PRODUCTION AND PERCEPTION OF THE VOICELESS SIBILANT
FRICATIVES IN TYPICALLY DEVELOPING CHILDREN WITH APPLICATIONS
FOR CHILDREN WITH CLEFT PALATE

Kerry Callahan Mandulak

A dissertation submitted to the faculty of the University of North Carolina at
Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of
Philosophy in Speech and Hearing Sciences in the Department of Allied Health
Sciences.

Chapel Hill
2009

Approved by:

Katarina Haley
David Zajac
Linda Watson
Elizabeth Crais
Ralph Ohde

ABSTRACT

Kerry Callahan Mandulak: Production and Perception of the Voiceless Sibilant Fricatives in Typically Developing Children with Applications for Children with Cleft Palate
(Under the direction of Katarina L. Haley)

The purpose of this study was to advance the current knowledge base regarding production and perception of the voiceless sibilant fricatives [s] and [ʃ] in two groups of ten typically developing children each, age 7 and 11. Developmental differences in production and perception were investigated, as well as the relationship between production and perception. A group of five children with repaired cleft lip and palate between 7 and 11 years of age was included in the study to determine if differences exist in perception or production in children with obligatory limitations in early development of speech production and perception skills compared to typically developing children.

The findings from the analyses of fricative production indicated that almost all typically developing children (95%) showed non-overlapping productive distinction between the voiceless sibilant fricatives, with varying degrees of token-to-token variability and variability in dynamic patterns of production. Developmental differences in production between the two age groups were found for fricative duration and coefficient of variation for [s] at midpoint. Differences in fricative perception were found between the TD-7 and TD-11 groups, with the older children displaying qualitatively steeper slopes on identification functions, and greater accuracy and less variability on tests of

iii

fricative discrimination compared to the younger children. No linear relationship was found between the participant's measures of fricative production and perception in the two age groups.

Children with repaired cleft lip and palate showed greater proportion of overlapping fricative production, but like the typically developing children, showed individual speaker variability in dynamic spectral patterns during production. In general, the participants in the CLP group showed monotonic crossovers in identification of the [s] – [ʃ] continuum despite most speakers showing no productive distinction. Fricative discrimination in the CLP group was similar to the performance of the TD-7 and TD-11 group, with older children in the CLP group demonstrating greater accuracy and less variability compared to the younger children in this group. Similar to the typically developing children, there did not seem to be a relationship between production and perception of the voiceless sibilant fricatives in the participants with repaired cleft lip and palate.

ACKNOWLEDGEMENTS

my dissertation project. Along the way, she has taught me many things: what it means to be a woman scientist; how to prioritize family; how to stand up for myself, in a respectful and humble way; how to problem solve the issues that plagued me throughout this journey; and how to be a person that can always be counted on to get the job done, meet a deadline, and produce quality work. She took on the responsibility of directing my dissertation without a second thought. She has always supported my ideas, even though our research, while it does overlap in some respects, focuses on widely different patient populations. She has been more than a role model to me – she has also been a wonderful friend.

David Zajac has also been a significant member of this Ph.D. journey. He and I have known each other for at least 8 – 9 years now; first, through our involvement in the American Cleft Palate – Craniofacial Association, and now, through working together in the doctoral program. David was generous enough to come up with a great idea for a "one-semester" research project that could be written up for the Cleft Palate – Craniofacial Journal.   This one-semester project turned into a 2-year project, but not before it also became my favorite comprehensive examination question and my very first published manuscript. David granted me first authorship, and this generosity will always be remembered as my very first publication. In addition, along the way, David was always good for a rousing debate about phonological awareness and reading skills, educating me about bat behavior, lively conversations about nasometry and clinical care of children with cleft lip and palate, and promising me very challenging questions for my oral comprehensive exam defense (What is the Treaty of Utrecht?) and my dissertation defense (Why is gravity the weakest

force?). Luckily, he never did test me on those questions, although I will admit, I was prepared to discuss the Treaty. David has supported me throughout my doctoral program, and I was always very encouraged by his belief in my abilities to finish this dissertation project.

Linda Watson was gracious to volunteer her time and effort to serve on my dissertation committee, and I was so thankful for her commitment to me and my doctoral career. Throughout my program, she has been a non-stop encourager and supporter of my efforts. I was fortunate to serve as a grant coordinator on one of her projects, and that allowed me the chance to work with her side by side and learn from her. Although my research area is not necessarily in her own area of interest, she was always willing to support me by serving also on my comprehensive examination committee. I will always be thankful for her leadership, encouragement, support, and friendship throughout this journey.

Elizabeth Crais was also willing to provide her time and efforts to serve on my dissertation committee. Betsy has truly been a significant source of encouragement and optimism for me from the very start of my consideration of entering the doctoral program. She was instrumental in convincing me that I could be successful in this program. Betsy has empathized with my struggles with Attention Deficit Disorder, supporting my progress or avoiding judgement when I admitted I had made no progress. Her neverending positive attitude was always a confidence builder for me.

I am indebted to the contributions of Ralph Ohde, professor in Hearing and Speech Sciences at Vanderbilt University, who agreed to serve on my dissertation committee. His expertise in developmental speech perception and

Speaking of cheerleading, there is no cheerleader more dedicated and unwavering as my best friend, Sarah Collins Prince. Regardless of where I was in the process, or how much I accomplished at any given time, I have written proof in the form of *many* cards reminding me of how I *was* going to finish, how I was made to do this, and that she had the utmost confidence in my abilities. I know I couldn't have done it without her. In addition, the ladies from my Grace Community Church Small Group, including Patti Bryce, Kelly Crider, Susan Resar, Suzanne Ward, and Katy Waugh, also deserve accolades for listening to several years of weekly updates regarding this process, and providing their own individual type of support and encouragement along the way, including pretending to enjoy talking about fricatives and cleft lip and palate.

I had the honor to spend one night a week, for 3 years, with a special group of high school girls, known as my Grace Community Church High School Youth Ministry Small Group. I met these girls as they were entering their sophomore year of high school, and I'm relieved that I was able to graduate from UNC before they graduated from college! Now they are college sophomores and my close friends. These special women, including Amy, Tracy, Megan, Caitlin, Hillary, Colleen, Hannah, and Caroline, always reminded me of who I really was when I had trouble remembering, and they spoke this truth to me every week, regardless of how much writing I had finished or articles I had read. Their love and support was invaluable in terms of enriching my life and providing me with so much joy during this difficult season.

In general, the support of all those I can call "friend" was so important for me. I am thankful for *all* of my friends who feigned interest in fricatives and

asked tirelessly about how my dissertation was going and what my dissertation was about. My colleagues at Duke, including Christine Beagle, Christie Range, Dan King, Ann Mabie, and Barron Suarez always had encouraging words, and I'm so grateful for the opportunity to continue my involvement in the Cleft Palate Board at Duke Medical Center during a few of my years during my program. Those Wednesday clinic days were sometimes the brightest spots in my week. Thanks to Marian Player to not only employing me for a year and a half, but also for supporting my data collection in any way she could, in addition to being a listening ear and constant support. My fellow Ph.D. students were always willing to listen, have lunch, chat, and make my time at UNC so much more fun – thank you Hillary, Eric, Angela, and Tom, and special thanks to Gary Martin, who was there every step of the way, in so many ways. Within the American Cleft Palate – Craniofacial Association, I had the pleasure of benefiting from long distance mentorship from so many wonderful speech – language professionals, including Helen Sharp, David Jones, Mary Hardin-Jones, Kathy Chapman, Karlind Moller, Mike Karnell, Jerry Moon, and Gip Seaver. Special thanks to my overseas-best-friend in David Fitzsimons. I also am thankful to the National Office staff of Nancy Smythe, Jeremy Baggish, Rafael Goldberg, Lisa Gist, and Morgan Gregson, who were always willing to let me crash at the office and do some work, or just stop by and say hello and lend a listening ear.

Within the Division, I am thankful to other faculty and staff who made the journey a little easier. Jack Roush always had a way of turning any negative I encountered into a positive, and I always knew he was on my side for whatever I needed. Lori Leibold was always willing to share her time, her advice and

counsel, and even a sample of IRB forms when I needed the support in a major time crunch. I appreciate the support and friendship from Lisa Domby and Ina Diana over the years.

To my future colleagues at Portland State – thanks for choosing me to join your faculty! I'm also thankful for the support I've received in the final stages of this process. Rik, Christina, Tom, Amy and Shelly – I thank you for your enthusiasm and encouraging emails over the summer as I finished my project; in addition, I'm thankful for the clinical faculty of Ellen, Claudia, Susan, and Desirae, because everyone was so gracious during my interview and excited about my imminent arrival!

Some final thanks – to the parents and children who participated in my study, some of whom I've had the pleasuring of knowing for the entirety of their child's life through the Duke Cleft Palate Team. My own family has always supported my decisions in life, and my father particularly always knew I would take this path. My sister always managed to have great ideas to support my studies, and I am appreciative to all of them for encouraging me throughout this process.

And last but not least, my David. I wouldn't have and couldn't have chosen anyone better to journey with me through this experience. Who else would have put up with my late nights, crazy schedule, attention deficit disorder, anxiety and irritability, and expert procrastination – and yet never judged me. Unconditional love and support was all I ever felt during this entire process. I am so lucky to have you as my very best friend and partner in this life.

TABLE OF CONTENTS

xiv

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

Typically developing children are in the process of honing both speech perception and production skills throughout early childhood and beyond. Although research focused on these skills has provided information regarding the course of development, there is still much that is unknown about how children advance their abilities in both speech perception and production beyond early childhood.

