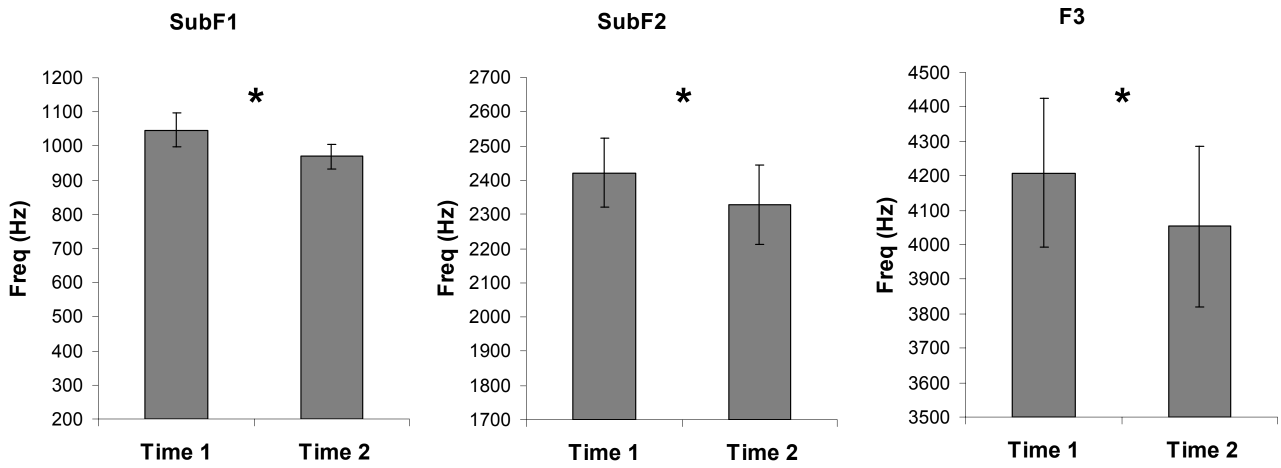


Figure 4.4 The mean frequency values of *SubF*1 (left), *SubF*2 (middle) and *F*3 (right) across children at Time 1 and Time 2. The mean ages at Time 1 and Time 2 are 35 months and 41 months, respectively. The differences between Time 1 and Time 2 are statistically significant ($p < 0.05$) for *SubF*1, *SubF*2 and *F*3.

Figure 4.5 depicts the overall trend of frequency changes in *SubF*1, *SubF*2 and *F*3 across children in the ages from 2 years and 6 months to 3 years and 8 months. The mean values of *SubF*1, *SubF*2 and *F*3 approximately linearly decreased with age.

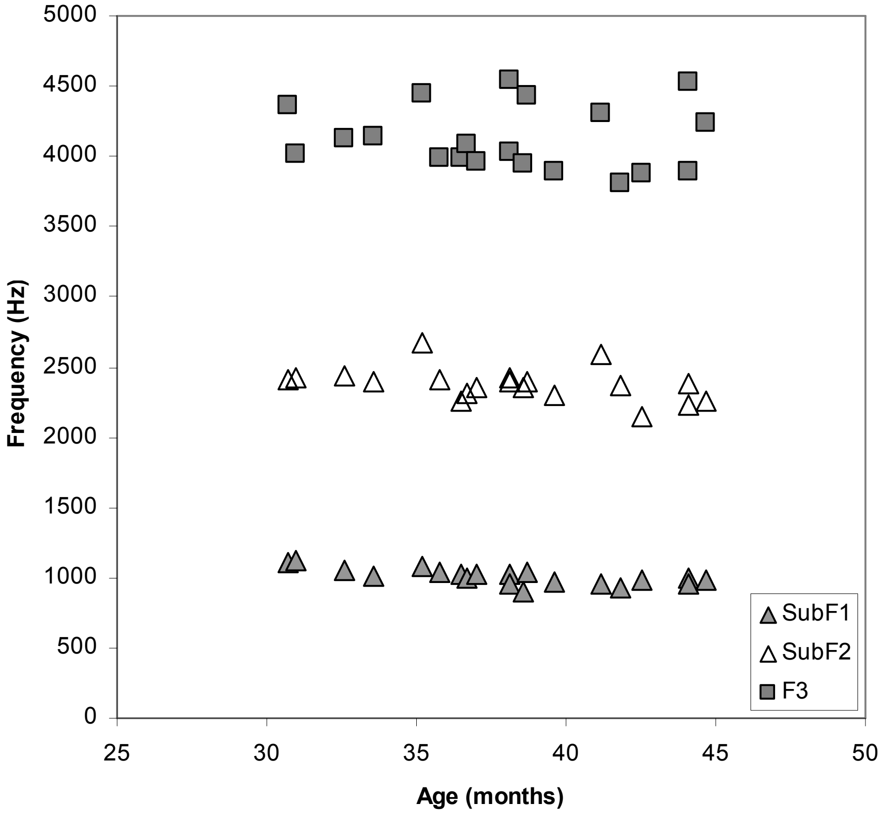


Figure 4.5 The means of *SubF*1 (shaded triangles), *SubF*2 (open triangles) and *F*3 (shaded squares) for children C01-C10 with chronological age in months. Each dot represents the mean values from one subject and from one session.

The data of means of *SubF*1, *SubF*2 and *F*3 were fitted into a linear regression model in the form of $y = ax + b$, where *y* is the value of *SubF*1, *SubF*2 or *F*3, *x* is the value of age in months, *a* and *b* are the estimated parameters (slopes and intercepts, respectively). The estimated values of developmental slope (*a*) for *SubF*1, *SubF*2 and *F*3 with age are given in Table 4.3. The results indicate that all frequency mean values of *SubF*1, *SubF*2, and *F*3 decreased with age over the time period. The decrease rates were similar in *SubF*1 and *SubF*2, the decreasing rates in *SubF*1 and *SubF*2 were larger than that in *F*3. These data will be discussed later in Section 4.5.

Table 4.3 The estimated values for the developmental slope across children in the ages 2 years and 6 months to 3 years and 8 months. The data were fitted into a liner regression model in the form of $y = ax + b$, where *x* is age in months.

y	***a*** **(slope, Hz/month)**
***F*3**	-4.4
***SubF*1**	-10.1
***SubF*2**	-11.2

In sum, for ten children ages between 2;3 – 3;9 years, the first and second subglottal resonances were estimated from speech data. The mean values of *SubF*1 were in the range of 900 – 1100 Hz, while those of *SubF*2 were between 2200 – 2700 Hz. In this time period, the first and second subglottal resonances decreased with age. The subglottal resonances were different for each individual.

4.3 Experiment #2 - Formant Frequency Changes in relation to Subglottal Resonances

4.3.1 Method

Acoustic Measurements

Formant Frequencies

For the ten children, *F*1 and *F*2 frequencies were measured from the low-front vowel /æ/ (e.g., in "cat") and from the low-back vowel /ɑ/ (e.g., in "bottle") or from the /ɑ/ part of the diphthongs (e.g., in "pie") at Time 1 and Time 2. The measurements were manually obtained from the locations of the *F*1 frequency maximum corresponding to the vowel landmarks (Stevens, 1998). The vowels /æ/ and /ɑ/ were chosen because these vowels are the closest to each other in the vowel quadrilateral with a feature contrast of [back]. This selection allows us to examine the developmental properties of quantal relations in more detail with age. The total number of tokens for *F*1 and *F*2 frequency analysis was 761, and the summary of the number of tokens is given in Table 4.4. First, from the spectrograms, the locations of the formants were broadly determined. Second, the values of the formant frequencies were obtained using the 12^{th} order LPC with a 32 ms Hamming window.

Table 4.4 The number of measurements for *F*1 and *F*2 of vowels /ɑ/ and /æ/ at Time 1 and Time 2. C01: Child 1, C02: Child 2 and so on. Except for C05, the age of a child speaker increases, as the number increases. We follow the notation of child subjects in Imbrie's database (2005).

		C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	Total
Time 1	**/æ/**	19	20	27	21	31	16	29	19	21	21	224
	/ɑ/	16	12	17	17	21	12	23	12	9	14	153
Time 2	**/æ/**	25	19	22	18	24	29	18	29	29	16	229
	/ɑ/	22	13	9	14	13	23	17	14	19	11	155
	total	82	64	75	70	89	80	87	74	78	62	761

Statistical Analysis

*F*1, *F*2 Shift over a Six Month Period (Linear Regression and t-test)

As described above, similarly, the mean frequencies of *F*1 and *F*2 across children for Time 1 and Time 2 were fitted to a liner regression line, to see the overall trend of the changes in the frequencies of *F*1 and *F*2 across children in the age range of 2;6-3;9 years. In addition, for each child, *p* values were calculated between a pair of data from Time 1 and from Time 2 to test if there are statistically meaningful changes over a 6-month period in *F*1 or *F*2.

Relations between Formant Frequencies and Subglottal Resonances

Percentage Following Quantal Relations

According to the quantal relations between formant frequencies and subglottal resonances in the vowels, *F*2 frequencies of the back vowel /ɑ/ are expected to be lower than *SubF*2 of the speaker, while *F*2 frequencies of the front vowel /æ/ are expected to be higher than *SubF*2**.** Similarly, *F*1 frequencies of low vowels /ɑ/ and /æ/ should be higher than *SubF*1. To quantify how *F*1 and *F*2 frequencies change in relation to *SubF*1 and *SubF*2 over a six month period, the percentage of tokens following quantal rules was calculated at Time 1 and Time 2 for each child.

The percentage following quantal relations can be defined as follows:

% following quantal relations in *F*2 for the vowel /æ/

$$= \frac{\text{The number of /ae/ tokens whose F2 is above SubF2}}{\text{The total numer of /ae/ tokens}} \times 100$$

% following quantal relations in *F*1 for the vowel /æ/

$$= \frac{\text{The number of /ae/ tokens whose F1 is above SubF1}}{\text{The total numer of /ae/ tokens}} \times 100 .$$

% following quantal relations in *F*2 for the vowel /ɑ/

$$= \frac{\text{The number of /a/ tokens whose F2 is below SubF2}}{\text{The total numer of /a/ tokens}} \times 100 .$$

% following quantal relations in $F1$ for the vowel /ɑ/

$$= \frac{\text{The number of /a/ tokens whose F1 is above SubF1}}{\text{The total numer of /a/ tokens}} \times 100 .$$

4.3.2 Results

A. *F*1, *F*2 Frequency Changes

Across Children

The mean values of *F*1 and *F*2 and their standard deviations for vowels /æ/ and /ɑ/ are summarized in Table 4.5 and Table 4.6. The measurements were obtained at Time 1 and Time 2 for each child. As shown in Figure 4.6**,** the mean values of *F*1 and *F*2 of both vowels slightly increased during a 6 month period. However, these changes were not significant (p = 0.20, 0.09 for *F*1 and *F*2 in /æ/, respectively; p = 0.24, 0.8 for *F*1 and *F*2 in /ɑ/, respectively.) Figure 4.7 and Figure 4.8 depict the overall trend of frequency changes of *F*1 and *F*2 for the vowels /æ/ and /ɑ/ across ten children in the age range of 30 to 45 months (2;6 to 3;9 years). The mean values of *F*1 and *F*2 were approximately constant with age.

Table 4.5 Mean *F*1, *F*2 frequencies and standard deviations in parentheses for the vowel /æ/ at Time 1 and Time 2 for each child. The means and standard deviations in the last column are inter-speaker means and deviations.

		C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	Mean
Time 1	***F1***	1124 (139)	1222 (127)	1266 (179)	1154 (203)	1217 (168)	1339 (107)	1205 (111)	1191 (169)	1163 (155)	1301 (101)	1218 (67)
	F2	2292 (188)	2626 (76)	2721 (194)	2256 (220)	2712 (186)	2362 (172)	2314 (226)	2712 (241)	2538 (178)	2501 (134)	2503 (186)
Time 2	***F1***	1342 (117)	1277 (133)	1419 (161)	1229 (299)	1240 (118)	1234 (197)	1235 (108)	1205 (123)	1126 (149)	1279 (183)	1258 (79)
	F2	2586 (166)	2614 (118)	2715 (238)	2452 (193)	2727 (152)	2530 (183)	2251 (130)	2648 (118)	2622 (145)	2614 (143)	2575 (140)

Table 4.6 Mean *F*1, *F*2 frequencies and standard deviations in parentheses for the vowel /ɑ/ at Time 1 and Time 2 for each child. The means and standard deviations in the last column are inter-speaker means and deviations.

		C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	Mean
Time 1	***F1***	1185	1318	1482	1222	1303	1338	1258	1247	1276	1344	1297
		(144)	(59)	(146)	(65)	(144)	(123)	(121)	(195)	(136)	(136)	(82)
	F2	1923	2111	2131	1888	2099	2069	2039	2045	2069	1885	2026
		(174)	(149)	(243)	(117)	(144)	(236)	(298)	(163)	(184)	(187)	(93)
Time 2	***F1***	1342	1386	1494	1282	1300	1300	1199	1279	1254	1384	1322
		(117)	(133)	(74)	(130)	(139)	(99)	(101)	(163)	(98)	(151)	(83)
	F2	2017	2046	2088	2085	2058	2313	1745	2027	1970	1978	2033
		(278)	(92)	(252)	(161)	(212)	(254)	(109)	(173)	(122)	(219)	(140)

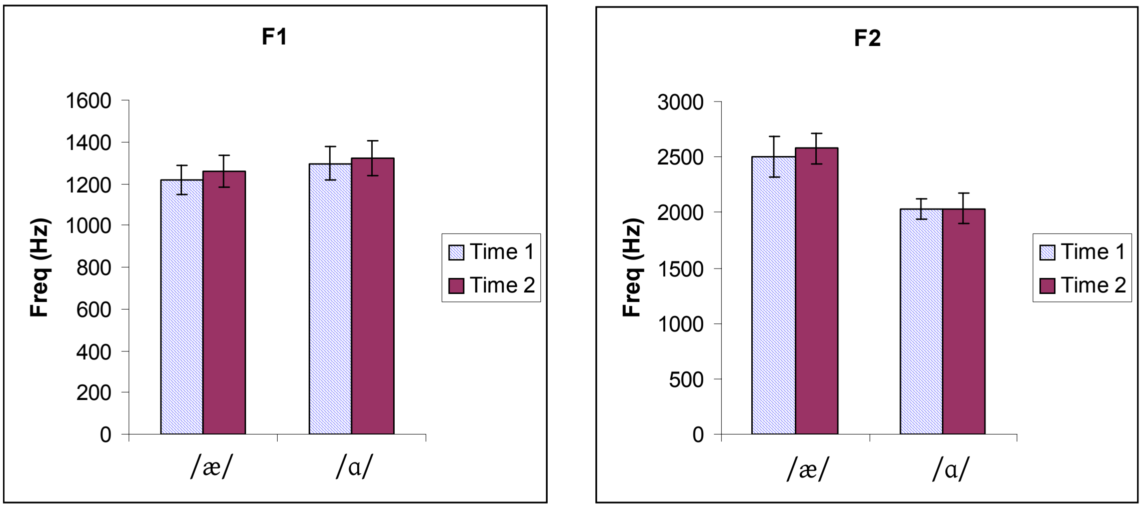


Figure 4.6 The mean values of *F*1 (left) and *F*2 (right) across children for vowels /æ/ and /ɑ/ at Time 1 and Time 2. The changes between Time 1 and Time 2 were not significant.

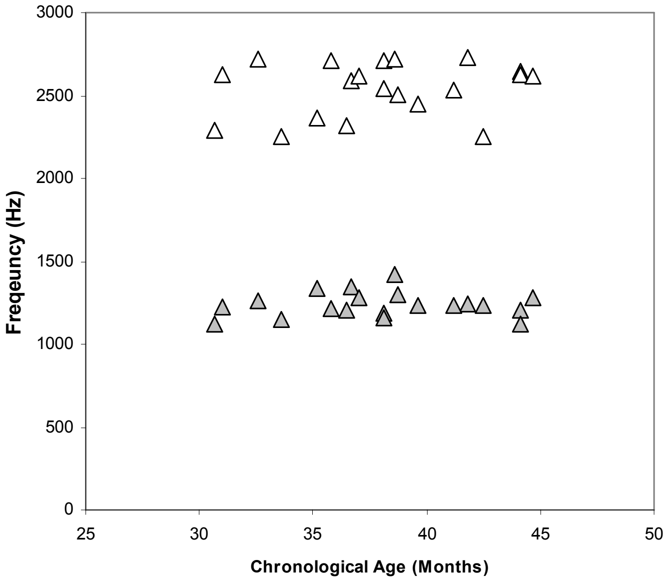


Figure 4.7 The means of *F*1 (shaded triangles) and *F*2 (open triangles) for the vowel /æ/ for children C01-C10 with chronological age in months. Each dot represents the mean values from one subject and from one session.