Evidence in the literature on speech perception and production skills based on studies with adult participants has indicated that these skills are not static even in the mature speech system (e.g., Shiller, Sato, Gracco, & Baum, 2009). In addition, research has shown that individual speakers show different perception and production styles, and that perception and production are interrelated (e.g., Perkell, Matthies et al., 2004).

It is possible that speech perception and production skills are interrelated in typically developing children, just like they seem to be for adults. Limited research has been done to provide information regarding the possibility of the relationship between perception and production in typically developing children. The pattern of development of these skills beyond early childhood is not well understood. Most research studies examining these skills in children have focused on analysis of group data rather than individual speaker or listener

patterns. This is problematic because it is the individual child that is the concern

for the speech – language pathologist.

Knowledge about the development of speech perception and speech

production and the relationship between them is particularly important for

studying speech disorders in children. The treatment of speech disorders is

contingent upon adequate diagnosis among one or many potential underlying

factors, including poor or abnormal phonological organization, deficits in motor

planning and execution, reduced or impaired perceptual abilities, and structural

variations and abnormality in dental and palatal morphology.

The current research study was designed to increase our understanding of

the range of perceptual and productive performance in individual typically

developing children, ages 7 and 11 years old. The current investigation

specifically focused on the phonetic contrast of the voiceless sibilant fricatives [s]

and [ʃ], as this particular contrast is known to emerge later than other contrasts

in typically developing children (e.g., Smit, Hand, Freilinger, Bernthal, & Bird,

1990). Both production and perception of [s] and [ʃ], in addition to the

relationship between these skills, were examined in this study. Developmental

differences between the two age groups were investigated as well.

As the model for a system with disordered speech, we examined a small

group of children with repaired cleft lip and palate. Prior to palate repair, early

motor experiences with speech in children with cleft palate are inevitably affected

by vocal tract differences, because the velopharyngeal incompetence caused by an

open cleft does not allow for motor practice of obstruent consonants, such as

stops and fricatives. In turn, this lack of practice may negatively affect the

development of stable auditory, proprioceptive, and tactile sensory targets, thereby perpetuating obstruent production difficulties even after the cleft has been repaired. Another risk factor for abnormal speech perception development in this population is the likelihood of hearing loss during critical periods of development. In addition, children with cleft palate are at high risk for eustachian tube dysfunction and middle ear effusion, introducing the possibility of fluctuating or persistent conductive hearing loss until the time of palate repair and/or the placement of pressure-equalizing tubes.

Production of fricatives is known to be particularly problematic for children with cleft palate, even after primary repair, and it is likely that factors related to early or persisting production and perception challenges contribute to the difficulty. Therefore, in the present study, children with repaired cleft lip and palate who present with distorted fricative production were compared to typically developing children on measures of production and perception of [s] and [ʃ]. Qualitative and quantitative differences in production and perception skills between the children with repaired cleft lip and palate and typically developing children were examined.

Summary

By examining the relationship among perceptual, structural, and motor variables likely to affect development of speech perception and speech production, clinicians may be able to identify underlying mechanisms that vary from child to child and use this information to guide the selection of effective interventions for children with speech disorders. To support such clinical applications, it is necessary to determine the nature of the relationship between

3

perception and production skills in typically developing children and examine the

range of normal variation.

CHAPTER 2

Literature Review

Introduction

The following chapter will provide background information regarding what is currently known about typically developing children's production and perception of fricatives. It will begin with an introduction to spectral moment analysis, the acoustic analysis method utilized in this study to examine children's fricative production. The literature will be reviewed with a focus on what is known about the development of fricative production and fricative perception in both individual children and groups of children. A summary of the research on the relationship between production and perception will be provided, including a review of a current model of speech motor control development that serves as the theoretical framework on which this study is based. Finally, information regarding what is known about perception and production in children with cleft palate will be reviewed.

Acoustic analysis of children's speech

Clinical interpretation of children's speech production skills relies on auditory perceptual analysis of the speech signal. In practice, the interpretation is typically mediated by orthographic and phonetic transcription. This method is convenient and easily accessible to the practicing clinician. Unfortunately, not all aspects of speech production are adequately captured through transcription

letters and symbols. For example, information regarding sound distortion is often lost by forcing a choice of one symbol or another to represent more complex variations. To preserve qualitative detail, the auditory perceptual analysis can be supplemented with acoustic analysis. With the advent of digital recording and analysis technology, many acoustic measures have become highly accessible to the practicing clinician. Some subjective judgment is required to determine speech segments for acoustic analysis, yet these measures provide additional information about the speech signal that may be unavailable to the ear of the practicing clinician when auditory perceptual analyses alone are utilized.

 Spectral moment analysis. One acoustic measure that has shown promise in objectifying and quantifying the acoustic signal is spectral moment analysis. In the acoustic study of voiceless sibilant consonants [s] and [ʃ], spectral moment analysis has become a particularly popular method. Forrest, Weismer, Milenkovic, & Dougall (1988) described the application of spectral moment analysis as a measure to analyze and quantify obstruent spectra (stops and fricatives) objectively. In the moment analysis approach, the aperiodic energy of the fricative is treated as a random probability distribution, with the first four moments of mean, variance, skewness, and kurtosis being derived from Fast Fourier Transformations (FFT) (Forrest et al., 1988). The first four moments characterize the concentration (or "centroid"), spread, tilt, and peakedness, respectively, of the fricative spectrum. The use of spectral moment analysis for investigating fricative consonant production is appealing for two reasons. First, it provides *quantitative* information about select qualities of the aperiodic noise spectra of fricatives that can be used for documentation and analysis purposes,

whereas the three-dimensional nature of spectrographic analyses mandates a qualitative interpretation. Second, it can document the presence or degree of categorical distinction in production of consonants that may or may not be perceptible to the ear, thereby providing a basis for comparison between disordered and normal speech. Spectral moment analysis successfully distinguishes sibilant ([s z ʃ ʒ]) from nonsibilant ([f v ɵ ð]) fricatives, and the sibilant fricatives from one another (i.e. [s] versus [ʃ]). It has not been as successful, however, to categorize the nonsibilant fricatives from one another. This may be due to the fact that sibilant fricatives are characterized by more distinctive spectral shapes, in comparison to the nonsibilant fricatives, which have a relatively flat spectrum. The nonsibilant fricatives have not been found to display acoustic properties that clearly differentiate between them compared to the sibilant fricatives (Jongman, Wayland, & Wong, 2000).

Spectral moments have been utilized by researchers over the past 20 years to study fricative production in both child and adult speakers for a variety of purposes. Investigations focused on determining an invariant measure for categorization of fricative consonants by place of articulation have employed this measure (Flipsen, Shriberg, Weismer, Karlsson, & McSweeny, 1999; Forrest et al., 1988; Fox & Nissen, 2005; Hagle, 2002; Jongman et al., 2000; Nissen & Fox, 2005). Moment analysis has also been used to substantiate developmental trends in typically developing children and to examine differences between child and adult speakers (Baum & McNutt, 1990; Nittrouer, 1995; Nittrouer, Studdert-Kennedy, & McGowan, 1989; Nittrouer, Studdert-Kennedy, & Neely, 1996). In

addition, researchers have utilized spectral moment analysis to investigate in adult speakers the relationship between acoustic properties of spoken fricatives and measures of fricative perception (Newman, 2003; Newman, Clouse, & Burnham, 2001; Perkell, Matthies et al., 2004).

The first spectral moment (FSM), specifically, is sensitive to changes in place of articulation for obstruent consonants, such as [s] and [ʃ]. Recent research has shown that spectral moments derived at fricative midpoint distinguish consistently between [s] and [ʃ] in normal adult speakers (Hagle, 2002), but not in speakers with impaired speech and language following stroke (Haley, 2002; Haley, Ohde, & Wertz, 2000).

In summary, spectral moment analysis is one acoustic analysis tool that shows promise in quantifying and objectively studying the speech signal, by providing additional information beyond auditory perceptual analysis or transcription alone. This type of analysis method can augment auditory-perceptual analysis in the study of speech production, by providing objective, quantitative information regarding specific aspects of the complex speech signal.

Fricative Production

Voiceless sibilant fricatives. The voiceless sibilant fricatives [s] and [ʃ] are characterized by high-intensity, aperiodic sound energy. This sound energy is generated by airflow from the lungs passing through a narrow constriction created in the vocal tract. The airflow passing through the constriction creates turbulence and therefore aperiodic noise. The [s] and [ʃ] are produced with distinctive frequency energy due to different places of constriction within the vocal tract. The consonant [s] is produced anteriorly in the vocal tract, by

8

creating a constriction using the tongue and alveolar ridge. The consonant [ʃ] is produced posterior to [s], created by a constriction made by the tongue approximating the palatal region (Kent & Read, 2002). These different articulatory configurations affect the acoustic output. An early study completed by Hughes & Halle (1956) found that the spectral energy in [s] was consistently higher than [ʃ]. The authors attributed this difference to variations in length of the vocal tract between the point of maximum constriction and the lips. The difference in spectral content between [s] and [ʃ], with [ʃ] exhibiting lower resonating frequencies, may also be due to the presence of a sublingual space (space under the tongue), created when the tongue approximates the palate (Perkell, Boyce, & Stevens, 1979; Perkell, Guenther et al., 2004). Regardless, [s] is known to be produced with greater frequency concentration at higher frequencies than [ʃ]. Therefore, the FSM, which reflects the spectral mean, is a useful measure for distinguishing between [s] and [ʃ].