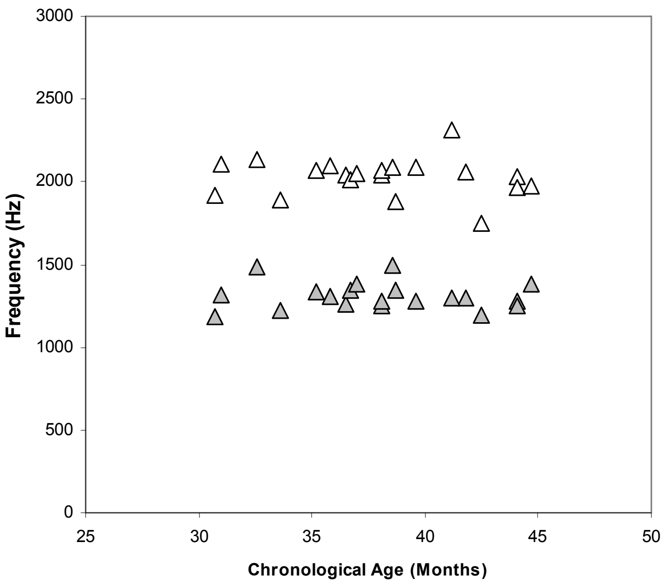


Figure 4.8 The means of *F*1 (shaded triangles) and *F*2 (open triangles) for the vowel /ɑ/ for children C01-C10 with chronological age in months. Each dot represents the mean values from one subject and from one session.

Again, the data of means of $F1$ and $F2$ were fitted into a linear regression model in the form of $y = ax + b$, where y is the value of $F1$ or $F2$, x is the value of age in months, a and b are the estimated parameters (slopes and intercepts, respectively). The estimated values of developmental slope (a) for $F1$ and $F2$ with age are given in Table 4.7. The results show that both of mean values of $F1$ and $F2$ of the vowel /æ/ increased with age, whereas both of mean values of $F1$ and $F2$ of the vowel /ɑ/ decreased with age. These data will be discussed later in Section 4.4.

Table 4.7 The estimated values for the developmental slope for $F1$ and $F2$ across children in the ages 2 years and 6 months to 3 years and 8 months. The data were fitted into a liner regression model in the form of $y = ax + b$, where x is age in months.

	y	***a* (slope, Hz/month)**
/æ/	***F*1**	0.9
	***F*2**	0.83
/ɑ/	***F*1**	-0.9
	***F*2**	-2.8

Individual Differences

For each child, t-test was performed to determine whether the mean values of $F1$ and $F2$ between at Time 1 and at Time 2 were significantly different or not. The mean value changes of $F1$ and $F2$ during a 6-month period for the vowels /æ/ and /ɑ/ are summarized in Figure 4.9 and Figure 4.10. Table 4.8 shows the p-values between the distributions of $F1$ and $F2$ frequencies at Time 1 and those at Time 2. The results turned out that $F1$ or $F2$ frequency significantly shifted ($p < 0.05$) over a 6-month period for some of children (C01, C03, C04, C06, C07 and C10); for other children (C02, C05, C08 and C09) these changes were not significant both in $F1$ and in $F2$. It is note worthy that in most of cases of significant changes (6/7 cases for /æ/, 3/4 cases for /ɑ/) $F1$ or $F2$ frequencies increased over the period.

Ten children were divided into two groups depending on their *p*-values as follows:

(1) Significant Change Group: children showing significant changes either in *F*1 or in *F*2 during a 6-month period.

(2) Non-significant Change Group: children showing non-significant changes both in *F*1 and *F*2 during a 6-month period.

It is noteworthy that the values of *F*1 or *F*2 of the vowel /æ/ significantly increased during the six month period for the children C01, C03, C04, C06 and C10. This observation is inconsistent with a prediction based on the anatomical growth in the vocal tract length. That is, the formant frequencies of *F*1 or *F*2 may increase as the vocal tract length increases with age.

Table 4.8 P-values of *F*1 and *F*2 of vowels /æ/ and /ɑ/ for each child for the test whether the *F*1 or *F*2 frequency values between Time 1 and Time 2 were different or not. (*) mark indicates there was a statistically significant difference in *F*1 or *F*2 between at Time 1 and at Time 2 with a significant level = 0.05.

		C01*	**C02**	**C03***	**C04***	**C05**	**C06***	**C07***	**C08**	**C09**	**C10***
/æ/	***F*1**	<.01*	0.19	<.01*	0.37	0.21	0.03*	0.35	0.57	0.40	0.67
	***F*2**	<.01*	0.71	0.92	<.01*	0.22	<.01*	0.27	0.29	0.09	0.02*
/ɑ/	***F*1**	<.01*	0.11	0.78	0.14	0.77	0.37	0.09	0.66	0.67	0.78
	***F*2**	0.21	0.21	0.68	<.01*	0.77	<.01*	<.01*	0.79	0.17	0.46

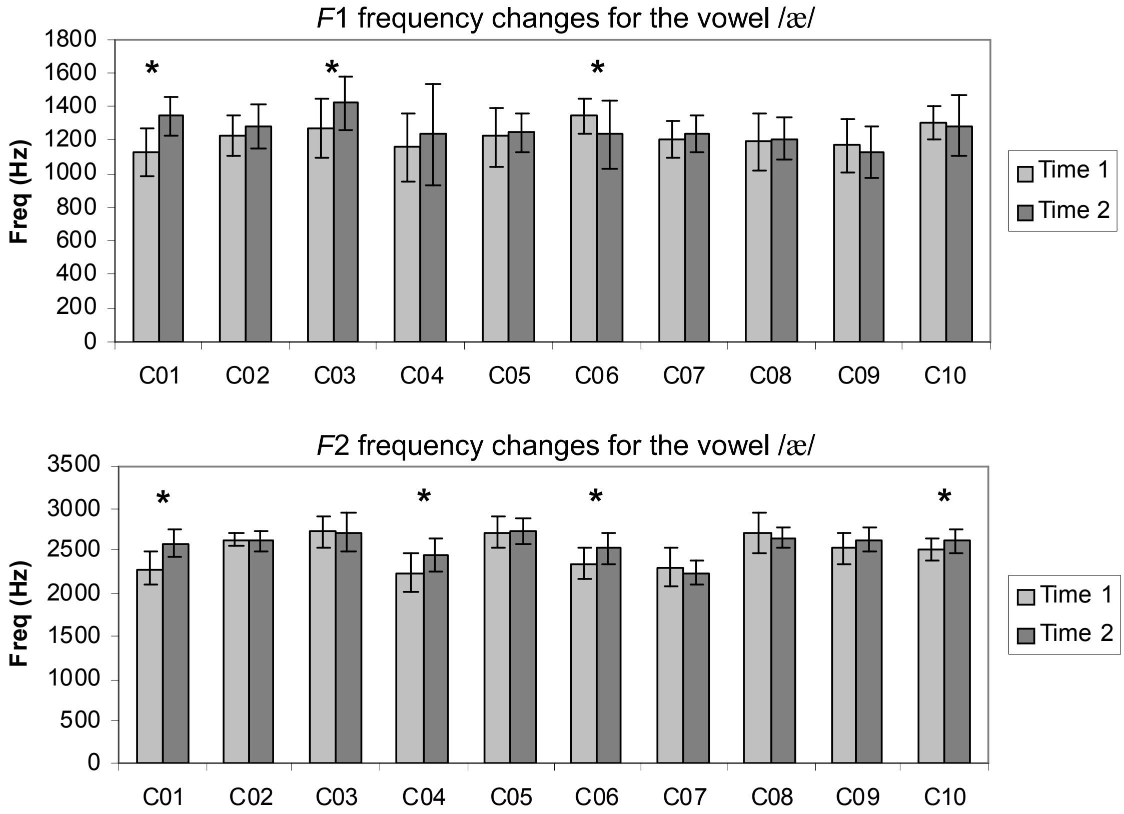


Figure 4.9 Mean *F*1 (upper) and mean *F*2 (lower) frequency changes over a 6-month period for the vowel /æ/ for children C01-C10. (*) There were significant changes in *F*1 or *F*2 between at Time 1 and at Time 2 ($p < 0.05$).

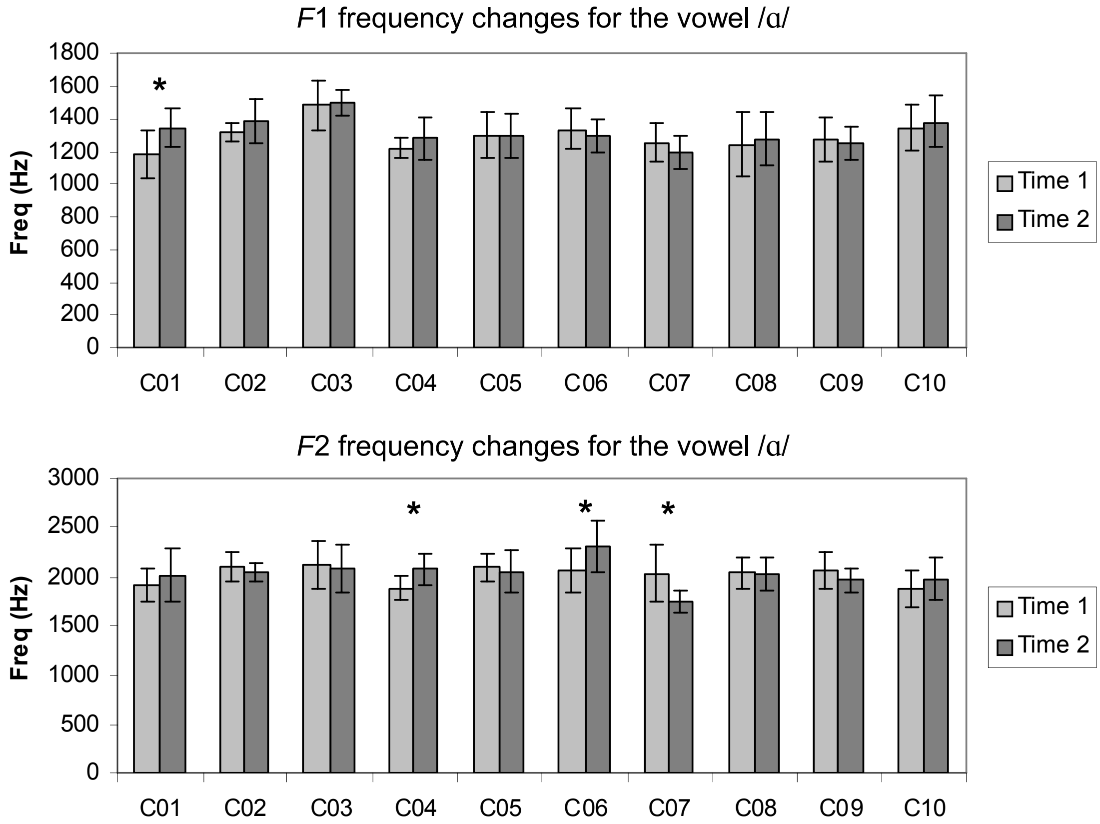


Figure 4.10 Mean *F*1 (upper) and mean *F*2 (lower) frequency changes over a 6-month period for the vowel /ɑ/ for children C01-C10. (*) There were significant changes in *F*1 or *F*2 between at Time 1 and at Time 2 ($p < 0.05$).

B. Relations between Formant Frequencies and Subglottal Resonances

The *F*1 and *F*2 changes over a 6-month period were examined in relation to speaker's *SubF*1 and *SubF*2. The ideal expectations based on quantal relations are: (1) *F*2 of /æ/ is higher than the speakers' *SubF*2, (2) *F*2 of /ɑ/ is lower than the speakers' *SubF*2 and (3) *F*1 of /æ/ or /ɑ/ is higher than the speaker's *SubF*1.

Figure 4.11 shows subject C01's *F*1-*F*2 plots for the vowels /æ/ and /ɑ/ along with her subglottal resonances at Time 1 and Time 2 with ellipses of 1.5 standard deviations. This child belonged to Significant Change Group. The figure on the left shows Time 1, age of

30 months (2;6 years), while the right figure shows Time 2, age of 36 months (3;0 years). Vertical lines are *SubF*1 at Time 1 and at Time 2 of the speaker, whereas horizontal lines are *SubF*2. For the vowel /æ/, many points were below *SubF*2 and below *SubF*1 at Time 1. However, both of *F*1 and *F*2 increased during a 6 month period, and they became above *SubF*2 and above *SubF*1 at Time 2. For the vowel /ɑ/, at Time 1, many points were below *SubF*1. However, *F*1 increased during the period, they became above *SubF*1, six months later. Compared with the ideal expectations based on quantal relations, the distributions of *F*1 and *F*2 shifted toward the ideal expectations during the period.

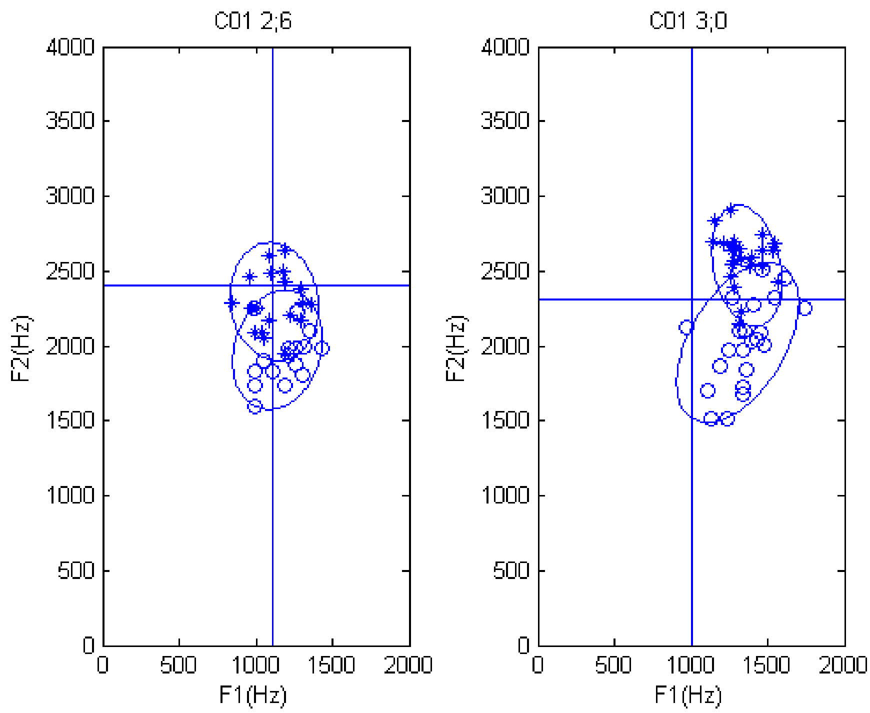


Figure 4.11 *F*1-*F*2 plot for vowels /æ/ and /ɑ/ for a subject C01 at Time 1 (left) and at Time 2 (right), with ellipses of 1.5 standard deviations. (Asterisk: vowel /æ/, Circle: vowel /ɑ/; Vertical line: *SubF*1 of C01, Horizontal line: *SubF*2 of C01; Time 1: age 2;6 years, Time 2: age 3;0 years.)

Figure 4.12 shows another example from the subject C02, who belonged to Non-significant Change Group. Left figure was obtained at ate age of 31 months, while right figure was obtained at the age of 37 months. In the left figure at Time 1, the spaces of the vowel /æ/ and /ɑ/ were already well separated along with *SubF*2. However, for *F*1 frequencies, several points were below *SubF*1. Six months later, this quantal relation was maintained and became more aligned with subglottal resonances by the increase of *F*1 frequencies. The *F*1-*F*2 plots for the rest of the children are given in Appendix C. Both for Significant Change Group and for Non-significant Change Group, the distributions of *F*1 or *F*2 frequencies changed toward quantal relations during a 6-month period. In Non-significant Change Group, quantal relations were already achieved at Time 1. However, more fine tunings were observed at Time 2. In Significant Change Group, the formant frequencies were deviated from the expectation at Time 1. However, the *F*1 and *F*2 frequencies shifted toward the expected positions at Time 2.

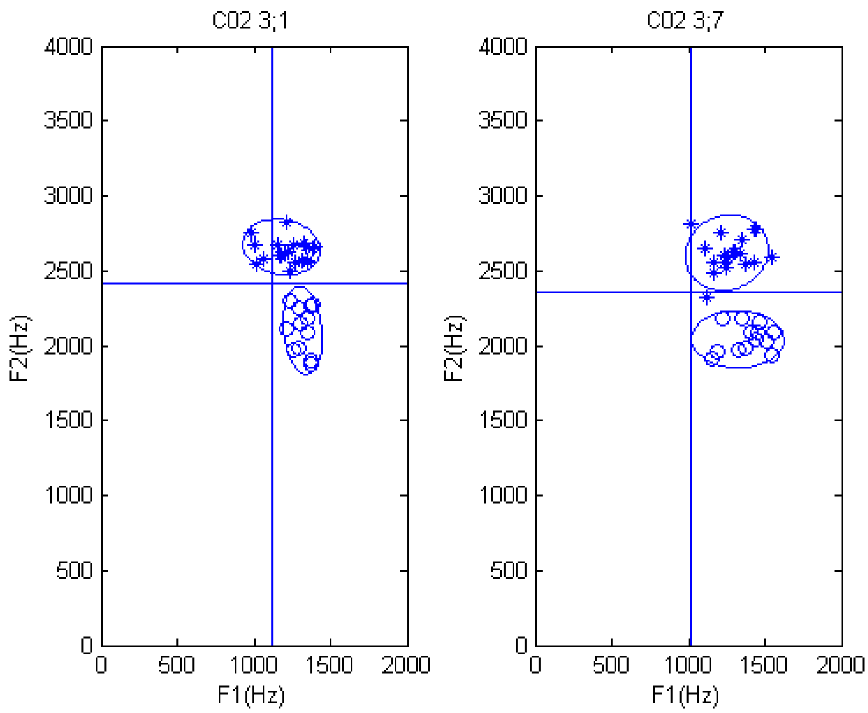


Figure 4.12 *F*1-*F*2 plot for vowels /æ/ and /ɑ/ for a subject C01 at Time 1 (left) and at Time 2 (right), with ellipses of 1.5 standard deviations. (Asterisk: vowel /æ/, Circle: vowel /ɑ/; Vertical line: *SubF*1 of C02, Horizontal line: *SubF*2 of C02; Time 1: age 3;1 years, Time 2: age 3;7 years.)