Differences in fricative production between adults and children. Studies that have utilized analysis of the frequency information contained in fricatives to examine production have compared adult and child speakers, with the goal to better understand how fricative production develops. Research has confirmed that the spectral energy of the fricatives [s] and [ʃ] produced by children contains higher frequency energy compared to adults (McGowan & Nittrouer, 1988; Pentz, Gilbert, & Zawadzki, 1979). This finding has been attributed to vocal tract length differences between adults and children, with larger adult vocal tracts resonating lower frequencies for fricatives compared to children.

A line of research by Nittrouer and colleagues (Nittrouer, 1995; Nittrouer et al., 1989; Nittrouer et al., 1996) explored the differences between adult and child speakers' fricative production using spectral moment analysis. Both child and adult speakers were found to produce a FSM distinction between [s] and [ʃ], and the magnitude of the distinction was found to increase with greater age of the subjects (Nittrouer et al., 1989). The child speakers were ages 3, 4, 5, and 7 years. The 7-year-old speakers produced a greater distinction between the fricatives than did the younger children, but a lesser distinction when compared to adult speakers (Nittrouer et al., 1989). While the results of FSM magnitude of distinction increasing with age were subsequently replicated in studies that followed (Nittrouer, 1995; Nittrouer et al., 1996) and by other researchers as well (Nissen et al., 2005), these studies did not study children older than age 7 years.

Variability in fricative production. Studies of variability in children's speech have sought to understand better the developmental changes in different acoustic parameters as age increases, and how these changes reflect an approach to adult-like speech characteristics. The focus of studies on variability in children's speech have included examining changes in fundamental frequency, formant frequencies, temporal parameters, and spectral parameters of children's speech compared to that of adults (Eguchi & Hirsch, 1969; Kent, 1976; Kent & Forner, 1980; Lee, Potamianos, & Narayanan, 1999; Munson, 2004; Smith, 1978; Smith, Kenney, & Hussain, 1996). Results of these studies have found greater token-to-token variability in the children's speech compared to adults, and variability that decreases as children increase in age. For example, in the work by Lee and colleagues (1999), the authors conducted a large scale, cross-sectional

study of acoustic parameters in children's speech. Lee and colleagues included 436 children age five to seventeen years of age and incorporated a group of 56 adults for comparison purposes. The authors examined the parameters of vowel duration, [s] duration and sentence duration, fundamental frequency, vowel formant frequencies and spectral-envelope variability in vowel production. There was a global trend of decreasing segmental durations of vowels, [s] production, and sentences between age 9 or 10 to age 12 or 13 that approached adult values, as well as a significant decrease in within-subject duration variability of [s] production between the ages of 8 to 14 years. The results of the study confirmed that decreasing magnitude and within-subject variability of all the acoustic parameters studied corresponded to an increase of age of the children, and this phenomenon could be generally accepted as a developmental effect of age on speech production.

Studies specific to acoustic parameters of fricative production have found similar results indicating children demonstrate greater variability in speech production tasks compared to adults. Munson (2004) investigated temporal and spectral variability of [s] produced by three groups of children, with mean ages of 3 years 11 months, 5 years 4 months, and 8 years 4 months, compared to adults. The children were found to produce [s] with greater temporal and spectral variability than the adult speakers, but unlike other studies of [s] duration, Munson did not find mean duration differences among any of the groups of children or between the children and adult speakers. In addition, although Munson found that the adults produced fricatives with less spectral variability than the child speakers, the child speakers did not display statistically significant

11

spectral variability measures among groups. Another study by Koenig, Lucero and Perlman (2008) investigated production variability in the fricatives [h], [s] and [z] in two groups of children, age 5 and 10 years, and adults. The authors found that the child groups demonstrated greater variability compared to adults on measures of amplitude variability (representing variability in airflow management during consonant production) and temporal variability. The children's [s] productions had the greatest amount of amplitude variability compared to the other two fricatives. This finding was interpreted by the authors as representing the articulatory complexity of [s] production.

Children with speech disorders display greater variability in acoustic parameters of [s] production compared to typical controls (Baum et al., 1990; Weismer & Elbert, 1982). Weismer and Elbert (1982) investigated [s] production in two groups of children between the ages of 4 and 6, and one group of adult speakers. There were seven children in each child speaker group, with one group representing children with normal articulation and the other representing children who misarticulated [s]. The investigators found that the children with incorrect [s] productions produced the consonant with greater temporal variability than the same-aged children with accurate [s] production. In addition, the children who produced [s] accurately demonstrated greater temporal variability than the adult speakers. Baum and McNutt (1990) expanded the findings of Weismer and Elbert (1982) by investigating temporal, intensity, and spectral parameters of [s] production in 2 groups of 6 and 8 year old children who produced a frontal misarticulation of [s] and those who produced [s] accurately. The authors found greater average inter-subject variability for the

measures of fricative duration and fricative spectra (using FSM analysis computed at the approximate midpoint of the fricative) for [s] production in the two groups of children who misarticulated [s] compared to the normal controls. In addition, the older subjects demonstrated less variability in the temporal measures of [s] production compared to the younger subjects, consistent with what was found in the Weismer and Elbert (1982) study.

Summary of research on fricative production. Acoustic studies of fricative production in children have provided evidence for a spectral distinction between the fricatives [s] and [ʃ] in child speakers (e.g., Nittrouer, 1995). In addition, the research has shown that variability in specific aspects of fricative production such as temporal and spectral parameters decreases as children become older (e.g., Lee et al., 1999). One limitation in this research is that while the cross-sectional studies reviewed have provided evidence of specific developmental trends, studies investigating individual speaker patterns are limited. The analysis of group data does not allow for the investigation of specific developmental patterns of individual children. Because absolute acoustic values can vary dramatically from speaker to speaker and because individual children, like individual adults, often behave differently from the group average, valid clinical applications require knowledge about the range and stability of normal performance. If typically developing children, like adults, produce a consistent spectral moment distinction between [s] and [ʃ], then it should be possible to determine the accuracy of that distinction regardless of speaker variability in absolute or relative values of the target phonemes.

Fricative Perception

Assessment of speech perception. Speech perception develops during early childhood. Research has documented changes in children's perception of fricatives within the first decade of life. During infancy and early childhood, the ability of children to extract meaning from the complex speech signal is honed as they acquire more experience with their native language (Nittrouer, 2002).

One widely-used method of assessing speech perception skills in children includes the creation of synthetic speech stimuli for use in perception tasks. Within these synthetic stimuli, specific acoustic parameters are varied. The parameters to be varied are chosen on the basis of their believed relevance to perception of specific cues within the speech signal. The acoustic parameters or cues being assessed are varied systematically along one or more acoustic dimensions that are believed to underlie a phonetic contrast (Strange, 1995). The synthetic speech stimuli can then be used in perceptual tasks such as identification or discrimination procedures, both of which assess categorical perception of phonetic stimuli. In identification testing, a listener is asked to label stimuli, presented one at a time, in a forced-choice response format, using phonetic targets provided by the examiner (for example, choosing whether a stimuli heard was "see" or "she"). Evidence of categorical perception is observed if the listener demonstrates an abrupt boundary between the phonetic contrasts being tested. Discrimination testing requires the listener to make a judgment regarding the comparison of two or more stimuli that are presented from a single continuum. Stimuli presented from the continuum can be very similar (for example, 1-step discrimination) or less similar (2- or 3-step discrimination). One

example of discrimination testing is an ABX paradigm (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). In an ABX paradigm, the listener is presented with three stimuli; the first two stimuli presented are different from one another, and the third stimulus is the same as either the first or the second. The listener is required to make a judgment regarding which of the first two stimuli presented is the same as the third stimulus presented. Evidence for categorical perception on discrimination testing would be observed if the listener demonstrated better discrimination of stimuli that cross the phonetic boundary versus those within the same phonetic boundary (Polka, Jusczyk, & Rvachew, 1995).

Development of fricative perception. With regard to fricative perception specifically, research has shown that younger children are more attentive to dynamic cues present in fricative + vowel segments (formant transitions between the fricative and the vowel), whereas older children and adults use static information (spectral content) contained in the fricative alone to make decisions regarding fricative identity. Nittrouer (1992) found that younger children, ages 3 - 7 years, attended to formant transitions between fricative and vowel segments (dynamic properties) rather than to the frequency content within the fricative (static properties) when making decisions about fricative identity. As an example, the 5-year-old children demonstrated shallower identification curves for [s] stimuli compared to adults, indicating an increased reliance on the vocalic portion of the stimuli to make decisions of fricative identify. These results have been replicated in other investigations by Nittrouer and colleagues (Nittrouer, 2002; Nittrouer, Crowther, & Miller, 1998). Evidence for a shift from attention to

15

dynamic properties to static properties was found by Nittrouer (2002) at approximately 8 years of age; performance on measures of fricative perception by children at age 8 years resembled that of adults (Nittrouer, 2002). Nittrouer and colleagues (Nittrouer, Manning, & Meyer, 1993) termed this phenomenon the "Developmental Weighting Shift," reflecting a change from a reliance on formant transitions to fricative noise to make decisions regarding fricative identity.

The concept of the Developmental Weighting Shift has been challenged by researchers who have found that the dynamic properties of consonant+vowel syllables may not be relevant cues for perception of all consonants. For example, Ohde and colleagues (Ohde, Haley, Vorperian, & McMahon, 1995) examined perception of stop-consonants in various vowel environments in consonant-vowel syllables in children age 5, 6, 7, 9, and 11 years compared to adult listeners. The results showed that children age 5 and older were not relying on the information contained in dynamic formant transitions any more than adults were when making decisions regarding stop-consonant identity (Ohde et al., 1995), but rather attended to properties of the noise burst from the stop consonant. However, dynamic formant transitions were found to be important for perception of the velar stop consonant [g] in children age 3 and 4 years of age (Ohde & Haley, 1997). These results, in comparison to the theory of developmental speech perception proposed by Nittrouer and colleagues (Nittrouer et al., 1993), indicate that children may develop different perceptual strategies for consonant perception at different developmental stages.

Individual perceptual performance. Similar to fricative production, current studies of fricative perception have analyzed group rather than individual

data. As an exception, Rvachew & Jamieson (1989) displayed identification functions of individual children to illustrate differences in perceptual performance between typically developing children and children with articulation disorders. The authors investigated the relationship between perception and production of [s] and [ʃ] in two groups of 5-year-old children and adults. One group of children presented with an articulation disorder, consisting of at least one misarticulated sound, and the other group of children were judged to have articulation skills within normal limits. Ten of the twelve children in the group of children with an articulation disorder demonstrated errors on [s] and/or [ʃ]. Only five of the twelve children in the misarticulating group demonstrated normal identification functions for a [s] - [ʃ] continuum, and three of these five children misarticulated [ʃ] and/or [s]. The results of this study indicated that the group of children with the articulation disorder demonstrated different perceptual skills from one another, and therefore could not be considered a homogeneous group. Therefore, the authors underscored the need for attention to individual participant performance, particularly when studying clinical populations, rather than examining averaged performance across participants.

Summary of research on fricative perception. Research on fricative perception in children has provided evidence for changes in perceptual strategies for fricative identification over the course of development in early and middle childhood years. Attention to individual patterns or abilities of perceptual skills within different age groups is limited in the current literature, particularly for children older than 7 to 8 years of age. Like fricative production, analyzing group data can provide information regarding developmental trends, but investigating

performance on perceptual tasks in individual typically developing children can provide information about the range of normal performance and normal variability.

Relationship between Production and Perception

Relationship between production and perception in typically developing children and adults. A speaker's perceptual performance may have a direct relationship to a speaker's productive abilities and vice versa. Shifting focus from group comparisons to individual performance can inform the study of specific strengths or deficits in either perception or production and the relationship between them. Evidence of a relationship between speech perception and production has been found in both child and adult speakers (Broen, Strange, Doyle, & Heller, 1983; Hoffman, Daniloff, Bengoa, & Schuckers, 1985; Monnin & Huntington, 1974; Ohde & Sharf, 1988; Perkell, Guenther et al., 2004; Perkell, Matthies et al., 2004). Many of the studies investigating this relationship in children have included groups of children with specific articulation disorders in addition to children with articulation skills within normal limits. For example, identification and discrimination of synthetic [r] and [w] have been examined in children who misarticulate these sounds, and children with normal articulation skills (Broen et al., 1983; Hoffman et al., 1985; Monnin et al., 1974; Ohde et al., 1988). Monnin & Huntington (1974) examined phonetic identification of four contrasts presented in word minimal pairs, including the [r] - [w] contrast, acoustically similar contrasts (such as [ʃ] - [tʃ]), acoustically dissimilar contrasts (such as [t] – [n]), and vowel contrasts, in 3 groups of children: 6-year-old children with normal articulation, 6-year-old children who misarticulated [r] in

18

the initial, medial, and final position of words, and 5-year-old children with normal articulation. The stimuli were presented in six conditions, ranging from no distortion (Condition 1), to maximum distortion (Condition 6). The consonants were distorted using "center clipping" (Monnin et al., 1974, p. 358), a type of amplitude distortion, and the vowels were distorted using high-pass filtering. The stimuli were distorted in order to decrease the redundancy of the speech signal and therefore increase the difficulty of the identification task across conditions. The children who misarticulated [r] demonstrated diminished performance on the identification task contrasting words containing [r] and [w] compared to the two groups of children with normal articulation, but performed as well as the children with normal articulation on identification tasks assessing all other consonants and vowel contrasts. The authors interpreted these results as providing evidence for a specific perceptual deficit in the children who misarticulated [r], related to their production abilities, rather than a general deficit in perceptual skills.

The study by Broen and colleagues (1983), mentioned earlier, investigated perception of three perceptual contrasts ([w] - [r], [w] - [l], [r] - [l]), synthetically generated, in 3-year-old children with normally developing articulation skills for their developmental age, and in 3-year-old children who had been diagnosed with delayed articulation development. Articulation development was assessed using a standardized test of articulation. Approximately 25% of the children in the normally articulating group demonstrated articulation errors on [l] and 63% misarticulated [r], whereas 83 % of the children in the delayed articulation group misarticulated [l], and 100% misarticulated [r]. On the identification task, which

utilized three minimal-pair contrasts (wake-rake, rake-lake, and wake-lake), both groups of children performed well, with the children in the delayed articulation group achieving 90% accuracy. However, the variance in perceptual performance between the two groups was the interesting finding, with the children in the delayed articulation group demonstrating greater intra-subject variability in responses compared to the group of children with normally developing articulation skills. The authors concluded that given the results of their study, some children, but not all, will have difficulty perceiving phonetic contrasts that they have yet to master.

In another study, Hoffman and colleagues (1985) used synthetic stimuli representing a continuum from [r] to [w] in order to examine identification and discrimination skills in two groups of 6-year-old children. One group misarticulated [r] as [w] in the initial position of words, and another group had normal articulation skills. A seven-step continuum of [re] – [we] was synthesized and used for the identification and 3-step discrimination trials. The children who misarticulated [r] achieved flatter identification functions, suggesting less consistency of responses compared to the children with normal articulation skills, but both groups demonstrated the same phonemic boundary between step 3 and 4 of the synthetic continuum. On the discrimination task, the children who misarticulated [r] demonstrated chance performance (mean = 57%), whereas the children with normal articulation achieved an average discrimination score of 80%. In their discussion, the authors questioned whether perceptual abilities had developed prior to mastery of production, as concluded in the Broen et al.

(1983) study, or if the ability to form perceptual categories may be shaped by a child's ability to produce the consonant.

Ohde and Sharf (1988) examined identification of a synthesized [r] - [w] continuum in two groups of 6 to 7 year old children, one of which consisted of children who misarticulated [r] and the other consisted of children with normal articulation. In addition, a group of adult participants were included for comparison purposes. The results of the study provided evidence for a perceptual deficit for identification of a [r] - [w] continuum in the group of children who misarticulated [r], similar to the results of previous studies. In addition, similar performance of the children who misarticulated [r] and the children with normal articulation was found on identification of a [b] - [w] continuum, consonants which both groups of children produced correctly. These findings provided evidence that the perceptual deficits in the children who misarticulated [r] were related to their production error, similar to the findings of Monnin and Huntington (1974). That is, they had difficulty identifying a continuum of a sound they misarticulated, whereas they performed the same as the normally articulating group on a contrast that was not incorrectly produced.

Fewer studies have examined the relationship between perception and production of the [s] and [ʃ] targets (see Rvachew and Jamieson, 1989, as an exception) in child participants. However, evidence for a relationship between perception and production has been found in normal adult speakers for the [s] - [ʃ] contrast. Perkell and colleagues (2004) examined the magnitude of distinction in production between [s] and [ʃ] in adult speakers, measured by spectral moment analysis, and discrimination skills of the [s] - [ʃ] contrast using

synthetic fricative noise and naturally produced vowel portions. The investigators used 2-step comparisons in order to assess auditory discrimination abilities, and separated the listeners into "high" and "low" acuity groups. The participants in the "high" group had achieved 100% accuracy on the discrimination task, and the participants in the "low" group had performance ranging from 60 – 90% accuracy. Perkell and colleagues found that the adults in the "high" acuity group produced the fricatives with a greater magnitude of difference in overall mean FSM compared to the adults in the "low" acuity group. The cross-subject correlation between discrimination score and sibilant contrast was found to be significant (r = .63, p < .01). The authors related their findings to the possible effect of this relationship in the development of speech in children; they stated, "in learning to maximize intelligibility, the child with higher acuity is better able to reject poor exemplars of each phoneme, and thus will adopt sensory goals for producing those phonemes that are further apart than the child with lower acuity" (Perkell, Matthies et al., 2004, p. 1267). Typically developing children may show a similar relationship between perception and production as the adults in the Perkell et al. (2004b) study; however, the relationship between perception and production in typically developing children has not yet been investigated.

    The DIVA model. The DIVA (Direction Into Velocities of Articulators) model (Guenther, 1995; Guenther, Hampson, & Johnson, 1998) provides a theoretical framework for investigating the relationship between speech perception and speech production in children. One central assumption of the DIVA model is that before a typically developing infant can produce a stable

speech sound target, the infant must be able to perceive that sound accurately and reliably (Guenther, 1995). The development of speech perception skill occurs within the context of the infant's native language. The DIVA model purports there are invariant auditory perceptual targets or "goal regions" which the speaker attempts to achieve (Guenther et al., 1998). This assumption is made given that auditory feedback information is available faster to the central nervous system than tactile and proprioceptive feedback during motor planning for speech production (Guenther et al., 1998). Therefore, speakers with finely honed speech perception skills may develop goal regions for speech sound movements that are more distinctly contrasted than speakers with less precise perceptual skills. If that is true, these speakers may be more likely to reject poorly produced speech targets than speakers with less precise speech perception abilities, whose goal regions for speech sound targets are not as well defined (Perkell, Guenther et al., 2004).