The mean formant frequency changes in relation to the speaker's subglottal resonances are also analyzed. Figure 4.13 shows the mean values of *F*1 and *F*2 in relation to *SubF*1 and *SubF*2 at Time 1 and Time 2 for the subject C04 in Significant Change Group. At Time 1, the *F*2 mean for /æ/ was below *SubF*2. However, six months later, the mean *F*2 became higher than *SubF*2. There is a crossing point of *F*2 and *SubF*2 at around 37 months between Time 1 and Time 2, which may correspond to the transition point into quantal relations. Similar pattern were also observed in C04 and C05, both of whom belonged to Significant Change Group.

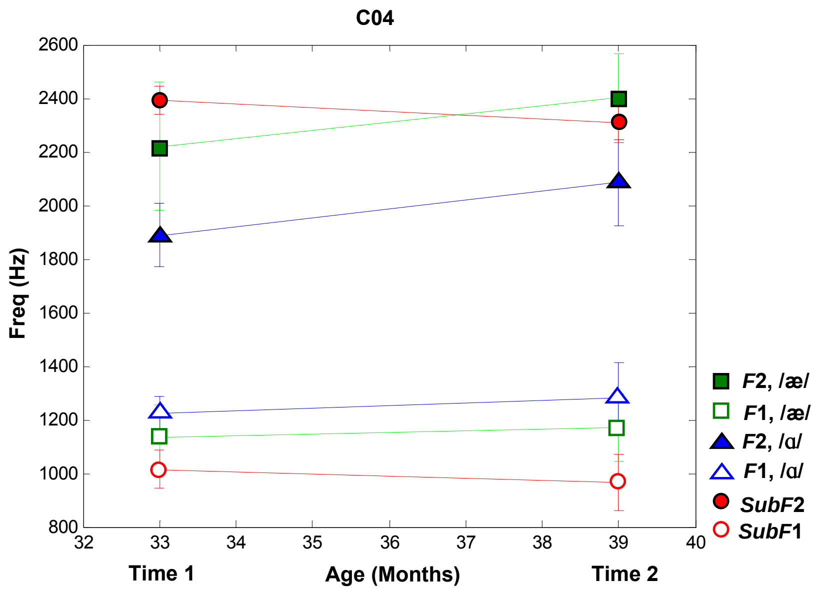


Figure 4.13 *F*1, *F*2, *SubF*1, and *SubF*2 frequency changes with age for the subject C04 in Significant Change Group. (Square: /æ/, Triangle: /ɑ/, circle: subglottal resonances; shaded: *F*2, open: *F*1; Time 1: 2;9 years, Time 2: 3;3 years)

For Non-significant Change Group, the mean frequency values of *F*1 and *F*2 already follows quantal relations at Time 1. At time 2, the F2 formant frequencies of back and those of front vowels were more separated in the boundary of *SubF*2. The *F*1 frequencies became more separated from *SubF*1. One of the examples is given in Figure 4.14, which shows the mean values of *F*1 and *F*2 in the relation to *SubF*1 and *SubF*2 at Time 1 and Time 2 for the subject C09 who belonged to Non-significant Change Group. At Time 1, the distributions of *F*2 of the vowel /æ/ were overlapped with *SubF*2. However, six months later these distributions were well separated from *SubF*2. The similar fine tunings in *F*1 were observed in C02 and C08.

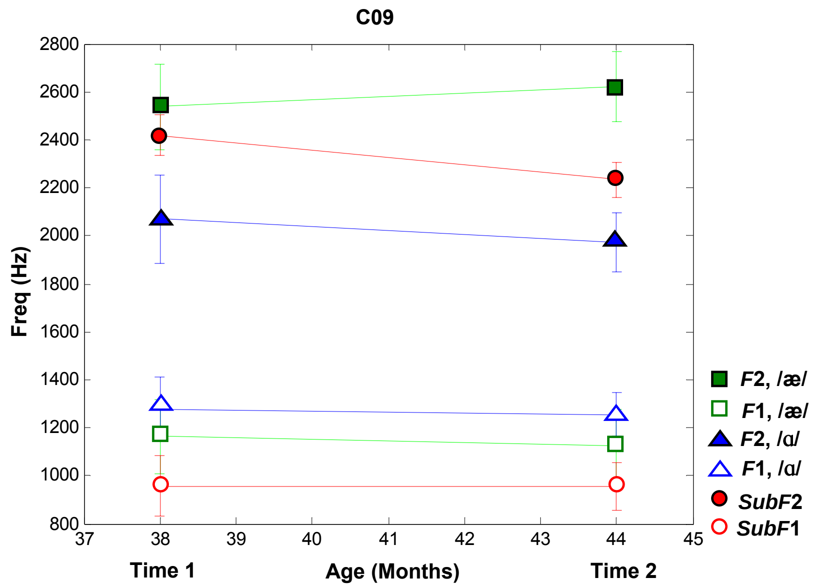


Figure 4.14 *F*1, *F*2, *SubF*1, and *SubF*2 frequency changes with age for the subject C09 in Non-significant Change Group. (Square: /æ/, Triangle: /ɑ/, circle: subglottal resonances; shaded: *F*2, open: *F*1; Time 1: 3;2 years, Time 2: 3;8 years)

Percentage Following Quantal Relations

To quantify how *F*1 and *F*2 change in relation to *SubF*1 and *SubF*2 over a 6 month period, the percentage of tokens following quantal rules were calculated at Time 1 and Time 2 for each child, as described in Section 4.2.

Figure 4.15 shows the percentage following quantal relations in the vowel /æ/ at Time 1 and Time 2 for Significant Change Group and for Non-significant Change Group. For all children in Significant Change Group, the percentage following quantal relations largely increased during a 6-month period. The percentage following quantal relations in the vowel was in the range between 40 – 86 % at Time 1, whereas the percentage at Time 2 was in the range 69 – 97 %. This implies that at Time 1, the formant frequencies of the vowel /æ/ for Significant Change Group were deviated from expected quantal relations. However, the formant frequencies shifted in the right direction during the period. As for Non-significant Change Group, however, the percentage following quantal relations was already high at Time 1. The values were in the range of 84 – 92 %. At Time 2, these values were remained in the similar range of 88 – 95 %. However, there was a slight increase of the percentage value.

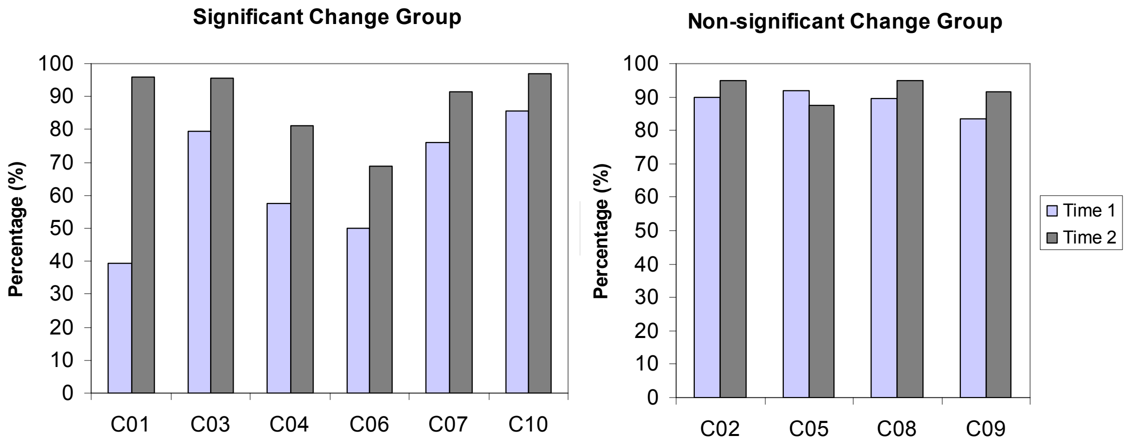


Figure 4.15 The percentage following quantal relations in the vowel /æ/ at Time 1 and Time 2 for Significant Change Group (left) and for Non-significant Change Group (right).

The percentages following quantal relations in the vowel /ɑ/ are shown in Figure 4.16. Both for Significant Change Group and for Non-significant change group, the percentages following quantal relations were already high at Time 1. The values were in the range 85 – 100 %. At Time 2, these values were remained in the same rage. There was little change of the percentage value.

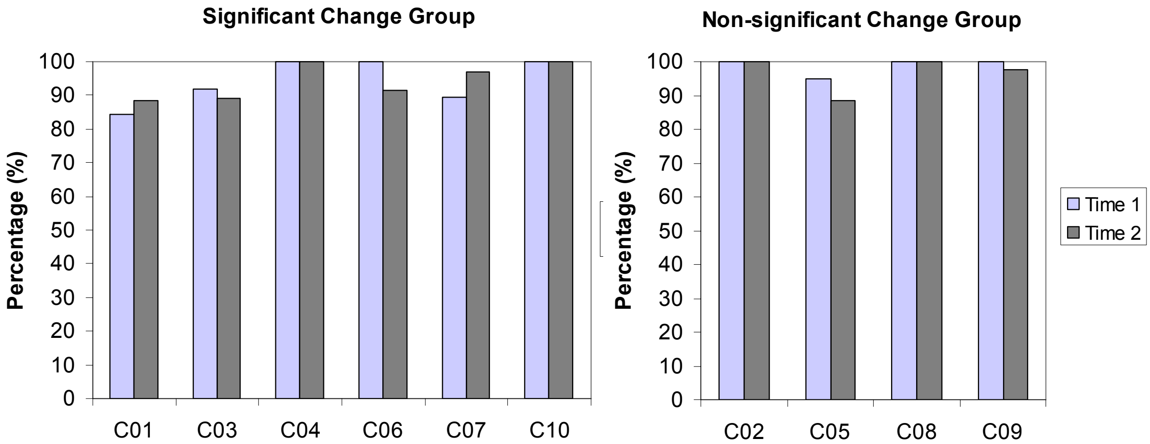


Figure 4.16 The percentage following quantal relations in the vowel /ɑ/ at Time 1 and Time 2 for Significant Change Group (left) and for Non-significant Change Group (right).

The percentage changes following quantal relations for the vowel /æ/ and for the vowel /ɑ/ with age from ten children are given in Figure 4.17. For the vowel /æ/, the percentage following quantal relations approximately increased with age, while the percentage for the vowel /ɑ/ was already very high at the early ages. The percentage for the vowel /ɑ/ was higher than the percentage for the vowel /æ/ over all age ranges. The percentage values for the vowel /æ/ became similar ranges with those for the vowel /ɑ/ at around 36 months (3 years).

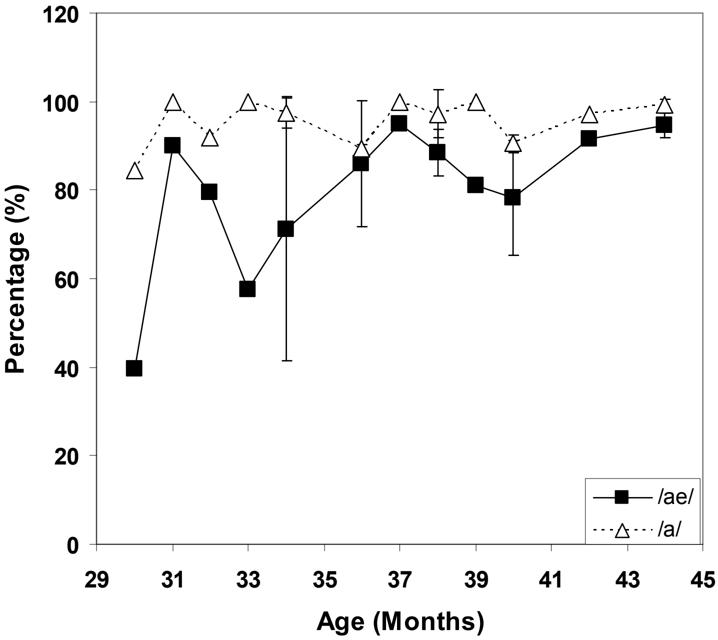


Figure 4.17 The percentage following quantal relations for the vowel /æ/ (square) and the vowel /ɑ/ (triangle) with age from ten children.

Figure 4.18 shows the percentage changes following quantal relations for *F*1 and for *F*2 with age from ten children. For *F*2 frequencies, the percentage following quantal relations approximately increased with age. The percentage following quantal relations also increased in *F*1 frequencies. The percentage values in *F*1 reached to 85% at earlier age about 32 months than those in *F*2. Beyond about 37 months (3;1 years), the percentages both in *F*1 and in *F*2 became in the similar ranges over 85%.

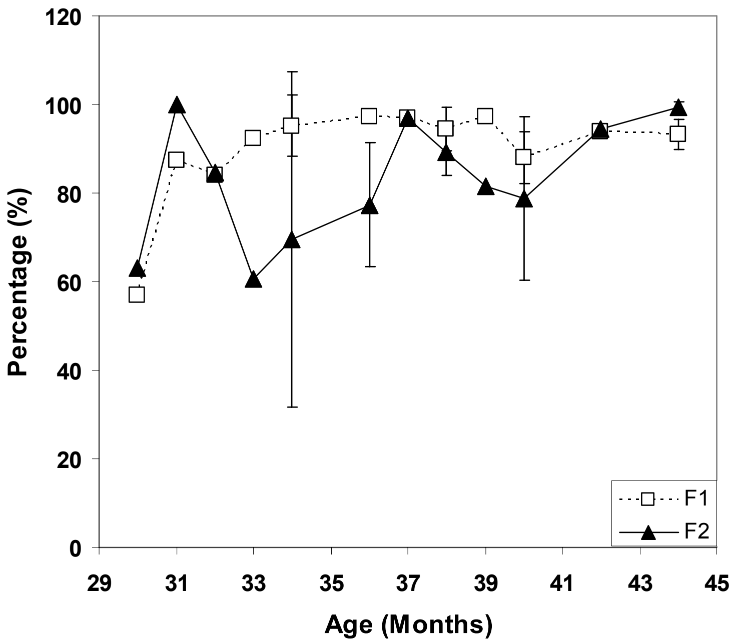


Figure 4.18 The percentage following quantal relations for *F*1 (square) and *F*2 (triangle) with age.

Figure 4.19 compares the development speed between in Significant Change Group and in Non-significant Change Group for the vowel /æ/. Since, as shown in Figure 4.17, in the vowel /æ/ there was an increase of percentage following quantal relations with age, the changes of percentages in the vowel /æ/ were examined in more detail with age depending on which group the children belonged to. The children in Non-significant Change Group showed high percentages over 85% at 31 months. The high percentage was maintained in the age range of 31 – 44 months. On the other hand, the percentage in Significant Change Group approximately increased with age. The minimum and maximum values were 40 and 97 %, respectively. Until 36 months, the percentage value sharply increased for Significant Change Group. Above 36 months, the value became to be in the similar range with that for Non-significant Change Group. Except for two measurement points from the children aged 31 months (C02 at Time 1) and 34 months (C05 at Time 1) in Non-significant change group, the percentage values showed this trend: the percentage largely increased with age until 36 months, and beyond 36 months the percentage change became small and saturated.