The assumption of the DIVA model that perception is intact when speech motor control is developing introduces the question of whether production abilities have an effect on early perceptual development. The concept that speech perception and production co-evolve in a child's early development is central to the model, with the idea that not every sound of the child's native language needs to be accurately perceived before speech motor learning can occur. However, a child's inability to produce certain sounds due to structural limitations, for example in children with unrepaired cleft palate, and the effect of production limitations on perceptual development, has not yet been investigated. Therefore,

23

to understand better specific relationships between perception and production, it may be informative to study children with cleft palate.

    Variables of the DIVA model possibly affected by cleft palate. Early development of speech motor control and perceptual development are likely affected by a cleft palate due to the nature of the defect. A cleft palate causes an alteration to the vocal tract that does not allow for the buildup of oral air pressure for the production of obstruent consonants, including the fricatives [s] and [ʃ]. Research on early speech sound development in children with cleft palate has demonstrated the absence of these consonants in a child's phonemic repertoire when a cleft palate is present (Chapman, 1991; Chapman, Hardin-Jones, Schulte, & Halter, 2001). While not directly stated in the DIVA model, the question may be raised of whether children who lack productive ability for obstruent consonants may have delayed or disordered perceptual development for these consonants. In addition, the issue of whether establishment of goal regions for auditory perceptual targets is delayed until the child has had the cleft repaired could be raised. Cleft palate repair typically occurs between 8 to 14 months of age in the United States, at an age when canonical and reduplicated babbling typically appears. The primary goal of palate repair is to establish normal velopharyngeal function in a child with a cleft palate. The palate repair closes the communication between the nasal and oral cavity and repositions the muscles of the soft palate, so the child has the ability to approximate the soft palate to the posterior pharyngeal wall and create oral air pressure to facilitate development of oral consonant production.

Children with an unrepaired cleft palate are prone to eustachian tube dysfunction, which introduces the possibility of fluctuating or persistent conductive hearing loss due to the presence of fluid in the middle ear. The effect of conductive hearing loss on the development of speech perception abilities in these infants is not known. In one study, 7- and 8-year old boys with cleft palate were compared to typically developing controls on performance on a perceptual speech discrimination task employing an ABX paradigm and words that differed on minimal initial and final consonant pairs (Finnegan, 1974). The children with cleft palate and a positive history of otitis media did not perform as well as the children with cleft palate and negative otological history on the test of speech discrimination in this study. Therefore, the poorer performance was attributed to a history of recurrent otitis media. In children without cleft palate, effects of severe, recurrent otitis media during early childhood on development of speech perception skills have been found in children with a history of otitis media (OM) with and without receptive / expressive language delays (Clarkson, Eimas, & Marean, 1989). In this study, the authors compared identification functions and discrimination scores for [b] – [p] stimuli differing in voice onset time (VOT) (duration of time between release of stop consonant and start of vowel production) of in three groups of 5-year-old children: one group of control children with no history of OM and language skills within normal limits, a second group of children with a documented history of OM and language skills within normal limits, and a third group of children with a documented history of OM and measurable language delays. In general, the children with OM and language delays showed significant differences on the identification tasks compared to the

25

control group and the group with OM only. Performance on the discrimination task, however, was markedly different for all groups, with the two groups of children with OM performing more poorly than the control group. In contrast, recent work by Paradise and colleagues (Paradise et al., 2005) found no difference in developmental outcomes, including speech and auditory processing measures, in children with persistent otitis media with effusion who received either early or late insertion of pressure equalizing (PE) tympanostomy tubes.

In addition, reduced or altered tactile sensation of the palate after cleft palate repair may affect somatosensory feedback in children with repaired cleft palate. The effect of resultant scar tissue and possible nerve damage to the mucosal surface of the hard palate after the initial repair may create sensory differences in children with repaired cleft palate compared to typically developing children. These sensory differences could potentially cause reduced feedback in children with repaired cleft palate compared to typically developing children, which may affect development of speech motor control, with respect to the DIVA model. Evidence for this claim in children has not yet been studied. Significantly reduced intraoral tactile perception for both surface alteration and shape alteration, however, in adults with repaired cleft palate compared to typical controls, has been found (Hochberg & Kabcenell, 1967).

Relationship between perception and production in children with repaired cleft palate. The relationship between speech perception and production in children with cleft palate has been explored by Whitehill and colleagues (2003), who provided evidence that children, age 4 to 12 years, with repaired cleft palate and posterior placement of alveolar targets ([k] for [t]

substitution) also showed a deficit in classification of a continuum of [t] – [k] targets. Stimuli for the speech perception evaluation were generated synthetically and systematically ranged from the word [tʰau] (Cantonese for "head") at one extreme of the continuum to the word [kʰau] (Cantonese for "ball") at the opposite end of the continuum. The performance of the children with repaired cleft palate and posterior placement was compared to the performance of two age-matched control groups. One control group consisted of children with repaired cleft palate without posterior placement of [t], but with other articulation errors, and the second control group consisted of typically developing children with no history of speech, language, hearing or learning problems. All groups were matched for age. Like the normal controls, children with cleft palate and accurate production of [t] showed a clear phonemic boundary between [t] and [k] in their identification functions. In contrast, the children with cleft palate and posterior placement performed at chance levels for all items on the continuum.

Several interpretations of these results were offered. First, the children with repaired cleft palate and posterior placement may have exhibited perceptual deficits specific to their production errors. Some support for this view from this study is that children with cleft palate without posterior placement were able to identify the continuum of [t] to [k] similarly to the normal control group. However, the evidence presented in this study cannot rule out the possibility that the children with cleft palate and posterior placement were also unable to classify consonants other than [t] and [k], indicating a broader perceptual deficit not limited to their erred phonemes. A second explanation is that children with cleft

27

palate, as a group, demonstrate delays in receptive and expressive language skills, and the perceptual deficit shown by the children with posterior placement represented a delay in perceptual skill acquisition rather than a persistent deficit. However, because the children with posterior placement in this study represented a wide range in age (4 – 12 years), the authors speculated that their poor performance on the identification task may not be explained by developmental delay alone. Third, the authors concluded that the results of the study "may be interpreted as supporting the view that speech errors that are originally phonetic in nature, as a result of structural abnormality, may lead to phonological errors" (Whitehill et al., 2003, pg. 459). This interpretation supports the idea that lack of early motor practice with fricatives in children with cleft palate may adversely affect perceptual development of auditory and somatosensory targets for later speech production skills.   A final explanation of children with cleft palate possibly demonstrating an underlying perceptual deficit, could be attributed to the perceptual errors found in the children in the study by Whitehill and colleagues (Ceponiene et al., 2000; Ceponiene et al., 1999; Cheour et al., 1999; Cheour et al., 1998). Two of these studies specifically (Ceponiene et al., 1999; Cheour et al., 1998) examined auditory short term memory in school-age children (age 6 to 10 years) using an objective test of electric brain activity, event-related potentials (ERPs). One component of ERPs is called mismatch negativity (MMN), which estimates auditory short term memory abilities. For example, in one study (Ceponiene et al., 1999), children with cleft lip and palate were divided into subgroups based on cleft type. The children with palate involvement (cleft palate only, submucous cleft palate,

unilateral cleft lip and palate) demonstrated auditory short term memory dysfunction, compared to healthy controls and children with no palate involvement (cleft lip with or without alveolar involvement). The authors concluded that the deficits in auditory short term memory could be contributing to language and learning problems in children with cleft palate. Delays in language development in children with cleft palate have been documented in the literature (see Kuehn & Moller, 2000, for an extensive review). Whitehill and colleagues (2003) highlighted the lack of existing research on the perceptual skills of children with cleft palate, and indicated that additional studies are clearly needed in order to better understand the relationship between perception and production of errors specific to children with repaired cleft palate.

Summary of research on the relationship between production and perception. The research completed thus far on the relationship of production and perception in children has found that children with speech production disorders show increased variability in perceptual performance and some show shallower identification curves, indicating deficits in phonetic categorization. In addition, evidence exists for a relationship between perception and production in children with speech errors, which may or may not be error specific.

The study by Perkell et al. (2004b) provided evidence that normal adult speakers show a positive relationship between perception and production of [s] and [ʃ], and individual variation exists between groups of adult speakers. We could expect that this relationship may also exist in typically developing children, in the context of the DIVA model.

It should be noted that none of the studies on perception and production of fricatives in children utilized acoustic analysis of speech production in order to objectify, quantify, or describe production status in more detail. Rather, auditory perceptual judgment of misarticulated phonemes and/or performance on standardized articulation tests were used to determine accuracy of speech production in order to determine relationships between production and perception. Information regarding specific aspects of the speech production characteristics may be lost when using auditory perceptual assessment alone, and the information gained from acoustic analysis may be able to be related directly to perceptual performance.

Summary

A significant amount of research has been done to describe both children's perception and production of fricatives. The research on children's production of fricatives has not shown an invariant production distinction between [s] and [ʃ] on an individual level and has not evaluated intra-speaker consistency. Of the few studies that have investigated the relationship between perception and production in children, acoustic analysis has not been utilized to objectify and quantify speech production skills. It is important to continue to explore the relationship between perception and production of fricatives in typically developing children, in order to learn about the possible existence of invariance in production across individual child speakers and likely variability in perceptual performance. What can be learned about the relationship between perception and production can be applied to the assessment and treatment of children with speech sound disorders, such as children with cleft palate, who have structural

abnormalities affecting production and may have deficits in early speech perception or differences in sensory feedback that affect development of speech motor control.

Purpose of the Research Study

The purpose of this research study was to determine if a relationship exists between individual children's perception and production of voiceless sibilant fricatives. An additional goal of this study was to advance the knowledge of developmental differences in perception and production at ages 7 and 11 years. Finally, production and perception of [s] and [ʃ] were examined in a small sample of children with repaired cleft lip and palate, in order to compare their performance to the performance of typically developing children.

Research questions. The research questions were the following: 1) Do typically developing individual children age 7 and 11 produce a reliable distinction between the voiceless sibilant fricatives, measured by the first spectral moment? 2) Are there acoustic differences in fricative production between these two age groups? 3) Are there perceptual differences between these two age groups? 4) Is there a relationship between typically developing children's perception abilities and production characteristics for the voiceless sibilant fricatives? 5) How does the production and perception of voiceless sibilant fricatives in children with repaired cleft palate compare to typically developing children?

CHAPTER 3

Methods

Participants

Three groups of children participated in the experiment. One group

consisted of 10 typically developing children, 7 years of age (mean age = 7 years,

8 months; range = 7 years, 4 months – 7 years, 11 months) (TD-7 group), and a

second group consisted of 10 typically developing children, 11 years of age (mean

age = 11 years, 3 months; range = 11 years – 11 years, 11 months) (TD-11 group).

Gender was evenly distributed within these two groups. The third group

consisted of 5 children with repaired cleft lip and palate (CLP group), between 7

and 11 years of age, and all participants in this group were male. All participants

were native speakers of English, and all passed an audiometric screening of the

octave frequencies between and including 500 and 4000 Hz at a level of 30 dB

HL. Parents of the participants were asked to estimate the approximate number

of otitis media (ear infection) events they could recall their children experiencing

in their lifetime. The incidence was recorded by the examiner along with the

demographic information for the participants. Please see Table 1 for the

demographic data regarding participants.

Normal speech and language skills were confirmed in the TD-7 and TD-11

group, and normal language skills were confirmed in the CLP group. In the TD-7

and TD-11 group, participants were considered to have normal speech and

language skills if performance on the Goldman-Fristoe Test of Articulation (GFTA-2; Goldman & Fristoe, 2000) was within one standard deviation of the mean and if the appropriate cut-off score, based on age, was achieved on the Clinical Evaluation of Language Fundamentals – Screening (CELF-4 Screen; Semel, Wiig, & Secord, 2004). In addition, the participants in the TD-7 and TD-11 groups were judged by the examiner, a certified speech – language pathologist with 11 years of experience in the assessment of children with and without cleft lip and palate, to produce perceptually accurate tokens of [s] and [ʃ]. In the CLP group, achievement of the cut-off score according to age on the CELF-4 screen indicated language development within normal limits. The GFTA-2 as a speech assessment requirement was not utilized for the CLP group, due to the fact that including children with speech errors (i.e., distorted productions of [s] and [ʃ] or a lack of auditory-perceptual distinction in production of the [s] - [ʃ] contrast) was a specific aim of the study with regard to participants in the CLP group. The results of the speech and language screening are detailed in Table 1.

All participants were screened for occlusal status. The participants were instructed to close their mouth normally and expose their anterior bite pattern (having the participant smile). The purpose of this examination was to evaluate if any of the typically developing participants had excessive overjet (excessive protrusion of the upper teeth over the lower teeth) or reverse overjet (malocclusion in which the lower teeth protrude past or overlap the upper teeth). In addition, dental status was observed for any missing teeth. It was assumed that a significant malocclusal pattern such as excessive overjet or an underbite could have an effect on the aperiodic sound source generated during fricative

production. In order to control for the possible effects of occlusal status on fricative production, typically developing children who displayed either of these anterior bite patterns would be excluded from participating in the study. None of the typically developing children demonstrated either of these anterior bite patterns. One participant in the TD-7 group (B711) was missing an upper front central incisor. Although this same examination was conducted for the CLP group, participants were not excluded for malocclusion patterns, as these patterns are exceedingly common in children with repaired cleft lip and palate. Two participants in the CLP group (CP101, CP102) demonstrated underbites with an observable crossbite (lower anterior teeth overlap upper anterior teeth laterally), one participant (CP103) had an end-to-end anterior bite pattern (upper teeth approximate lower teeth along the incisal edges), one participant (CP104) had normal molar occlusion but a protruding premaxillary segment, and one participant had an end-to-end anterior bite pattern with an appreciable gap at the alveolar cleft site. In addition, all participants in the CLP group were judged to present with adequate velopharyngeal function due to the absence of hypernasal resonance and absence of consistent audible nasal air emission during speech production. This judgment was made by the primary investigator who has 11 years of experience in the speech assessment of children with cleft lip and palate.

Two 7-year-old typically developing participants were excluded from the study. Both participants did not pass the speech assessment due to the amount of developmental articulation errors produced. One participant in the TD-7 group (B707) was enrolled in the study initially and completed the entire study protocol. Although this participant passed the speech screening portion by

achieving a standard score within normal limits given his age and the normative values provided, he was observed to produce a frontal lisp on productions of [s] production. A decision was made to exclude the data from this participant because he produced a distortion on one of the targets under investigation. The results from this participant were not included in the final data analysis.

Table 1. Participant demographics and results of speech and language screening. Age is expressed in total months (years : months). OME is expressed in the approximate number of occurrences of otitis media (ear infection). Occlusal Status column indicates "OK" if no observable overjet or underbite existed or reflects description of occlusal status. GFTA column reflects raw score / standard score. CELF-4 Screen column reflects raw score / cutoff score for normal performance according to age (7-year-olds = 16, 11-year-olds = 19).

Group	Participant	Age	Gender	Occlusal Status	OME	GFTA	CELF-4 Screen
TD – 7	G701	88 (7:4)	Female	OK	0	1 / 104	21 / 16
	G702	92 (7:8)	Female	OK	0	0 / 105	24 / 16
	G703	93 (7:9)	Female	OK	0	3 / 99	25 / 16
	G704	93 (7:9)	Female	OK	1	0 / 105	26 / 16
	G705	95 (7:11)	Female	OK	1	0 / 105	20 / 16
	B706	93 (7:9)	Male	OK	5	0 / 105	27 / 16
	B708	89 (7:5)	Male	OK	4	0 / 109	22 / 16
	B709	92 (7:8)	Male	OK	10	0 / 105	24 / 16
	B710	94 (7:10)	Male	OK	10	0 / 108	24 / 16
	B711	88 (7:4)	Male	OK; missing one central incisor	3	0 / 109	20 / 16
TD-11	G1101	143 (11:11)	Female	OK	12	0 / 101	32 / 19
	G1102	139 (11:7)	Female	OK	3	0 / 101	29 / 19
	G1103	139 (11:7)	Female	OK	1	0 / 101	30 / 19
	G1104	132 (11:0)	Female	OK	6	0 / 101	30 / 19
	G1105	134 (11:2)	Female	OK	3	0 / 101	31 / 19
	B1106	133 (11:1)	Male	OK	0	0 / 105	31 / 19
	B1107	132 (11:0)	Male	OK	1	0 / 105	29 / 19
	B1108	132 (11:0)	Male	OK	3	0 / 105	34 / 19
	B1109	132 (11:0)	Male	OK	3	0 / 105	33 / 19
	B1110	135 (11:3)	Male	OK	0	0 / 105	32 / 19
CLP	CP101	134 (11:2)	Male	Underbite and crossbite	0	N/A	27 / 19
	CP102	115 (9:7)	Male	Underbite with crossbite and rotated central incisor	7*	N/A	24 / 17

CP103	98 (8:2)	Male	End to End	2*	N/A	27 / 18
CP104	85 (7:1)	Male	Protruding premaxillary segment and alveolar clefts	3	N/A	21 / 16
CP105	108 (9:0)	Male	End to End, significant gap noted at alveolar cleft site	7	N/A	26 / 17

* indicates PE tubes present in either one ear or both ears at time of testing

Sample size estimate. Sample size was estimated using a power analysis to calculate number of cases required to discover that the correlation is statistically significantly different from there being no correlation. For single-sided power of 80%, an alpha level of 0.05, and an effect size of $r = 0.5$, a sample size of 23 was estimated. Based on the recent research of Perkell et al. (2004b) and this power analysis, a sample size of 20 was adopted for the current study.

Procedures

Each participant in the TD-7 and TD-11 group completed the hearing screening, speech evaluation, and language screening before testing. Participants in the CLP group were required to complete the hearing and language screening only. Following these tasks, the participants completed a perceptual identification screening task. These tasks were completed in the participant's home or at the UNC Division of Speech and Hearing Sciences Laboratory Suite.

If the participants passed all of these screening items, then the participants and their parents were asked if they wished to complete the study protocol. If the participants and parents agreed to continue, then the participants completed the study protocol, including a production task, a perceptual identification task, and a perceptual discrimination task at the UNC

36

Division of Speech and Hearing Sciences Laboratory Suite, which included a

sound treated IAC booth. The study protocol was completed for all participants

within 4 weeks of the initial screening procedures. Most participants (23 out of

25) completed the screening tasks and study protocol in one testing session,

which lasted approximately 2 – 2 ½ hours in total. The complete study protocol,

including the perceptual identification screening task, will be detailed in the

sections that follow.