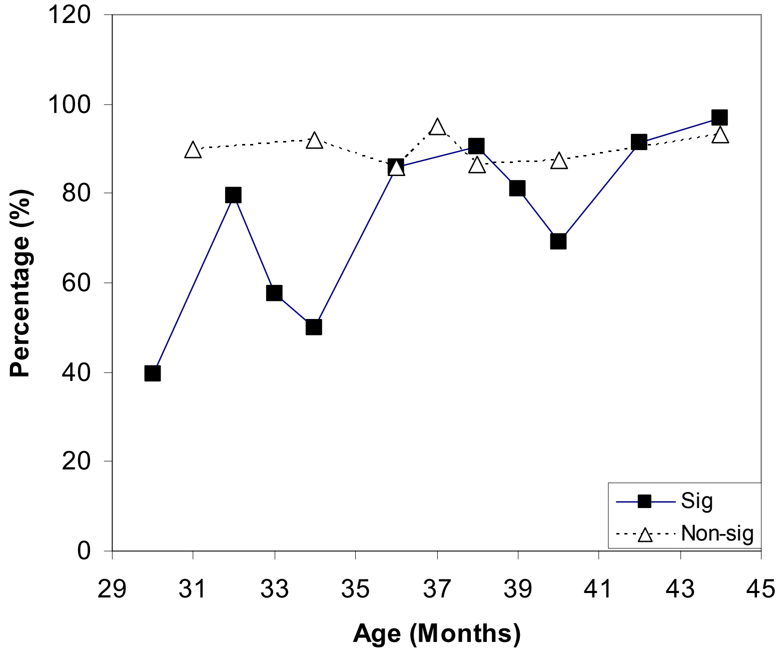


Figure 4.19 The percentage following quantal relations for the vowel /æ/. Significant Change Group vs. Non-significant Change Group

This may imply that the development speed of these two children is faster than other children. From the graph, the percentage from C02 at Time 1 (31 months) is similar with the percentage at around 36 months from other children, which might imply that her development speed was at least about 5 months faster than others. Locally minima in the percentage for Significant Change Group correspond to the measurements from the same subject C06. For this child, the development speed was estimated to be about 6 months slower than other children. Based on these observations, quantal relations for the vowel /æ/ in the children of Significant Change group appeared at about 3 years.

In sum, $F1$ and $F2$ changes were relative to the speaker's subglottal resonances in this age range. The quantal development between $F1$ and $SubF1$ and between $F2$ and $SubF2$ was observed. The transition to the expected quantal relation appeared at around 3 years.

4.4 Experiment #3 Predictions: Vocal Tract Growth vs. Quantal Relations

The commonly known hypothesis about formant changes with age is that the formant frequencies decrease linearly as the vocal tract length increases (Kent, 1976). The prediction based on this hypothesis was compared with the prediction based on quantal relations in this experiment. How the vowel formants would change 6 month later was predicted using the method based on (1) quantal relations, and based on (2) vocal tract growth. Then the prediction results were compared between them.

4.4.1 Method

Prediction Based on Quantal Relations

The prediction method based on quantal relations is illustrated in Figure 4.20. To predict the data for Time 2, the distributions of $F1$ and $F2$ at Time 1 were shifted so that the center of each distribution was a certain amount of distance away from the subglottal resonances at Time 2. The shifted center for Time 2 (Mean $F1_{\text{Time 2}}$, Mean $F2_{\text{Time 2}}$) was calculated as follows:

$$\text{Mean } F1_{\text{Time 2}} = SubF1_{\text{Time 2}} + \beta_1 \text{ Std } F1_{\text{Time 1}}.$$

$$\text{Mean } F2_{\text{Time 2}} = SubF2_{\text{Time 2}} + \beta_2 \text{ Std } F2_{\text{Time 1}}, \text{ for the vowel /æ/.}$$

$$SubF2_{\text{Time 2}} - \beta_2 \text{ Std } F2_{\text{Time 1}}, \text{ for the vowel /ɑ/.}$$

,where $SubF1_{\text{Time 2}}$ is $SubF1$ at Time 2, Std $F1_{\text{Time 1}}$ is the standard deviation of $F1$ at Time 1, $SubF2_{\text{Time 2}}$ is $SubF2$ at Time 2, Std $F2_{\text{Time 1}}$ is the standard deviation of $F2$ at Time 1,

$0 < \beta_1, \beta_2 < 2$, depending on the degree of children's quantal development at Time 1.

The distributions of formants at Time 2 were assumed to be the same as Time 1. $F1$ and $F2$ frequencies were randomly generated by Gaussian distributions with these mean and variance values. These were used as the predicted data for comparison with the real data at Time 2.

Prediction Based on Vocal Tract Growth

To predict the data based on vocal tract growth, the distribution of $F1$ and $F2$ at Time 1 was shifted in relation to vocal tract growth. Because $F3$ frequencies were obtained in the vowel /ʌ/, the shape of the vocal tract was assumed to be an approximately uniform tube. The detailed procedure is explained in Figure 4.21. The centers of $F1$ and $F2$ distributions at Time 1 were shifted by the scaling factor α, which was determined by the ratio of the $F3$ mean at Time 2 to the $F3$ mean at Time 1. The standard deviations of formants at Time 2 were assumed to be the same as Time 1. $F1$ and $F2$ frequencies were randomly generated by Gaussian distributions with these mean and variance values. These were used as the predicted data for comparison with the real data at Time 2.

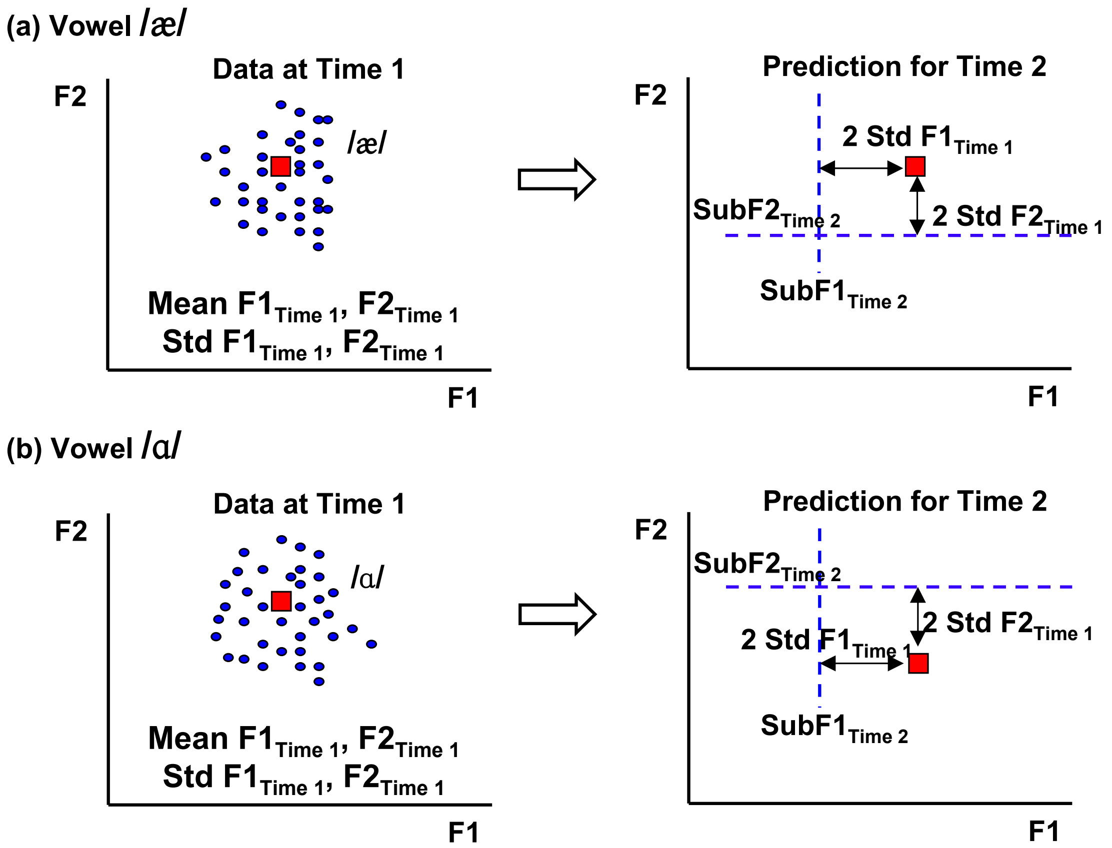


Figure 4.20 Prediction method based on quantal relations. (a) Predictions for the vowel /æ/. The left figure depicts the data of *F1* and *F2* frequencies measured at Time 1. Each dot represents the frequencies of *F1* and *F2* of each token. The red square at the center represents the mean values of the measured data. From the data, the standard deviations of *F1* and *F2* frequencies ($Std\ F1_{Time\ 1}$, $Std\ F2_{Time\ 1}$) were obtained. The right figure depicts the prediction for Time 2 for the same child. The vertical dashed line is the first subglottal resonance at Time 2 ($SubF1_{Time\ 2}$). The horizontal dashed line is the second subglottal resonance at Time 2 ($SubF2_{Time\ 2}$). The red square at the center represents the shifted center distributions of *F1* and F2 to be two standard deviations away from the first and second subglottal resonances at Time 2 ($\beta_1 = \beta_2 = 2$.) For the vowel /æ/**,** the mean of *F2* at Time 2 is above $SubF2_{Time\ 2}$. (2) Predictions for the vowel /ɑ/. The left figure depicts the data of *F1* and *F2* frequencies measured at Time 1. The right figure depicts the prediction for Time 2. The procedure is similar to the procedure for vowel /æ/**.** However, for the vowel /ɑ/**,** the mean of *F2* at Time 2 is below $SubF2_{Time\ 2}$.

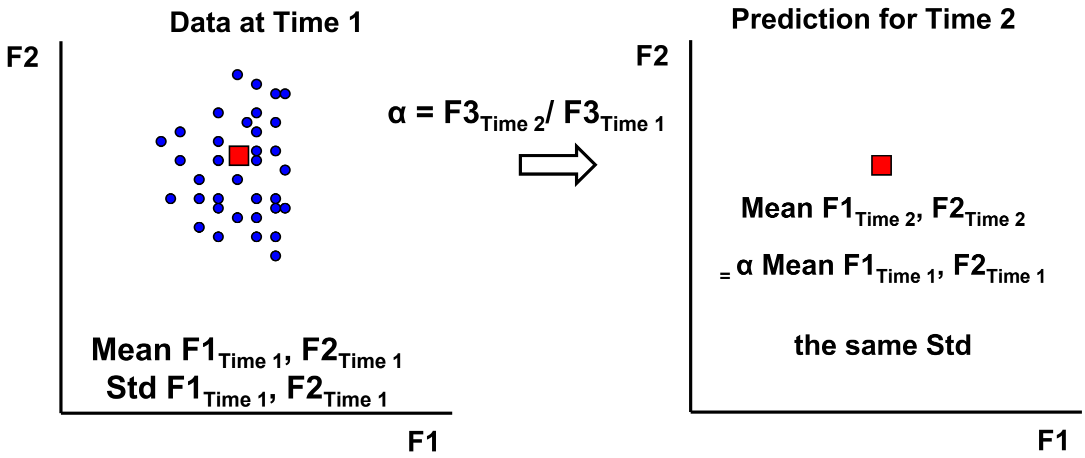


Figure 4.21 Prediction method based on the vocal tract growth. The left figure depicts the data of *F*1 and *F*2 frequencies measured at Time 1. Each dot represents the frequencies of *F*1 and *F*2 of each token. The red square at the center represents the mean values of the measured data. Mean $F1_{\text{Time 1}}$,$F2_{\text{Time 1}}$ are the mean of F1 and the mean of *F*2 at Time 1. From the data, the standard deviations of *F*1 and *F*2 frequencies (*Std* $F1_{\text{Time 1}}$, *Std* $F2_{\text{Time 1}}$) were obtained. The right figure depicts the prediction of the center of *F*1, *F*2 distributions for Time 2, for the same child. The red square at the center represents the shifted center distributions of *F*1 and *F*2. The shifted ratio, α was determined by the ratio of the *F*3 mean at Time 2 to the *F*3 mean at Time 1. The centers of the distributions of *F*1 and *F*2 frequencies for Time 2 were shifted by the ratio. The standard deviations of predicted data for Time 2 were the same as the standard deviations of real data at Time 1.

4.4.2 Results

As shown above, the *F*1 and *F*2 frequencies significantly increased for several children during a 6 month period. The results are opposite to the predictions based on the commonly known hypothesis that the formant frequencies linearly decrease with age. Therefore, the prediction based on this hypothesis was compared with the prediction based on the quantal hypothesis, as described in Section 4.2. For each child, the *F*1 and *F*2 formant frequencies at Time 2 were predicted based on these two hypotheses, and the prediction results were compared with real data from Time 2.

To predict the data at Time 2 based on quantal relations, the distribution of $F1$ and $F2$ at time 1 was shifted so that the center of the distribution was one or two standard deviations away from the subglottal resonances at Time 2. The value of β_1 was taken to be 2, while the value of β_2 was 1 or 2 depending on the degree of children's quantal development at Time 1. One standard deviation was used when percentages following quantal relations at Time 1 were lower than 80%, which was considered as to be still in progress of quantal development (C01, C03, C04, C06 and C07, see Figure 4.15.) If one sided Gaussian distributions is assumed, about 84% of the data are above subglottal resonances or below subglottal resonances. Two standard deviations were used for children for whom the quantal relations were already achieved at Time 1 (C02, C05, C08, C09 and C10). About 98% of the data were assumed to be in the right side, and there will be little overlap between the vowels /æ/ and /ɑ/.

Figure 4.22 compares the prediction result based on quantal relations with the prediction result based on vocal tract growth. The goodness of the predictions was quantitatively compared by calculating the distance between the center frequency of $F1$ and $F2$ of the real data at Time 2 and the predicted center frequency based on each method.

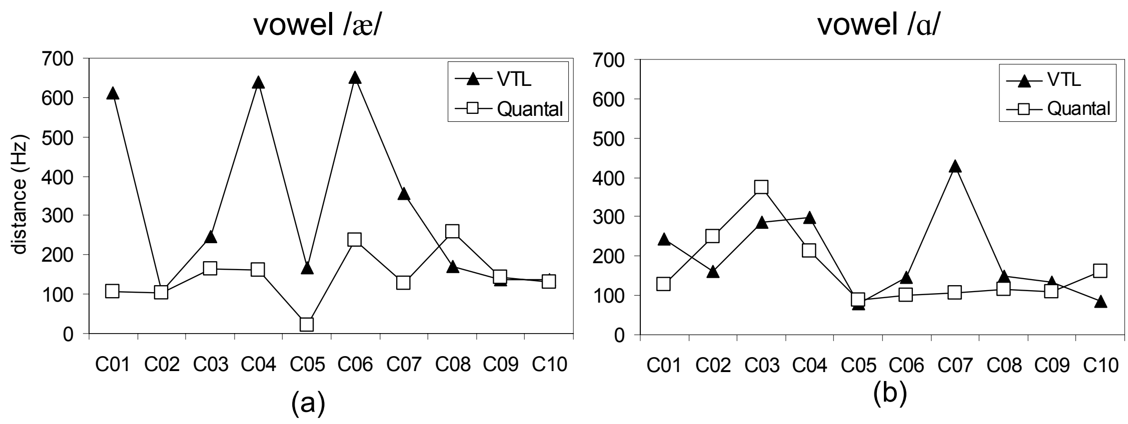
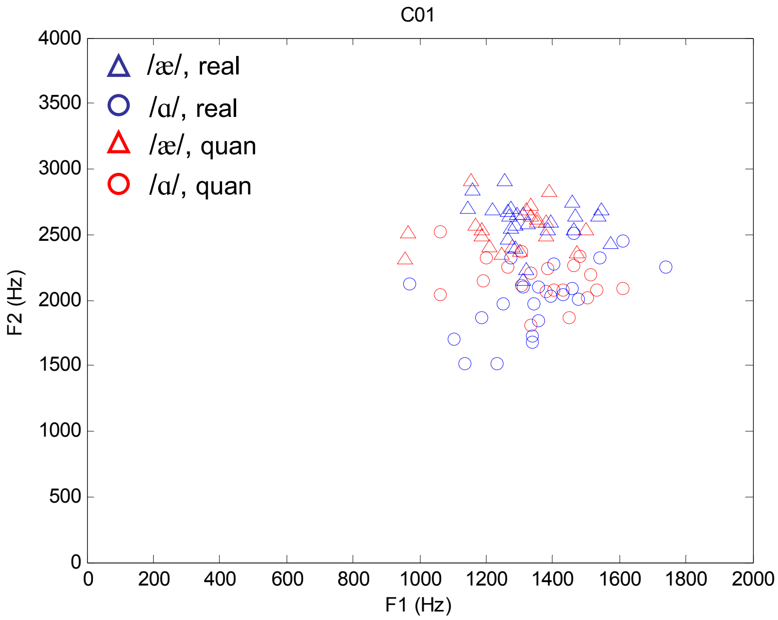


Figure 4.22 The distance between predicted centers and the ground truth centers of real data for the vowels /æ/ and /ɑ/. Triangle represents the distances between predictions based on vocal tract growth and the real data, while square represents the distances between predictions based on quantal relations and the real data.

For the vowel /æ/, the center distances between the prediction and real data were much larger in vocal tract growth hypothesis than in quantal relation hypothesis for all children except for C08. For children C01, C04, C06 and C07, the differences between two prediction results were very large.