Perceptual Identification Screening

Equipment. A laptop computer (IBM ThinkPad R40) and the perceptual

Identification Task module of the SpeechMeasures (Haley, 2008) software

program was utilized for data collection. SpeechMeasures is a program designed

for speech elicitation and perceptual and acoustic quantification. The perceptual

identification screening task was completed in the sound treated IAC booth.

Participants wore headphones (Sennheiser, HD 250 Linear II) in order to block

out distractions from the environment and to allow them to focus on the task at

hand.

Procedures. A participant was presented with 10 tokens each of the

endpoints of a synthesized [si] – [ʃi] and [su] – [ʃu] continuum (40 tokens total),

over headphones at 80dB SPL in two separate identification tasks, blocked by

vowel. Specific information regarding the two continua will be presented in a

later section. Tokens were pre-ordered in a pseudo-randomized order in each

task (see Appendix 1 for the order of the tokens). For either task, a cartoon

picture and orthographic representation of the words "see" or "she" and "sue" or

"shoe" was presented on an 8 inch x 11 inch sheet of paper, with each picture

being approximately 4.5" by 4.5" inches and the words in 64-point font (Appendix 2 contains the pictures representing the words for both continua). A participant was required to identify, through finger point, which word he/she heard presented. A participant was seated at a table with the representative pictures for response in front of the child. The principal investigator sat next to the child in order to clearly view their pointing response. The participant's finger point response to either picture was recorded by the examiner clicking on the word representing the child's response on the laptop computer screen (a visual representation of the perceptual identification task module from a pilot version of the SpeechMeasures software is contained in Appendix 3). A participant was required to identify 90% of the endpoints (36/40) correctly in order to continue with the experiment. All participants achieved 90% or greater accuracy on this portion of the screening procedures.

Production Task

Equipment. Following the perceptual identification screening task, the production task was administered. A laptop computer (IBM ThinkPad R40) and Speech Measures were used to elicit and record speech stimuli for the Production task. The production task was completed in the sound treated IAC booth. Speech stimuli were recorded with a unidirectional condenser microphone (AKG C520) positioned approximately 1 inch from the participants' mouth, digitized with an Edirol UA-25EX external sound card at a sampling rate of 48 kHz and quantization of 24 bits, and low pass filtered at 22 kHz. The sampling rate of 48 kHz was chosen in order to capture the high frequency energy contained in the child participants' fricative productions, and to eliminate the possibility of any

artificial cutoff of spectral content in the participants' productions. Stimuli were recorded and saved directly to the laptop.

Procedures. The speech stimuli utilized for the production task consisted of the words "see," "she,", "sue," and "shoe" ([si], [ʃi], [su], and [ʃu]). The fricative + [i] or fricative + [u] represented two different coarticulatory configurations, as the resonating frequencies of the fricatives [s] and [ʃ] could be differently affected by the lip rounding required for articulation of [u], thus adding to the external validity of the production findings. A pseudo-random order of production tokens was created in order to eliminate the same token being presented twice in a row to any participant. The order was the same for all participants. The order of the words is listed in Appendix 4. A participant, wearing the head-mounted microphone, sat at a table next to the examiner. A participant was instructed that the investigator would present the stimulus word, and he/she would respond by saying the token in the carrier phrase, "Say … again." A carrier phrase was used in order to have the participant say the word in a more natural speaking context, rather than simply repeating the word after the examiner. This carrier phrase was written by the examiner on an 8 x 11 inch sheet of paper to serve as a guide for the participants if needed. The specific instructions given to participants for the carrier phrase and token word included "say it nice and clear as if you were telling someone to "say 'the word' again." No specific instructions were given regarding rate of speech for the participants to use, and no specific model of production to indicate stress or emphasis of any word within the carrier phrase was provided for the participants. Twenty-five tokens of each production were recorded, for a total of 100 productions per

participant. The principal investigator, who has a Northeastern American dialect, produced the target words for all participants. All participants completed 4 – 8 practice trials in order to become familiar with the procedure. If a participant produced the wrong word in the sentence, or made adjustments in production during the carrier phrase (i.e. "say [sʃu] again") then the token was re-recorded at the end of the task. Although all efforts were made to have all participants say the words in the specified order, some participants occasionally had to repeat the same word in the carrier phrase several times in a row at the end of the task, depending on their individual errors during the production task, due to this procedure. The production task lasted approximately 5 – 10 minutes, depending on the speaking rate of the individual participant and the number of tokens that were required to be re-recorded.

   Acoustic analysis. Acoustic analysis of the production tokens for each participant was completed after the entire study protocol was finished. Several measurements were computed. Each participant's first spectral moment time history for each fricative production was calculated using the "moments" command in the CSpeech software (Milenkovic & Read, 1997). Analyses were completed using a 20-ms Hamming window with a 10-ms step. The sibilant portion of the utterance was inspected visually using wide-band spectrographic analysis and waveform analysis. The onset and offset of the voiceless fricative segment was determined based on continuous aperiodic energy in the wide-band spectrographic analysis as well as onset and offset of periodicity in the surrounding vowel segments. The cursor was placed so that any region of periodicity was avoided. The placement of the cursors for analysis of the FSM is

shown in Figure 1. The mean and standard deviation of the first spectral moment of each token ("see, she, sue, shoe") was measured across the entire fricative. From this measurement the difference between the mean first spectral moment for [s] and [ʃ] was calculated. In addition, the first spectral moment was determined at fricative midpoint (defined as the middle analysis window or an average of the middle 20 ms, if the marked segment included an even number of analysis windows) and averaged across speaker. The range of the first spectral moments for [s] and [ʃ] was recorded for each speaker at fricative midpoint. The range was also qualitatively evaluated using visual inspection of the time history plots. The intra-speaker variability in production of [s] and [ʃ] for each speaker was estimated using the coefficient of variation (COV; ratio of standard deviation to the mean). This measure was used rather than standard deviation because it is a dimensionless number, can be used to compare data with different means, and provides a normalized measure of dispersion compared to standard deviation, which must always be evaluated in the context of the mean of a data set. An estimation of fricative duration was derived from the FSM analysis. The number of windows analyzed, multiplied by 10, represented fricative duration to the nearest 10 ms. The fricative duration was averaged across each token ([si], [ʃi], [su], and [ʃu]) for each speaker.

Figure 1. Waveform and spectrographic analysis of the carrier phrase "Say 'see' again" from a participant in the TD-7 group. Cursors are placed at the fricative segment in order to avoid any area of periodicity from the previous and following vowel segment.
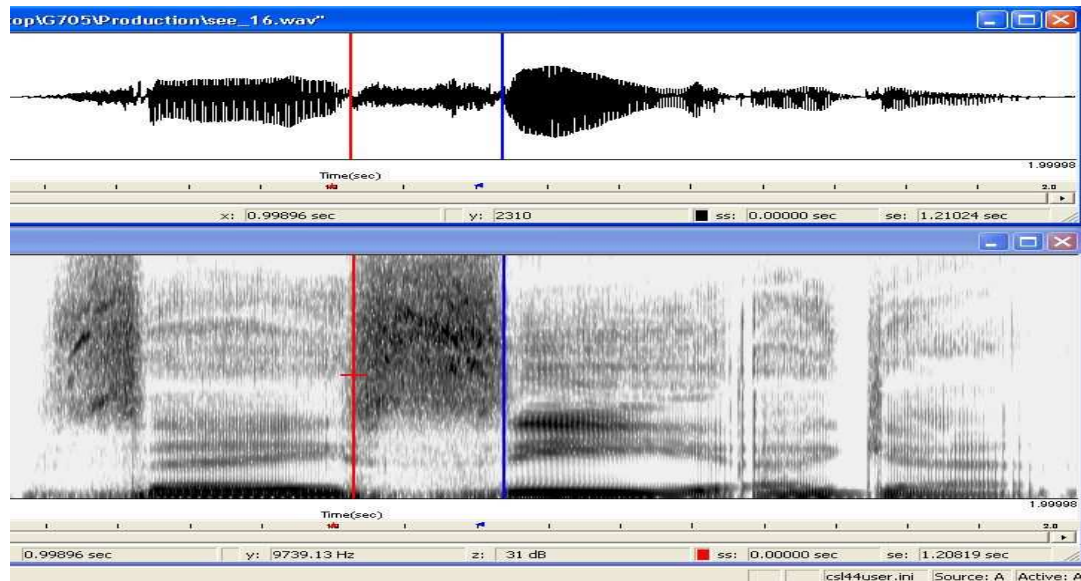
Of the 2500 productions collected, only 14 utterances across 8 participants were unusable for data analysis. The unusable utterances were caused by recording errors within the SpeechMeasures software. Therefore, a total of 2486 utterances were analyzed.

Data reliability. Inter-observer agreement was calculated through re-analysis of the first spectral moment and estimated fricative duration. An undergraduate student from the University of North Carolina at Chapel Hill Division of Speech and Hearing Sciences, familiar with both waveform and spectrographic analyses, was trained on the procedures for measuring first spectral moments and fricative duration. The student was instructed on the keystroke commands used for the Cspeech program, the procedures for deriving the waveform and spectrogram for speech stimuli, and the conventions for placing cursors so as to avoid regions of periodicity on either side of the fricative segment. The student then computed 20 FSM for a participant jointly with the

42

principal investigator. These training stimuli were not included in the final reliability analysis. Twenty percent of the utterances from each speaker in order to ensure equal distribution across speakers and fricative + vowel targets. Five random targets each from each speaker's [si], [ʃi], [su], and [ʃu] productions were selected for re-analysis. For the first spectral moment measures, the Pearson product moment correlation was .99, and point-to-point agreement was 94% when agreement was defined as measures within .15 kHz. For the duration measures, the Pearson product moment correlation was .92, and point-to-point agreement was 88% when agreement was defined as measures within 20 ms.