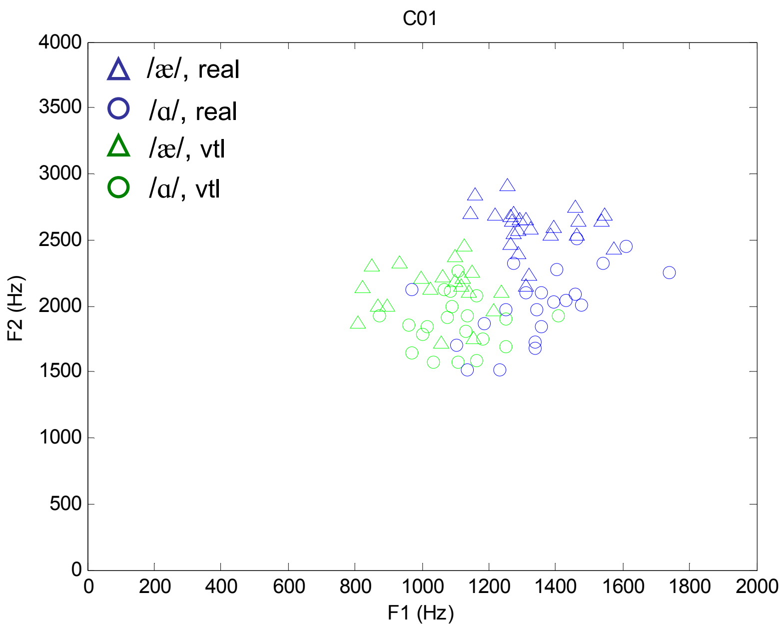
It is noteworthy that all these children belong to Significant Change Group. The results suggest that the prediction based on quantal relations is better than the prediction based on vocal tract growth, for children in this age range. In addition, the significant changes in $F1$ or $F2$ frequencies over a 6 month period cannot be captured by vocal tract growth hypothesis. For the vowel /ɑ/, the differences of center distances between these two predictions were small, as shown in Figure 4.22(b).

Examples of predicted $F1$ and $F2$ distributions based on these two methods are given in Figure 4.23 and Figure 4.24 from two children C01 and C04, respectively. The upper panel in Figure 4.23 compares $F1$-$F2$ plots from the real data and from the prediction based on quantal relations for C01. *P*-values were also calculated between the real data and predicted data. The two distributions did not differ significantly. On the other hand, there were significant differences in the distributions between the real data and the prediction based on vocal tract growth, as shown in the lower panel. All *p*-values were smaller than 0.05. Similar observations were made for C04, as shown in Figure 4.24. There were significant discrepancies between predicted data by vocal tract growth and real data.



		p-value
/æ/	*F*1	0.08
	*F*2	0.38
/ɑ/	*F*1	0.85
	*F*2	0.06

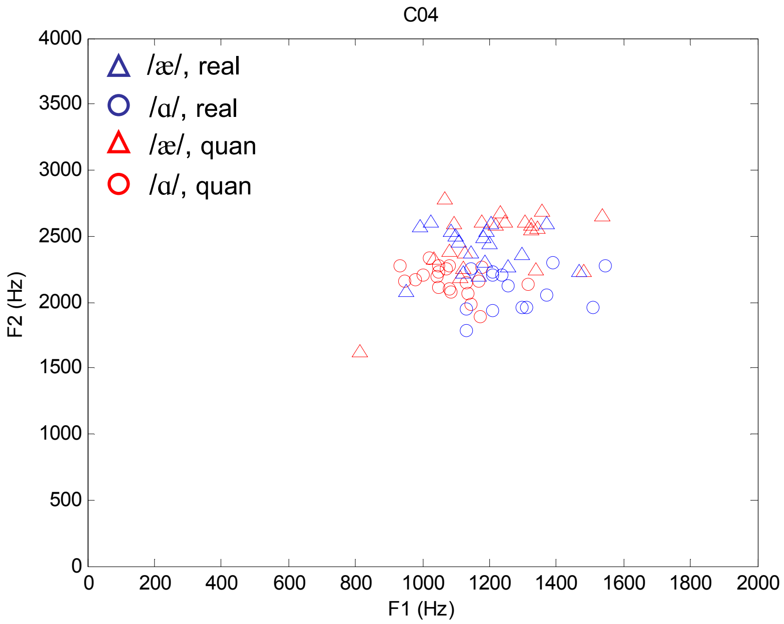
(a)



		p-value
/æ/	*F*1	<.01*
	*F*2	<.01*
/ɑ/	*F*1	<.01*
	*F*2	<.05*

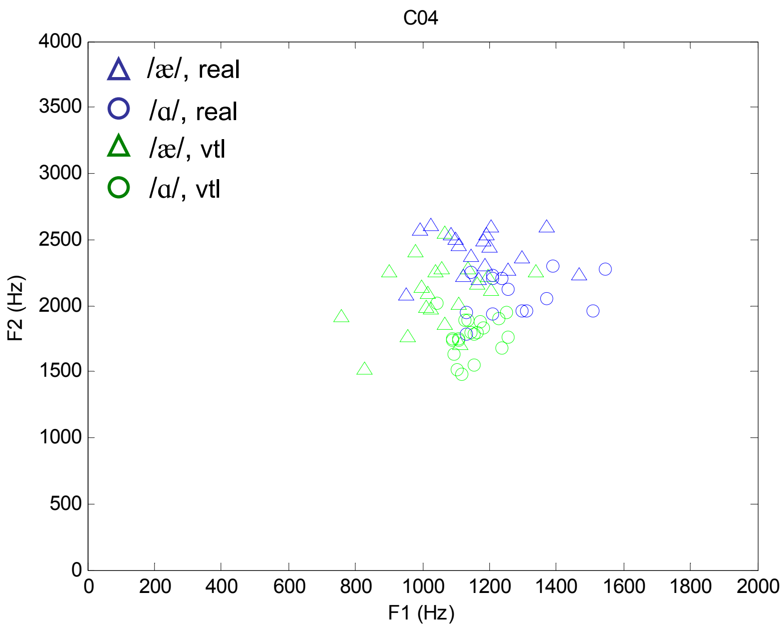
(b)

Figure 4.23. *F*1-*F*2 plots and *p*-values from real data at Time 2 for the subject C01 and from the predictions for them. (a) the real data vs. predictions based on quantal relations (b) the real data vs. predictions based on vocal tract growth.



		p-value
/æ/	*F*1	0.34
	*F*2	0.53
/ɑ/	*F*1	<.01*
	*F*2	0.10

(a)



		p-value
/æ/	*F*1	<.01*
	*F*2	<.01*
/ɑ/	*F*1	<.01*
	*F*2	<.01*

(b)

Figure 4.24 *F*1-*F*2 plots and *p*-values from real data at Time 2 for the subject C04 and from the predictions for them. (a) the real data vs. predictions based on quantal relations (b) the real data vs. predictions based on vocal tract growth.

In sum, the difference between predicted data and the real data for /æ/ was smaller in the prediction based on quantal relations; the difference for /ɑ/ was comparable. In contrast, the differences between prediction based on vocal tract growth and the real data were significant. Thus, the significant changes in $F1$ or $F2$ during a 6 month period could be explained by quantal relations, while vocal tract growth failed to predict.

4.5 Discussion and Conclusion

Previous studies of adult speech have suggested that the subglottal system plays important roles in defining vowel features. The main objective of this study was to explore the role of subglottal resonances in defining the vowel features for young children from ages 2;6 to 3;9 years. For these children, the characteristics of vowel formant frequencies were analyzed to examine how formant frequencies change during a 6 month period relative to child speakers' subglottal resonances, and when quantal relations appear. Three questions were addressed in this study: (i) How do the first and second subglottal resonances change in the age range of 2-3 years? (ii) How do $F1$ and $F2$ change relative to the subglottal resonances in this age range for each child? (iii) Do quantal relations better predict $F1$ and $F2$ changes than vocal tract length?

The data in Section 4.2 show that $SubF1$ and $SubF2$ for these children decrease with age. The mean values of $SubF1$ are in the range of 900 to 1100 Hz, while those of $SubF2$ are in the range of 2200–2700 Hz. These frequency values vary from one individual to another because they depend on the size of the subglottal system. The ranges of the frequency values of $SubF1$ and $SubF2$ agree with those obtained from the data using an accelerometer. Due to the differences in the size of the subglottal system between adult speakers and children speakers, these subglottal resonant frequencies are much higher than those of adult speakers. The mean frequencies of $SubF1$ and $SubF2$ are 600 and 1550 Hz, respectively, for adult male speakers; for adult female speakers $SubF1$ and $SubF2$ are 700 and 1650 Hz, respectively**.** The frequencies of $SubF1$ and $SubF2$ for these children significantly decrease over a 6-month period in this age range. The $F3$ frequencies also decrease with age. The decrease rates in $SubF1$ and in $SubF2$ are similar to each other as shown in Table 4.7, but larger than that in $F3$. This might imply that the anatomical development speed in the vocal tract and in the subglottal system are different, since $F3$ is related to the vocal tract length and the subglottal resonances are related to the subglottal system.

The inter-subject mean of $F3$ at Time 1 (average age = 35 months) was 4200Hz. This value agrees with the $F3$ mean value of children at the ages of 3 year olds reported by Vorperian and Kent (2007). If the configuration of the vocal tract is assumed to be a uniform tube, the relation between $F3$ and the vocal tract length is given as $5c/4l$~F3 where c is the velocity of sound and l is the vocal tract length. The vocal tract length is expected to be in the range of 9.7 to 11.6 cm based on this relationship between l and F3. This range is also in agreement with values more directly measured using MRI, which are 9.6–11.5 cm for children in the age range of 2 to 3 years (Vorperian et al., 2005).

The data in Section 4.3.2 show that both of $F1$ and $F2$ means of children are constant or slightly increase over a 6 month period. These observations are opposite to the expectations based on vocal tract growth. The analysis for individual speakers shows that the $F1$ or $F2$ frequency shift over a 6 month period is significant for six out of ten children, and often in the opposite direction to the decrease of frequencies. The examination of the relationship between the first two formants ($F1$, $F2$) and the first two subglottal resonances ($SubF1$, $SubF2$) reveals that these formant changes are relative to the speaker's subglottal resonances. At the earlier ages before 3 years as shown in Figures 4.15 and 4.16, the $F1$ and $F2$ values deviated from the expected quantal relations. At the later age of 3 years and above, considerable agreement with the expected values of $F1$ and $F2$ in relation to the subglottal resonances was obtained. The transition to the expected quantal relation appeared to occur at around 3 years. Until the quantal relation is achieved, $F1$ or $F2$ significantly changed during a 6-month period (See Figure 4.9-4.11, Figure 4.15 and Appendix C.) These observations cannot be simply explained by vocal tract length changes. For every child, the distribution of the vowel /æ/ and the distribution of the vowel /ɑ/ are more distinctively separated at Time 2. For children whose vowel spaces of /æ/ and /ɑ/ are overlapped at the beginning, apparent separation between the vowel spaces is observed six months later.

Ishizuka and his colleague (2007) pointed out that the $F1$ values for two Japanese children remain unchanged and high between 20 and 40 months for the vowel /ɑ/ even though for other non- low vowels $F1$ decrease, and this result is unexplainable. However, the hypothesis of quantal relation between *F*1 and *SubF*1 predicts that for a low vowel, *F*1 increases until *F*1 is above *SubF*1, and after the period of when quantal relation occurs *F*1 would decrease with vocal tract growth.

We found that the time when *F*1 is in the expected place in relation to *SubF*1 is earlier than the time when *F*2 is in the expected place in relation to *SubF*2. The *F*1 frequencies are related to jaw movements and mouth openings, which are visible in child speakers. Young children may also use this visible information in acquisition of speech sound related to the [high]/[low] feature. On the other hand, *F*2 is related to the backness of the tongue body. Therefore, this [back] feature may be learned based solely on auditory feedback (Imbrie, 2005). In addition, our finding suggests that young children develop acquisition of the vowel /æ/ later than the acquisition of vowel /ɑ/.

In addition, the data in Section 4.3.2 imply considerable variability in the developmental speed from one individual to another. Even though there is an overall trend that quantal relations develop with age, several children present large deviations from expected development. For example, a younger child C02 already shows quantal relations at the age of 31 months, while an older child C06 does not show this even at the age of 40 months. This result supports the observation that for young children longitudinal studies are more appropriate than cross-sectional studies, as suggested by Smith and colleagues (1996).

There is evidence that the percentage following quantal relations is related to the articulatory maturation of children. For example, C01 and C06, whose percentages following quantal relations were relatively low at Time 1, failed to produce the diphthong /ɑɪ/. However, 6 month later, they were able to properly produce the diphthong. Similarly, C06 who had low percentage value at Time 1 also failed to produce the diphthong, and often added reduced vowels following the coda such as in "tub". This result suggests clinical applications for indication and diagnosis for speech disorder or speech immaturity for young children (see Chapter 5, applications). Compared to the current methods of articulation tests, such as the Goldman-Fristoe Test of Articulation-2, a more quantitative test would be possible by using children's subglottal resonance information. The data in Section 4.4 show that quantal relations predict better the $F1$ and $F2$ frequency changes over a 6 month period than vocal tract growth for each child. The differences between predicted data based on quantal relations and the real data for /æ/ are smaller. The significant changes in $F1$ or $F2$ during a 6 month period can be explained by quantal relations, while vocal tract growth fails to predict. This result implies that the significant changes in $F1$ and $F2$ of real data do not mainly arise from the vocal tract length changes, because the discrepancies between prediction based on vocal tract growth and the real data are significant. The results suggest that the acquisition of vowel feature contrasts is primarily dependent on articulatory learning, rather than vocal tract growth, which agrees with the conclusion by Clement and Wijnen (1994) for German children ages 2-4 years.

In conclusion, quantal articulatory-acoustic relations between formants and subglottal resonances have been shown to define several vowel and consonant distinctive features, including the vowel features [back] and [low]. Measurements of $F1$ and $F2$ for productions of the vowels /ɑ/ and /æ/ were made on a number of occasions for ten children in the age range 2;6 to 3;9 years. Measures of $SubF1$ and $SubF2$ for each child were obtained from the locations of discontinuities in the $F1$ and $F2$ trajectories in diphthongs and in vowel-consonant transitions. At the earlier ages for these children, the $F1$ and $F2$ values deviated from the expected quantal relations. During the six month period in which the measurements were made, considerable agreement with the expected values of $F1$ and $F2$ in relation to the subglottal resonances was obtained. The transition to the expected quantal relations appears to occur about at the age of 3 years.

Chapter 5. Summary and Future Directions

In this chapter, our findings of this thesis research are summarized chapter by chapter. General discussions relating to the findings and the future directions are addressed. Furthermore, several examples of applications are suggested.

5.1 Summary and Conclusion

The subglottal resonances are the natural frequencies of the airways below the glottis. Near these resonant frequencies due to subglottal coupling, the output of speech sound from the mouth shows irregularities, and it has been suggested that spectral peaks are placed to avoid these frequency regions, resulting in a quantal boundary between feature values such as front and back (Stevens, 1998; Chi and Sonderegger, 2007). In this thesis research, the hypothesis that subglottal resonances govern feature boundaries was tested, by exploring the relations among subglottal resonances, vowel formants and features for three populations: adult speakers of American English; adult speakers of; and children learning American English. The general hypothesis of this thesis research was that subglottal resonances form unstable frequency regions of quantal theory. Specifically, the first subglottal resonance forms an unstable region at $F1$ frequencies of vowels while the second subglottal resonance forms an unstable region at $F2$ frequencies; these unstable regions construct vowel feature boundaries [low] and [back]. In addition, it was hypothesized that the subglottal resonances may play a role in defining these features in other languages as well as in English since these subglottal quantal effects on speech signals are expected to be universal and independent of language.

In Chapter 2, we determined whether acoustic coupling between the first subglottal resonance ($SubF1$) and the $F1$ frequency for vowels creates a region near $SubF1$ in which the $F1$ prominence shows an irregularity. The time course of $F1$ in relation to $SubF1$ was examined for certain diphthongs and several monophthongs produced by a number of speakers of English using the database of Chi and Sonderegger (2004, 2007). For the diphthongs, a discontinuity in $F1$ or a dip in amplitude of the $F1$ prominence was observed as it passed through *SubF*1, while for the monophthongs *F*1 was usually above *SubF*1 for [+low] vowels and below *SubF*1 for [-low] vowels. A preliminary further study of data from the literature on *F*1 for vowels from various languages showed that the boundary between *F*1 values of [+low] vowels and those of [-low] vowels agrees with the average value of *SubF*1 obtained from the laboratory study with English. This finding provides evidence for a defining quantal articulatory-acoustic relation for the distinctive feature [low]. In addition, these observations provide stronger evidence for the quantal hypothesis, verifying that it is not a coincidence that the second subglottal resonance happens to lie near the frequency boundary between back vowels and front vowels.

In Chapter 3, it was tested whether *SubF*2 plays a role in the feature contrast in vowels independent of language. As preliminary research, we made recordings of speech and subglottal signals simultaneously for several adult Korean speakers. We found acoustic irregularities in *F*2 near *SubF*2 in Korean vowels. The boundary between [+back] and [-back] vowels agrees with speakers' *SubF*2. These observations support our hypothesis that *SubF*2 plays a role in defining feature contrast of [back] in other languages as well as in English since subglottal coupling effects on speech sound are theoretically the same, independent of language. In addition, we tested which of these hypotheses is correct for a low vowel in Korean, which has only one low vowel with no contrast of [+back] and [-back]: (1) *SubF*2 is avoided, and the low vowel is always [+back] or (2) *SubF*2 is avoided but the vowel is front or back depending on adjacent consonants. The measurements of *F*2 and *SubF*2 were obtained in the context /CaC/, where C is a consonant. We found that *SubF*2 was always avoided for the low vowel. If the adjacent consonants were labial or velar, *F*2 of the low vowel was below *SubF*2, whereas if the consonants were alveolar, *F*2 of the vowel was above *SubF*2. These findings suggest that *SubF*2 may have a role in assimilation of the feature [back] in Korean, as opposed to leading to a contrast for the feature in English.

In Chapter 4, we explored the role of subglottal resonances in defining the vowel features [low] and [back] for young children ages 2;6–3;9 years. Measurements of *F*1 and *F*2 for productions of the vowels /ɑ/ and /æ/ were made on a number of occasions for ten children in this age range using Imbrie's database (Imbrie, 2005). Measures of *SubF*1 and *SubF*2 for each child were obtained from the locations of discontinuities in the *F*1 and *F*2 trajectories in diphthongs and in vowel-consonant transitions. At the earlier ages for these children, the *F*1 and *F*2 values deviated from the expected quantal relations; that is the frequency boundaries between low vowels and non-low vowels or frequency boundaries between back vowels and front vowels do not always agree with subglottal resonant frequencies of child speakers. During the six month period in which the measurements were made, considerable agreement with the expected values of *F*1 and *F*2 in relation to the subglottal resonances was obtained. The transition to the expected quantal relations appeared to occur about at the age of 3 years for the most of children. This development patter was inconsistent with an account in terms of simple anatomical increase in the size of the child's vocal tract.

In conclusion, this thesis research shows that the frequency boundaries between low vowels and non-low vowels or frequency boundaries between back vowels and non-back vowels agree with the first and second subglottal resonant frequencies of speakers, respectively. It might be expected that these quantal relations between formant frequencies and subglottal resonances in vowels are valid in other languages as well. These observations provide evidence that the first two subglottal resonances play a role in defining the vowel feature boundaries for [low] and [back], respectively. For young children age 2-3 years in the stage of language development, the experimental results show that the frequency boundaries between low vowels and non-low vowels and between back vowels and front vowels usually do not agree with subglottal resonances of child speakers until 3 years. Yet, the frequency boundaries show agreeable with the subglottal resonances of speakers with increasing age. These findings suggest that the first two subglottal resonances may play a role in development of defining the feature boundaries for [low] and [back] in producing vowels for young children. These three sets of observations provide evidence that subglottal resonances play a role in defining vowel feature boundaries, as predicted by Stevens' (1972) hypothesis that contrastive phonological features in human languages have arisen from quantal discontinuities in articulatory-acoustic space. The mechanism of how formant frequencies of young children change according to their subglottal resonances requires further study.

5.2 General Discussions and Possible Future Research

Other Distinctive Features

This thesis focused on certain vowel features within the quantal framework; especially, the experimental results in Chapter 2 suggested that the first subglottal resonance might play a role in defining the vowel feature [low] and the sound categorical boundary between low vowels and non-low vowels may be defined by the location of $SubF1$. Recently, Hunt examined acoustic differences between glides and high vowels and the resulting a feature boundary between glides and high vowels (Hunt, 2009). Her experimental results show that glides are significantly different from high vowels in terms of their constriction degree in the vocal tract, resulting in the air pressure build-up behind the constrictions and changes in the glottal source, thus causing a distinction between these two sound segments. Figure 5.1 shows different classes of sound segments according to the relative degree of constrictions in the vocal tract. As seen in Figure 5.1, another categorical boundary exists besides the feature boundaries addressed above - the boundary between high vowels and non-high vowels. It is expected that a non-linear relationship between articulatory parameters and acoustic parameters defines the feature [high], and the boundary lies between high vowels (e.g., /i/ and /u/) and non-high and non-low vowels (e.g., /o/ and /e/) in standard vowels, which is marked as (b) in the figure. In quantal theory, distinctive features can change from [+feature] to [-feature] or from [-feature] to [+feature] in the following two conditions, as discussed in Chapter 1 (Halle and Stevens, 1991). For the sound segments classified as articulator-free features, the sign of the feature changes as the main sound source changes for phonation of the segment. The distinction between glides and high vowels can fall into this category in Hunt's research. On the other hand, for the feature [high], since this feature is classified as articulator-bound features, the main sound source is expected to be the same between [+high] and [-high] segments. It is also supposed that the configurations of the vocal tract are significantly different or a certain side branch is coupled to the vocal tract filter, resulting in coupling effects between resonators, as similar as in the feature [low]. However, the source of distinction between high vowels and non-high vowels are unknown, and this can be a topic for future research.

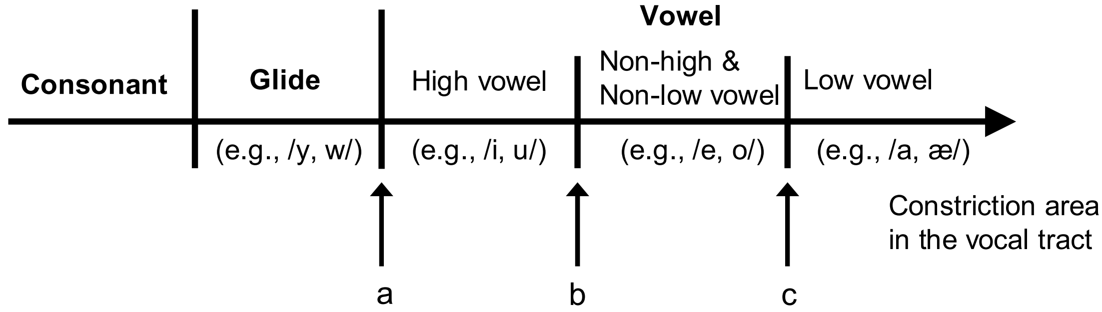


Figure 5.1 Features according to the degree of constrictions in the vocal tract (the idea of this figure is adapted from Hunt 2009). **a**: The boundary between glides and high vowels from Hunt 2009, **b**: the boundary between high vowels and non-high vowels, which is defined as the feature [high], **c**: the boundary between low vowels and non-low vowels, which is defined as the feature [low] from this thesis research.

There are also many different vowel features such as [round] and [tense], however, this thesis research was limited to the feature [low] and [back]. Exploring other vowel features within the framework of distinctive feature theory and quantal theory can be the possible future directions.

Cross-linguistic Study

The data in the first and second experiments of Chapter 3 for Korean vowels, which were obtained as the first step toward cross-linguistic study, suggest that the subglottal resonances may play a role in Korean. The results provide preliminary evidence that the second subglottal resonance has the same effects on the speech signals, independent of languages. Though this thesis study was limited to the vowel features [low] and [back], there are many other features which are common for many different languages in the world. For example, Korean stop consonants are classified as consonant features [labial], [alveolar] and [velar] according to the place of constriction similar to English stop consonants. Therefore, cross-linguistic study of these universal features within the framework of quantal theory can give insight into why the universal sound categories exist for various languages in the world and characteristics of the universally common sounds.

In addition, the data in Chapter 3 show that *F*2 shift depending on the place of adjacent consonants is often observed, and the *F*2 shift direction is related to *SubF*2: If the adjacent consonants are alveolar, $F2 - SubF2 > 0$, whereas if the adjacent consonants are velar or labial, $F2 - SubF2 < 0$. These findings suggest that *SubF*2 may have a role in assimilation of the feature [back] in Korean, as opposed to leading to a contrast for the feature in English. As future research, for the vowels where the feature [back] has not been specified in a certain language (e.g., /a/ in Korean, Mandarin Chinese, Japanese, or /ə/ in English), similar study of *F*2 shift of the middle vowels is suggested by the relation to *SubF*2 in various consonant environments. These findings can possibly be applied into automatic speech recognitions using cross-linguistic database.

Subglottal Acoustics

The results of subglottal resonances from several adult Korean speakers in Chapter 3 suggest that the second subglottal resonances may be correlated with the speaker's height. In addition, the subglottal resonances obtained from several child speakers age of 2-3 years in Chapter 4 show that the subglottal resonances approximately linearly decrease with age. This result implies that the size of subglottal system may increase with age because the subglottal resonances are presumed to be related to the size of the subglottal system of a speaker. It is expected that the subglottal resonances decrease as the size of the subglottal system increases. Though the linear correlation between the second subglottal resonance and the height was suggested in this study, the subglottal system is a very complicated system which consists of trachea, bronchi and lungs, and every air way is divided into two small airways. Currently, little is known about acoustic characteristics of the subglottal system depending on the physical characteristics of the subglottal system (e.g., the characteristics of wall tissues in the lower airways). Detailed acoustic study of the subglottal system can possibly give answers for the following questions: how can physical changes due to lung surgery have an effect on the subglottal acoustics? What acoustic changes might result in speech? How broad is the width of avoided frequency band due to subglottal coupling depending on speakers?

It was assumed that in this thesis study the subglottal resonance value is almost constant during the phonation of different vowels for a given speaker, and the effects from different configurations in the vocal tract on the subglottal resonances are minimal. Preliminary experimental results, which are not included in this thesis, show that for the standard vowels, the subglottal resonant value relatively does not change in the various vowel environments for a given speaker, compared to the formant frequencies. The value of COV (standard deviation/mean) was in the range of 0.05 for subglottal resonances across vowels, whereas COV was in the range of 0.5 for formant frequencies across vowels for a given speaker. Future expanded experiments and theoretical modeling could examine the range of subglottal resonances for various contextual environments and could explore subglottal resonance changes depending on a particular vowel.

Quantal Development and Children's Speech

In Chapter 4, the development of the features [low] and [back] in production of vowels was explored for young children of age 2-3 years. This work was a first step in studies of the development of production in the perspective of quantal theory. A speaker needs to learn about which articulator should be used and how the articulatory configurations should be, for producing a speech sound segment; phonological information should be mapped to articulatory information in this process. The findings of this thesis research can give insight into understanding of correlations between phonologically defined features and physical properties of the speech production system, the effects of physical properties of the speech production on articulatory controls and categorical perceptions of speech sound segments in speech development. Therefore, similar studies on quantal development of other features for young children are required in future. For example, for the feature [anterior], the development of a fricative consonant /s/ ([+anterior]) and /š/ ([-anterior]) can be explored in the production of young children. As discussed in Chapter 1, the boundary between /s/ and /š/ can be theoretically explained by quantal theory. It is expected that the unstable frequency region might be determined by the length of a front cavity and the length of a back cavity, which depend on the tongue size, oral cavity size, and the position of tongue constrictions. Therefore, the changes in these articulatory-acoustic parameters with age can possibly be interesting future research topics.

The data shown in Chapter 4 imply that the acquisition of the quantal divisions between vowel categories is gradually developed in young children. This observation raises the following questions: How do children know to align the boundaries of vowels with quantally unstable regions? What aspects of the acquisition system are innate or which involve learning? Would all children arrive at the same boundaries, or whether particular children might happen to choose different dividing lines? Would some children fail to use *SubF*2 because they do not have a strong subglottal effect in their own speech? For some children with phonological delay to be late in realizing quantal boundary (e.g., C06, see Appendix C, Fig C-3), why they fail to recognize the proper boundary? To answer these questions, the mechanism of how formant frequencies of young children change according to their subglottal resonances requires further study. For example, our preliminary analysis for the subject C06 with phonological delay in realizing quantal boundaries of vowels shows that for the child, frequency discontinuities were rarely observed in the *F*2 time course due to *SubF*2. This observation might imply that this child has a week subglottal coupling effect on speech, causing the child fail to use *SubF*2 as the boundary between vowel categories. This possibility should be tested for future research.

This thesis research in Chapter 4 focused on the vowel development using Imbrie's database (Imbrie, 2005), and the data show that adult-like distinctions between front vowels and back vowels are achieved by most children by the age of 3 years. In Imbrie's research, which focused on the development of stop consonant production, the findings indicate that the development is still in progress toward adult-like production both in articulation and phonation. The results of both of two studies show that for each speaker an adult-like articulatory configuration develops with age. In addition, the data in Chapter 4 show that even though there is an overall trend that quantal relations develop with age, there is considerable variation in the developmental speed from one individual to another. Several children present large deviations from expected development. A large number of studies on speech development in children use cross-sectional data. Our results in Section 4.3.2 show that the averaged values of formant frequencies do not change significantly in the age of 2-3 years across speakers though there are significant changes in frequencies during a six month period for each speaker. This result implies that the averaged parameters across subjects at the same age might obscure important information of individual differences. Such a conclusion can drive us to misinterpretations of data, and suggests that for young children longitudinal studies are more appropriate than cross-sectional studies. So far, cross-sectional data are more available than longitudinal data, and more cross-sectional studies have been performed than longitudinal studies. Therefore, more longitudinal databases need to be built, and more longitudinal studies should be carried out, especially for children younger than two years old. Such studies are relatively rare due to difficulties of data collection. Because the data of this thesis research cover the age range of 2-3 years, expanded work using speech data from children younger than two-year old would be useful to understand how their productions develop from babbling to categorical speech sounds of vowels. Also, possible future directions could include accelerometer signal recordings for children younger than 2 years, testing correlations between this research result and listener's perception, examining other vowel pairs such as /e/ and /o/, and studying subjects with speech disorder.

Automatic Detection of Subglottal Resonances

In this thesis research, the subglottal resonances were obtained by using two methods. In Chapter 2 and Chapter 3, the subglottal resonances were measured from accelerometer signals. In Chapter 4, the subglottal resonances were manually obtained from the locations of frequency discontinuities in speech signals due to subglottal effects. In addition, it was shown that the estimated subglottal resonance values from the second method agree with the subglottal resonance values from accelerometer signals in Chapter 2 and Chapter 3. However, few databases of accelerometer signals are currently available, and, therefore, reliable automatic detection methods of subglottal resonances from speech databases are required. This would allow us to use a large number of currently existing speech databases. For example, our preliminary experiment results, which are not included in this thesis, show that the bandwidths of formants (B1, B2) seem to increase near subglottal resonances because the energy in the vocal tract filter would be absorbed by the subglottal system at the subglottal resonances. More reliable methods of detecting subglottal resonances from speech signals by using the bandwidths of formants can be important directions for future research (Wang et al., 2008a,b).

5. 3 Applications

The findings in this study can be applied to various areas as addressed in Chapter 1, such as modeling of sound acquisition in young children, clinical areas, and automatic speech recognition. Examples of several possible applications are suggested here.

Clinical Areas

The finding of these studies can be applied to clinical areas such as speech testing for maturity and/or disorders, and speech therapy. In Chapter 4, the observations that formant frequencies of children of early ages deviated from the expected quantal relations suggest that formant frequencies of immature speech or disordered speech will deviate from the expected quantal relations. For example, the formant frequencies of speakers with hearing loss deviate largely from those of normal speakers. The $F1$ and $F2$ frequencies are lower because of lack in hearing feedback (Angelocci et al., 1964). The difference between $F1$ and $SubF1$ or between $F2$ and $SubF2$ can quantitatively give information about the degree of deviation from normal speech. In addition, the information can be used for speech therapy as a guideline: if $F1 < SubF1$ for a speaker is observed for low vowels, the tongue body of the speaker needs to be lowered, and if $F2 < SubF2$ is observed for front vowels, the tongue body needs to be more fronted.

Furthermore, the results may potentially be used in developing non-invasive devices for clinical examinations concerning the conditions of subglottal systems. Since the subglottal resonance values are assumed to vary with changes in the shape of the trachea, for example, changes due to lung surgery, subglottal resonances of the patients who have had surgery might deviate from those of people in normal condition (Lulich, 2006).

Automatic Speech Recognition

Subglottal resonances can be used in automatic speech recognition for speaker normalization. For example, due to large differences in acoustic characteristics between adult speakers and children speakers, the performance of the automatic speech recognizer using adult training data can be severely degraded. Recent study has reported that the speaker normalization technique based on the subglottal resonances for children speakers shows better performance than conventional techniques using the vocal tract length (Wang et al, 2008a,b).

Appendix A

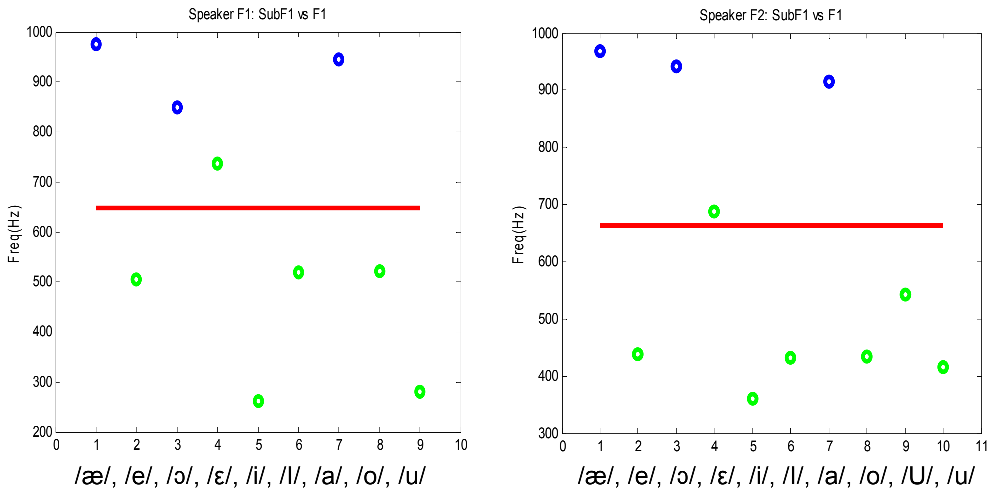


Fig A-1. *F*1 frequencies of English monophthong vowels for female speakers F1 (left) and F2 (right). Each point represents the averaged *F*1 value from four tokens. Horizontal lines represent *SubF*1 values from Table 2.2. Blue dots: [+low] vowels; Green dots: [-low] vowels.

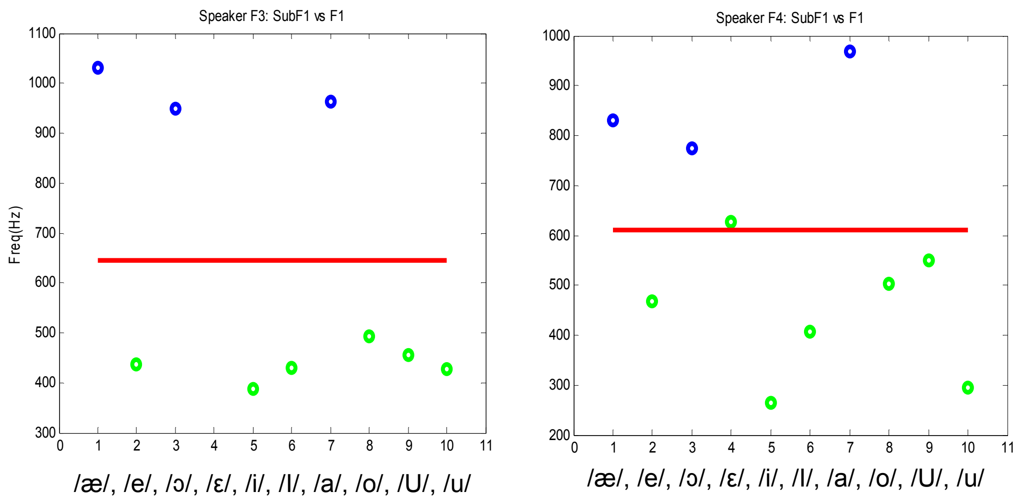


Fig A-2. *F*1 frequencies of English monophthong vowels for female speakers F3 (left) and F4 (right). Each point represents the averaged *F*1 value from four tokens. Horizontal lines represent *SubF*1 values from Table 2.2. Blue dots: [+low] vowels; Green dots: [-low] vowels.

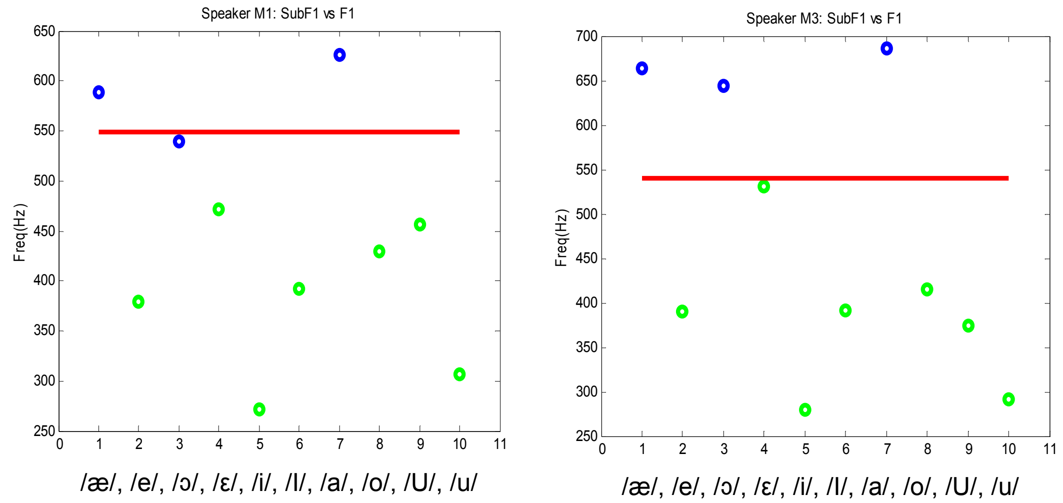


Fig A-3. *F*1 frequencies of English monophthong vowels for male speakers M1 (left) and M3 (right). Each point represents the averaged *F*1 value from four tokens. Horizontal lines represent *SubF*1 values from Table 2.2. Blue dots: [+low] vowels; Green dots: [-low] vowels.

Appendix B

Table B-1. Languages, number of subjects and averaged values of the first two formant frequencies of standard vowels for male speakers. The English vowel /ɑ/ was classified as /a/ in this table.

language	Subject number		vowel					
			a	æ	e	i	o	u
Spanish (Cervera et al., 2001)	10	**F1**	718		502	331	533	376
		F2	1479		1872	2241	1156	773
English (Yang, 1996)	10	**F1**	694	687	469	286	498	333
		F2	1127	1743	2082	2317	1127	1393
Hindi (Khan et al., 1994)	2	**F1**	665		535	390	485	505
		F2	1145		2075	2510	1120	1265
Brazilian Portuguese (Escudero et al., 2009)	10	**F1**	683		357	285	372	310
		F2	1329		2028	2198	804	761
European Portuguese (Escudero et al., 2009)	10	**F1**	661		355	284	363	303
		F2	1365		1987	2161	843	814
Bengali (Ray and Ghoshal, 1996)	3	**F1**		681	374	304	438	327
		F2		1663	1935	2095	1015	935
Japanese (Tokura et al., 1992)	1	**F1**	750		469	281	468	312
		F2	1187		2031	2281	781	1219
German (Heid et al., 1995)	16	**F1**	690		383	324	416	328
		F2	1339		2076	2071	927	946
Dutch (Pols et al., 1973)	50	**F1**	795		407	294	487	339
		F2	1301		2017	2208	911	810
Korean (Yang, 1996)	10	**F1**	738		490	341	453	369
		F2	1372		1968	2219	945	981
French (Gendrrot and Adda-Decker, 2005)	10	**F1**	650		350	300	400	350
		F2	1300		1950	2050	900	850

Table B-2. Languages, number of subjects and averaged values of the first two formant frequencies of standard vowels for female speakers. The English vowel /ɑ/ was classified as /a/ in this table.

language	Subject number		vowel					
			a	æ	e	i	o	u
English (Yang, 1996)	10	**F1**	857	825	521	390	528	417
		F2	1255	2059	2536	2826	1206	1511
Hindi (Khan et al., 1994)	1	**F1**	800		620	540	590	600
		F2	1900		2620	3010	1870	1820
Brazilian Portuguese (Escudero et al., 2009)	10	**F1**	910		425	307	442	337
		F2	1627		2468	2676	893	812
European Portuguese (Escudero et al., 2009)	10	**F1**	781		402	313	422	335
		F2	1662		2508	2700	921	862
Korean (Yang, 1996)	10	**F1**	986		650	344	499	422
		F2	1794		2377	2814	1029	1021
French (Gendrrot and Adda-Decker, 2005)	10	**F1**	750		350	450	450	350
		F2	1550		2050	2400	950	850

Appendix C

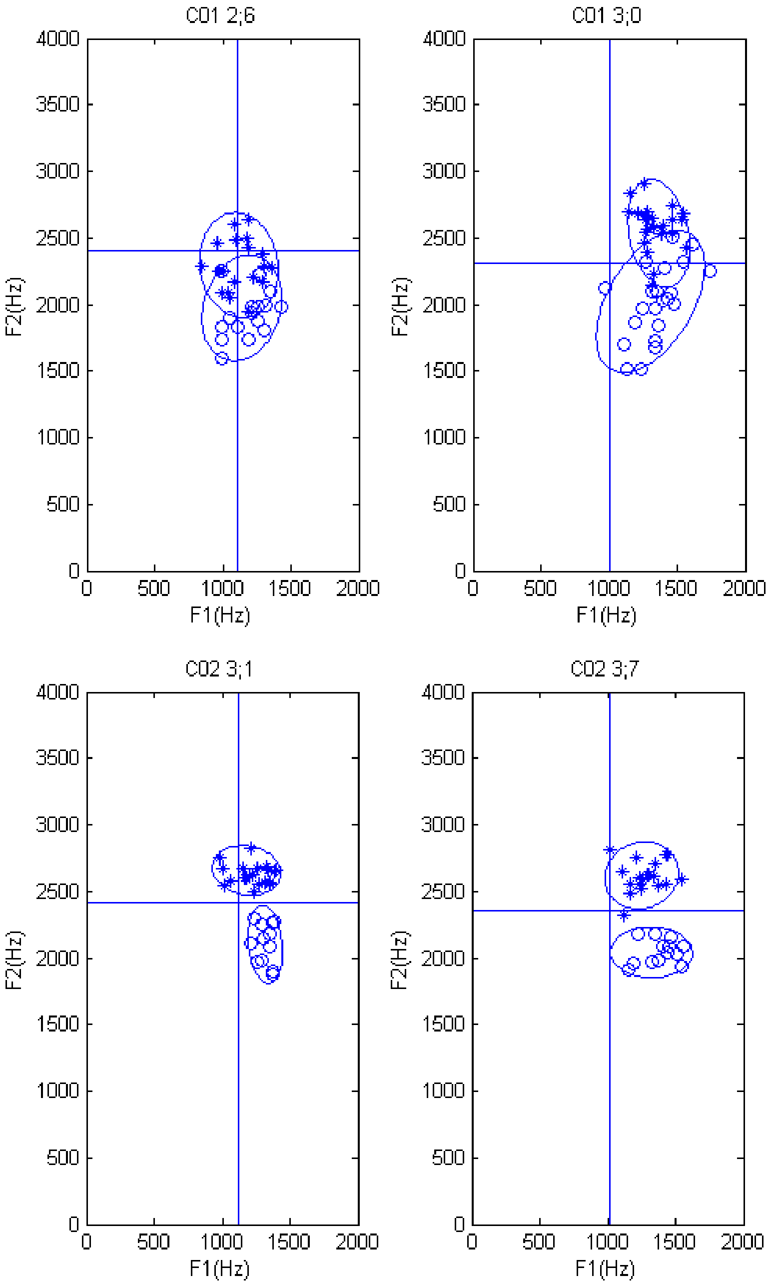


Fig C-1. *F*1-*F*2 plot for low vowels /æ/ and /ɑ/ for the subjects C01 and C02.

Asterisk: vowel /æ/, Circle: vowel /ɑ/.

Vertical line: *SubF*1 for each speaker, Horizontal line: *SubF*2 for each speaker.

Upper: C01, Lower: C02.

Left: Time 1, Right: Time 2 (six month later).

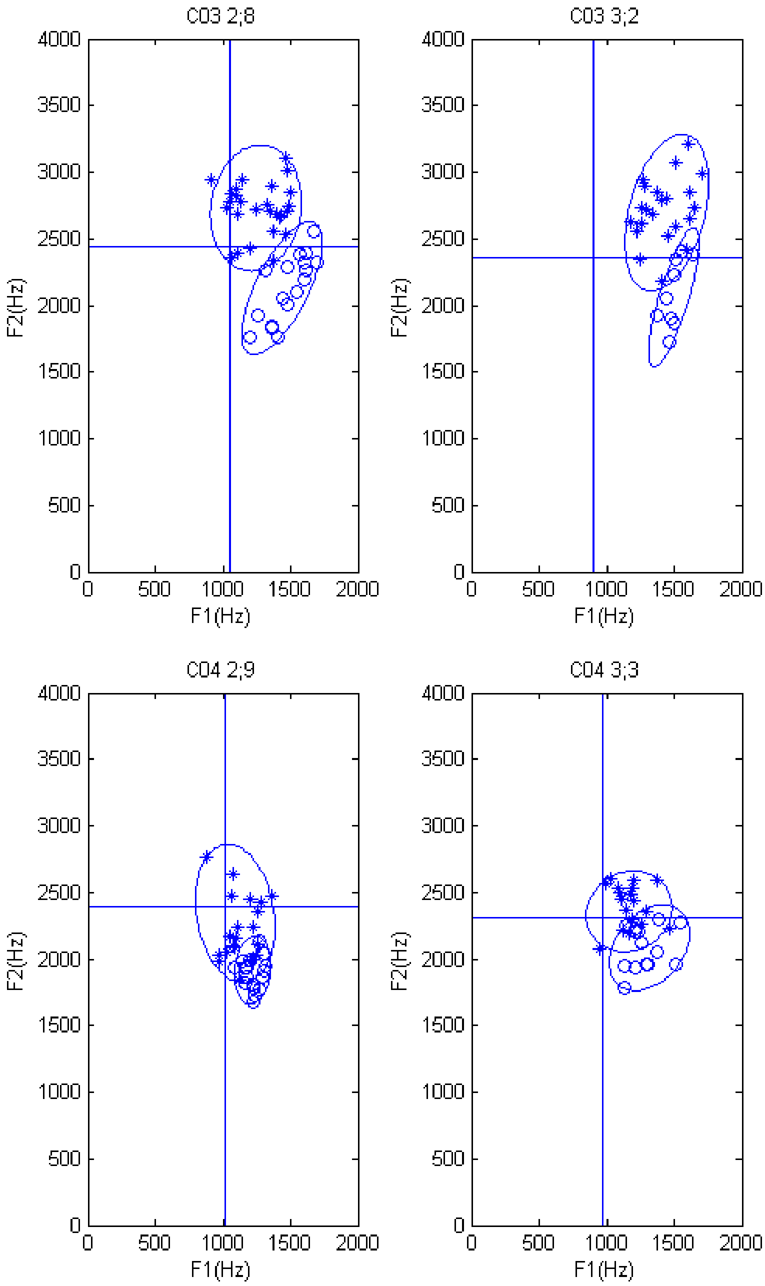


Fig C-2. *F*1-*F*2 plot for low vowels /æ/ and /ɑ/ for the subjects C03 and C04.

Asterisk: vowel /æ/, Circle: vowel /ɑ/.

Vertical line: *SubF*1 for each speaker, Horizontal line: *SubF*2 for each speaker.

Upper : C03, Lower: C04.

Left: Time 1, Right: Time 2 (six month later).

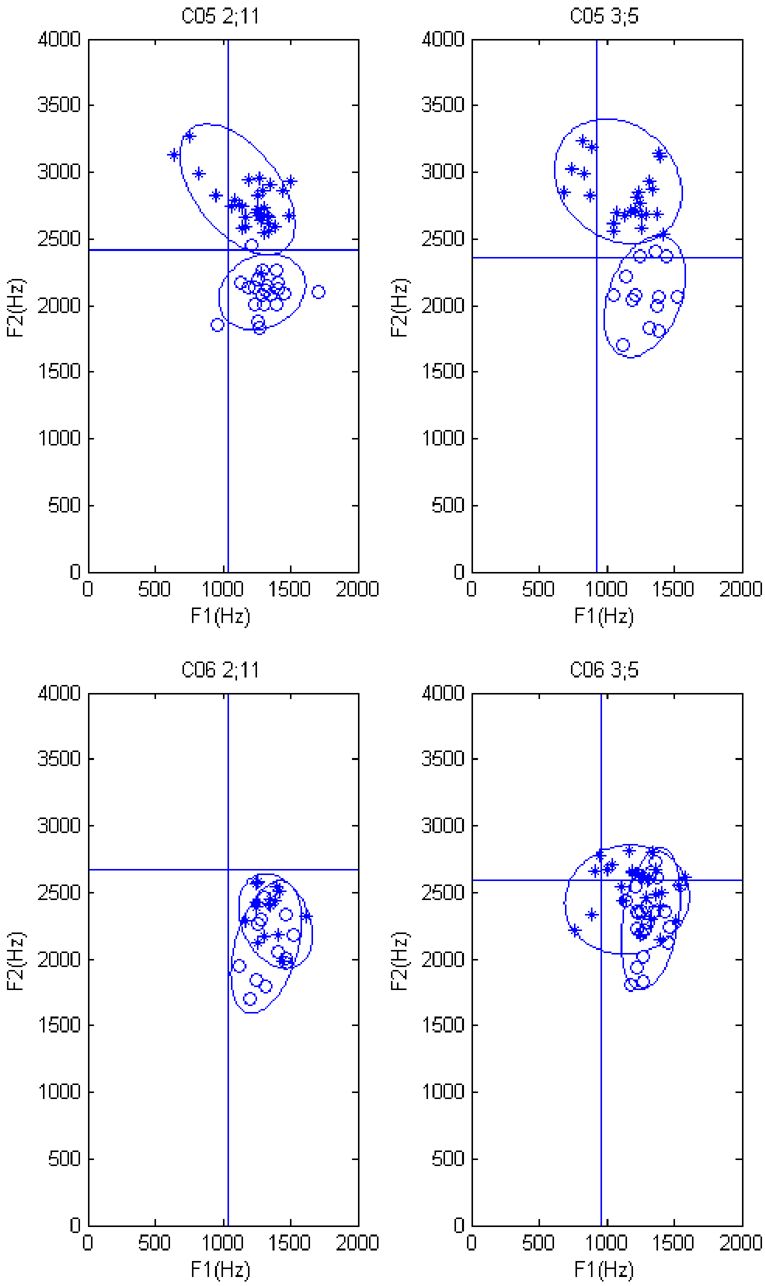


Fig C-3. *F*1-*F*2 plot for low vowels /æ/ and /ɑ/ for the subjects C05 and C06.
Asterisk: vowel /æ/, Circle: vowel /ɑ/.
Vertical line: *SubF*1 for each speaker, Horizontal line: *SubF*2 for each speaker.
Upper : C05, Lower: C06.
Left: Time 1, Right: Time 2 (six month later).

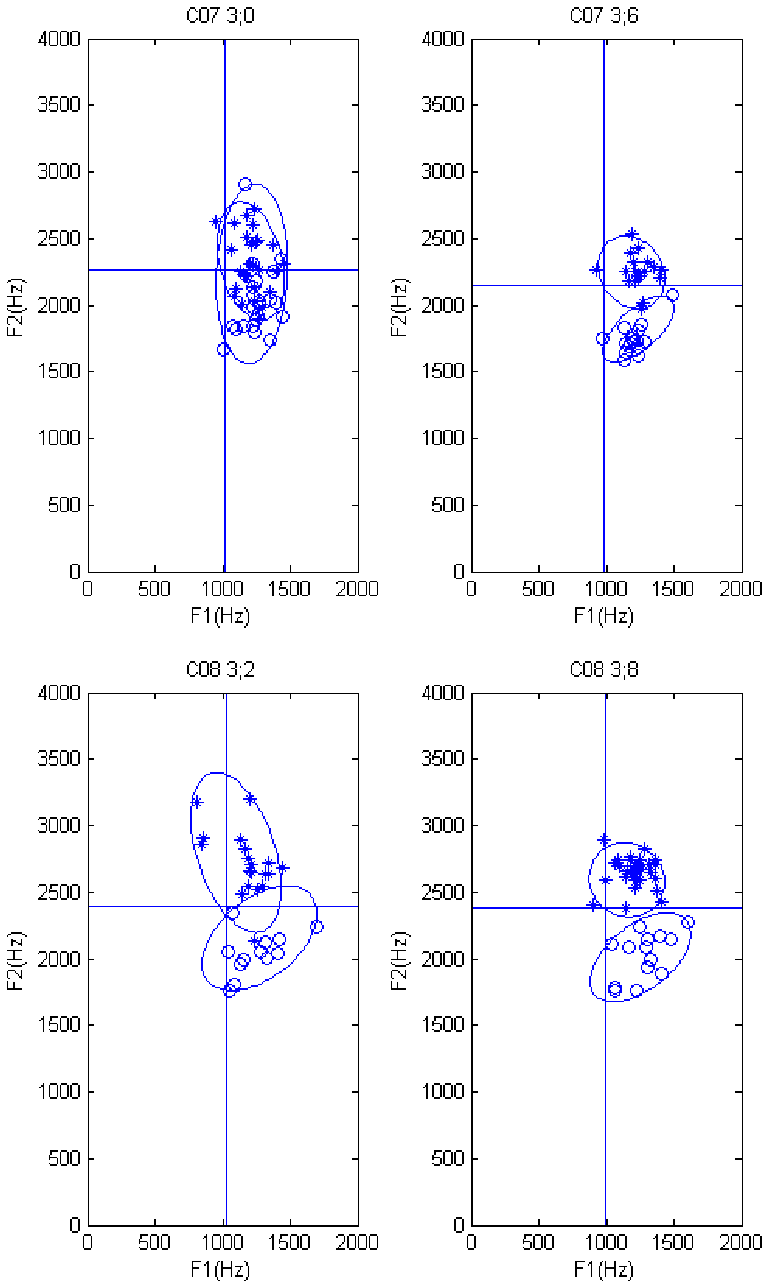


Fig C-4. *F*1-*F*2 plot for low vowels /æ/ and /ɑ/ for the subjects C07 and C08.

Asterisk: vowel /æ/, Circle: vowel /ɑ/.

Vertical line: *SubF*1 for each speaker, Horizontal line: *SubF*2 for each speaker.

Upper : C07, Lower: C08.

Left: Time 1, Right: Time 2 (six month later).

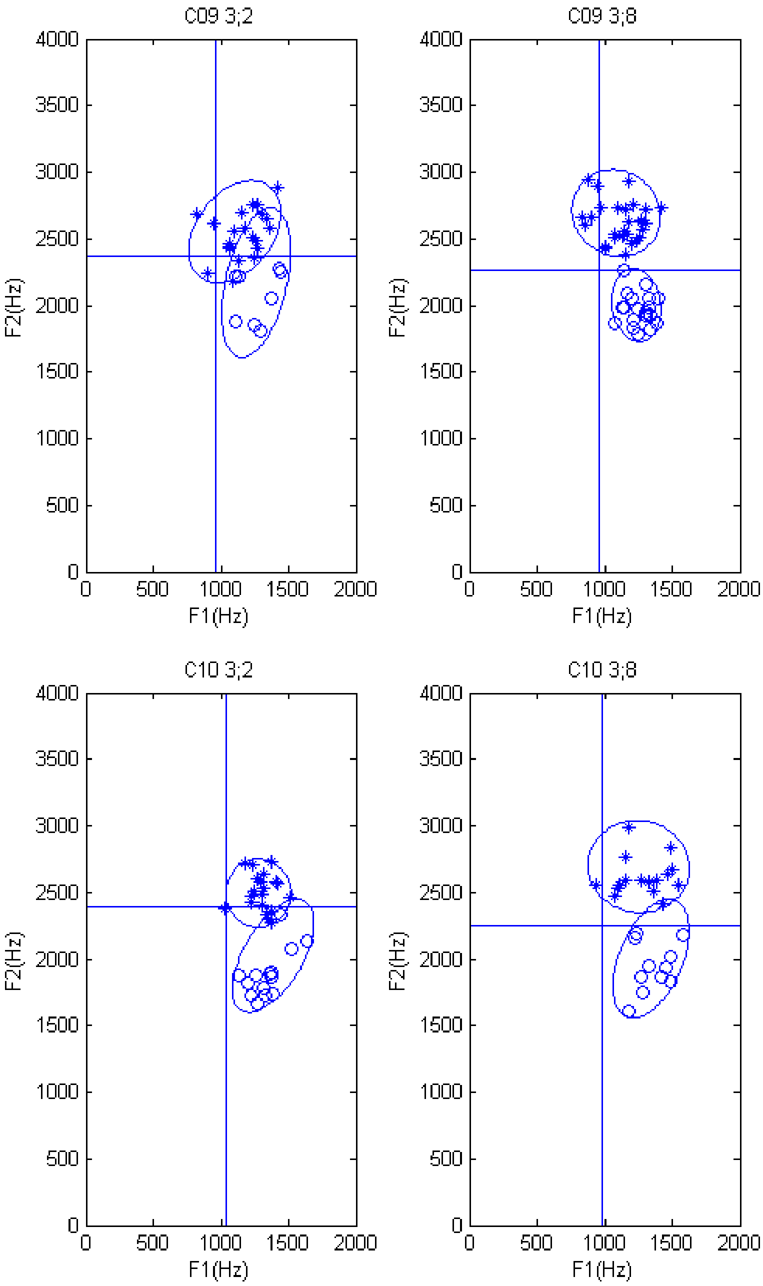


Fig C-5. *F*1-*F*2 plot for low vowels /æ/ and /ɑ/ for the subjects C09 and C10.

Asterisk: vowel /æ/, Circle: vowel /ɑ/.
Vertical line: *SubF*1 for each speaker, Horizontal line: *SubF*2 for each speaker.
Upper : C09, Lower: C10.
Left: Time 1, Right: Time 2 (six month later).

References

Ahn, S. (1998) An introduction to Korean phonology. Hanshin Publishing Co., Seoul, Korea.

Ahn, S and Iverson, G. K. (2005) Structured imbalances in the emergence of the Korean vowel system. Historical Linguistics, 175-193.

Angelocci, A. A., Kopp, G. A., and A. Holbrook (1964) The Vowel Formants of Deaf and Normal-Hearing Eleven- to Fourteen-Year-Old Boys. Journal of Speech and Hearing Disorders. 29: 156-170.

Bond, Z. S., Petrosino, L., and Dean, C. R. (1982). "The emergence of vowels: 17 to 26 months," J. Phonetics 10, 417–422.

Cervera T. Miralles, J., and Gonzalez-Alvarez, J. (2001) Acoustical Analysis of Spanish Vowels Produced by Laryngectomized Subjects. (2001) Journal of Speech, Language, and Hearing Research, 44, 988-996

Cheyne, H. (2002) "Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Chi, X., and Sonderegger, M. (2004) Subglottal coupling and vowel space. Journal of the Acoustical Society of America, 115:2540–2540.

Chi, X. (2005) The quantal effect of sub-glottal resonance on vowel formant. RQE report, MIT.

Chi, X., and Sonderegger, M. (2007) Subglottal coupling and its influence on vowel formants. Journal of the Acoustical Society of America, 122:1735-1745.

Chomsky, N., and Halle, M. (1968) The sound pattern of English. New York; Harper & Row.

Clement, C. J., and Wijnen, F. (1994) Acquisition of Vowel Contrasts in Dutch. Journal of Speech and Hearing Research, 37: 83-89.

Cranen, B., and Boves, L. (1987) On subglottal formant analysis. Journal of the Acoustical Society of America, 84, 734–746.

Eguchi, S., and Hirsh, I. J. (1969). Development of speech sounds in children. Acta Otolaryngologica, 257, 5–51.

Escudero, P., Boersma, P., Rauber, A. S., and Bion R. (2009) A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. Journal of the Acoustical Society of America.

Fant, G., Ishizaka, K., Lindqvist, J., and Sundberg, J. (1972) "Subglottal formants," KTH Speech Transmission Laboratory Quarterly Progress and Status Report 1, 1–12.

Fitch, W. T., and Giedd, J. (1999) Morphology and development of the human vocal tract: A study using magnetic resonance imaging. Journal of the Acoustical Society of America, 106:1511-1522.

Gendrot1, C. and Adda-Decker, M. (2005*)* Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German. Interspeech, Lisbon, Portugal

Goodell, E. W., and Studdert-Kennedy, M. (1993) Acoustic evidence for the development of gestural coordination in the speech of 2-year-olds: A longitudinal study. Journal of Speech Hearing Research 36:707-727.

Habib, R. H., Chalker, R. B., Suki, B., and Jackson, A. C. (1994). Airway geometry and wall mechanical properties estimated from subglottal input impedance in humans. Journal of Applied Physiology, 77(1):441–451.

Halle, M. and Stevens, K. N. (1991) Knowledge of language and the sounds of speech. In J. Sundberg, L. Nord, and R. Carlson (Eds.), Music, Language, Speech and Brain, Basingstoke, Hampshire: Macmillan Press, 1-19.

Hanson, H. M., and Stevens, K. N. (1995). Sub-glottal resonances in female speakers and their effect on vowel spectra. In Proceedings of ICPhS-95 (Vol. 3, pp. 182–185). Stockholm.

Hanson, H. M. (1997) Glottal characteristics of female speakers: Acoustic correlates. Journal of the Acoustical Society of America, 101 (1) , 466-481.

Harper V. P., Kraman, S. S., Pasterkamp, H., and Wodicka, G. R. (2001) An acoustic model of the respiratory tract. IEEE Transactions on Biomedical Engineering, 48(5):543–550.

Heid, S., Wesenick, M.-B., and Draxler., C., (1995) Phonetic analysis of vowel segment in the PhonDat database of spoken Germen. Proceedings of the XIIth International conference of Phonetic Sciences, Stokholm. 4: 416-419

Honda, K., Takano, S., and Takemoto H. (2009) Effects of side cavities and tongue stabilization: Possible extensions of the quantal theory Journal of Phonetics (2009)

Hunt, E. H. (2009) Acoustic Characterization of the Glides /j/ and /w/ in American English. Ph. D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Imbrie, A. (2005) Acoustical study of the development of stop consonants in children. Ph. D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Ishizaka, K., Matsudaira, M., and Kaneki, T. (1976**)** Input acousticimpedance measurement of the subglottal system. Journal of the Acoustical Society of America, 60, 190–197.

Ishizuka, K., Muitani, R., Kato, H., and Amano, S. (2007) Longitudinal developmental changes in spectral peaks of vowels produced by Japanese infants. Journal of the Acoustical Society of America, 121:2272-2282.

Jakobson, R. C., Fant, G. M., and Halle, M. (1952) Preliminaries to speech analysis: The distinctive features and their correlates. Acoustics Laboratory Technical Report 13, MIT, Cambridge MA; reprinted by MIT Press, Cambridge MA, 1967.

Jung, Y., Stevens, K. N. (2007) Acoustic articulatory evidence for quantal vowel categories: The feature [low]. Journal of the Acoustical Society of America, 122:3029.

Khan I., Gupta, S. K., and Rizvi., S. H. (1994) Formant frequencies of Hindi vowels in /hVd/ and C1VC2. Journal of the Acoustical Society of America, 96 (4), 2580-2582.

Kent, R. D. (1976) Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. Journal of Speech Hearing Research 19:421-447.

Kent, R. D., and Murray, A. D. (1982). Acoustic features of infant vocalic utterances. Journal of the Acoustical Society of America, 72, 353–365.

Lee, S., Potamianos, A., and Narayanan, S. (1999). Acoustics of children’s speech: Developmental changes of temporal and spectral parameters. Journal of the Acoustical Society of America, 105, 1455–1468.

Lulich, S. M. (2006**)** The role of lower airway resonances in defining feature contrasts, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Lulich, S. M., Bachrach, A., and Malyska, N. (2007) A role for the second subglottal resonance in lexical access. Journal of the Acoustical Society of America. 122:2320-2327.

Lulich, S. M. (2009) Subglottal resonances and distinctive features. J Phonetics.

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Madsack A., S. M. Lulich, W. Wokurek and G. Dogil. (2008) Subglottal resonances and vowel formant variability: A case study of High German monophthongs and Swabian diphthongs. In Proceedings of LabPhon11, 91-92.

Mou, X. (2006) Nasal codas in Standard Chinese - a study in the framework of the distinctive feature theory, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Peterson, G. E., and Barney, H. L. (1952). Control methods used in a study of the vowels. Journal of the Acoustical Society of America, 24, 585–594.

Pols, L. C. W, Tromp, H. R. C., and Plomp R. (1973) Frequency analysis of Dutch vowels from 50 male speakers. Journal of the Acoustical Society of America, 53(4), 1093-1102

Ray K. S., and Ghoshal, J. (1996). Approximate reasoning approach to pattern recognition. Fuzzy Sets and Systems, 77, 125-150

Smith, B. L., Kenney, M. K., and Hussain, S. (1996) A longitudinal investigation of duration and temporal variability in children's speech production. Journal of the Acoustical Society of America, 99: 2344-49.

Sonderegger, M. (2004) Subglottal coupling and vowel space**.** B.S. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Stevens, K.N. (1972) The Quantal nature of speech; Evidence from articulatory-acoustic data. In P.B. Denes and E.E. David Jr. (Eds.) Human Communication: A Unified View, New York: McGraw-Hill, pp 51-66.

Stevens, K. N. (1989) On the quantal nature of speech. Journal of Phonetics, 17, 3-46.

Stevens, K. N. (1998) *Acoustic Phonetics*. MIT Press, Cambridge, Massachusetts.

Stevens, K. N. (2003). Acoustic and perceptual evidence for universal phonological features. In Proceedings of the 15th international congress of phonetic sciences, Barcelona. 33-38

Vorperian, H. K., Kent, R. D., Lindstrom, M. J. Kalina, C. M., Gentry, L. R., and Yandell, B. S. (2005) Development of vocal tract length during early childhood: A magnetic resonance imaging study. Journal of the Acoustical Society of America, 117: 338-350.

Vorperian, H. K., and Kent, R. D. (2007) Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. Journal of Speech Language Hearing Research 50:1510-1545.

Wang, S., Alwan, A., and Lulich, S. M. (2008) Speaker normalization based on subglottal resonances. In Proc. ICASSP, 4277-4280.

Wang, S., Lulich, S. M., and A. Alwan (2008) A reliable technique for detecting the second subglottal resonance and its use in cross-language speaker adaptation. In Proc. Interspeech.

Yang, B. Y. (1996) A comparative study of American English and Korean vowels produced by male and female speakers. Journal of Phonetics. 24(2): 245-261.