Perceptual Identification Task

    Equipment. The equipment for the Perceptual Identification Task was the same as the equipment used for the Perceptual Identification Screening. This included the laptop computer (IBM ThinkPad R40) and the perceptual Identification Task module of SpeechMeasures (Haley, 2008) software program . The perceptual identification task was completed in the sound treated IAC booth. A participant wore headphones (Sennheiser, HD 250 Linear II) in order to block out distractions from the environment and to allow them to focus on the task at hand.

    Stimuli. Speech stimuli were created for both the perceptual identification and perceptual discrimination tasks. The speech stimuli for the identification and discrimination tasks consisted of synthesized fricative noise ([s] and [ʃ]) and vowel portions ([i] and [u]). These two vowels were chosen to represent both a front and back vowel context, and corresponded directly to the tokens used in the production task. Two continua, each consisting of seven stimuli, were

synthesized using the cascade-parallel synthesis mode of the Sensimetrics High-

Level Parameter Speech Synthesis System for Windows (HLsyn; Sensimetrics

Corporation, 2004).

Parameters used to synthesize the fricative portion of the stimuli were

modeled after the [sit] – [ʃit] continuum used by Rvachew & Jamieson (1989), in

their study on the relationship between perception and production of fricatives in

children with both normal and disordered articulation. The continuum from [s]

to [ʃ] was created using formant amplitude values that were an approximate

linear interpolation of the values of the endpoints [Step 1 ([s]) to Step 7 ([ʃ])] for

5 intermediate stimuli, creating a 7-step continuum. Several outside sources,

including the director of another speech science research laboratory and doctoral

students from this laboratory, were consulted to ensure that the endpoints of the

stimuli were subjectively good representations of the target syllables. The

director and students listened to several alternatives and chose stimuli that were

best representative of the endpoints of the [s] - [ʃ] continuum. The values for the

sibilant synthesis parameters are displayed in Table 2.

Table 2. Parameters for sibilant synthesis. Formant (F) and bandwidth (B) parameters
are expressed in Hz; amplitude of frication formants (A) are expressed in dB.
Parameters that were held constant during the sibilant portion include the following: F1
= 310Hz; B1 = 200 Hz; F4 = 3300 Hz; B4 = 250 Hz; F5 = 3850 Hz; B5 = 200 Hz; F6 =
4900 Hz.

Step	F2	F3	B2	B3	B6	A3	A4	A5	A6
Step 1: [s]	1465	2567	87	217	2000	9	8	8	51
Step 2	1528	2598	89	229	1782	17	15	15	50
Step 3	1590	2629	91	242	1587	25	21	21	49
Step 4	1653	2660	93	255	1414	33	28	28	48
Step 5	1715	2690	95	269	1260	41	35	35	47
Step 6	1778	2721	98	284	1122	49	41	41	46
Step 7: [ʃ]	1840	2752	100	300	1000	57	48	47	45

The duration of the fricative segments was 210 ms. The intensity pattern during the fricative is shown in Table 3. The fricative intensity increased 20 dB for 170 ms (0 – 170 ms), remained constant for 20 ms (170 – 190 ms), and decreased 5 dB for 20 ms (190 – 210 ms). These temporal parameters were based on the work of Nittrouer & Studdert – Kennedy (1987), who created synthetic fricative stimuli to study fricative perception in children. The amplitude of frication and amplitude of voicing values were based on the work of Hedrick & Ohde (1993), who provided specific values for fricative amplitude using a synthetic continuum of [s] and [ʃ] and the vowels [i] and [u]. The amplitude of voicing represented the transition from the voiceless sibilant fricative to the vowel portion of the stimuli.

Table 3: Intensity Pattern of Frication and Voicing during Fricative + Vowel Stimuli. AF = amplitude of frication; AV = amplitude of voicing. AF and AV are expressed in dB.

Time	AF	AV
0	35	0
170	55	0
190	55	0
210	50	49
220	0	70

The vowel portions of the stimuli were synthesized versions of [i] and [u], creating stimuli representing [si], [ʃi], [su], and [ʃu] ("see, she, sue, shoe"). The duration of the vowel was 350 ms, based on established values in Nittrouer & Studdert-Kennedy (1987), who investigated perception of fricative+vowel syllables in children 3 – 7 years of age. The vowel parameters were modeled after the work of Whalen (1981), Nittrouer & Studdert-Kennedy (1987), and Hedrick & Ohde (1993). All three studies provided values for synthetic fricative stimuli

using the vowels [i] and [u]. F1 for both the [i] and [u] vowels remained at 310Hz

(value of F1 during frication). The F2, F3, and duration of transition from

frication to the steady state portion of the vowel were based on established values

for the endpoints of [s] - [ʃ] continua using the vowels [i] and [u] published by

Nittrouer & Studdert-Kennedy (1987). The endpoints for both the [i] and [u]

vowel, in both the [s] and [ʃ] contexts were input into the HLSyn software, and

the values for the 5 intermediate stimuli were created using a linear interpolation

of the endpoint values. The vowel parameters are shown in Tables 4 and 5 for the

continua from [si] - [ʃi] and [su] - [ʃu], respectively. Figure schematics of the

endpoint and middle stimuli for each of the continua are displayed in Figures 2

and 3. Actual spectrograms for all stimuli are provided in Appendices 5 and 6.

Table 4: Vowel synthesis parameters for [i]. Formant onset and steady state are expressed in Hz, and transition duration is expressed in ms.

	Onset of F2	F2 Transition Duration	Steady State F2	F3 Onset	F3 Transition Duration	Steady State F3
Step 1 ([s])	1870	80	2050	2430	140	2600
Step 2	1903	67	2053	2447	127	2597
Step 3	1937	53	2057	2463	113	2593
Step 4	1970	40	2060	2480	100	2590
Step 5	2003	27	2063	2497	87	2587
Step 6	2037	13	2067	2513	73	2583
Step 7 ([ʃ])	2070	0	2070	2530	60	2580

Table 5: Vowel synthesis parameters for [u]. Formant onset and steady state are expressed in Hz, and transition duration is expressed in ms.

	F2 Onset	F2 Transition Duration	Steady State F2	F3 Onset	F3 Transition Duration	Steady State F3
Step 1 ([s])	1620	300	1000	2340	170	2130
Step 2	1673	300	1000	2342	152	2133
Step 3	1727	300	1000	2343	133	2137
Step 4	1780	300	1000	2345	115	2140
Step 5	1833	300	1000	2347	97	2143
Step 6	1887	300	1000	2348	78	2147
Step 7 ([ʃ])	1940	300	1000	2350	60	2150

Figure 2. Schematic representing Step 1 ([si]), Step 4, and Step 7 ([ʃi]) of the [si] - [ʃi] continuum, respectively. The y-axis represents frequency in Hz, and the x-axis represents time in ms. Vowel formants are labeled accordingly.

Figure 2, continued.

Figure 3. Schematics representing Step 1 ([su]), Step 4, and Step 7 ([ʃu]) of the [su] - [ʃu] continuum, respectively. The y-axis represents frequency in Hz, and the x-axis represents time in ms. Vowel formants are labeled accordingly.
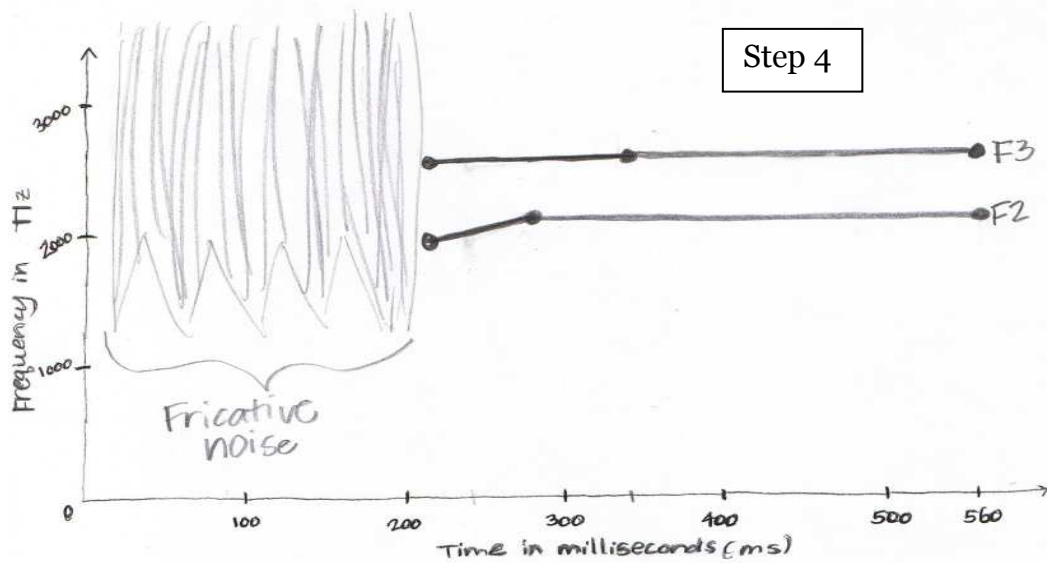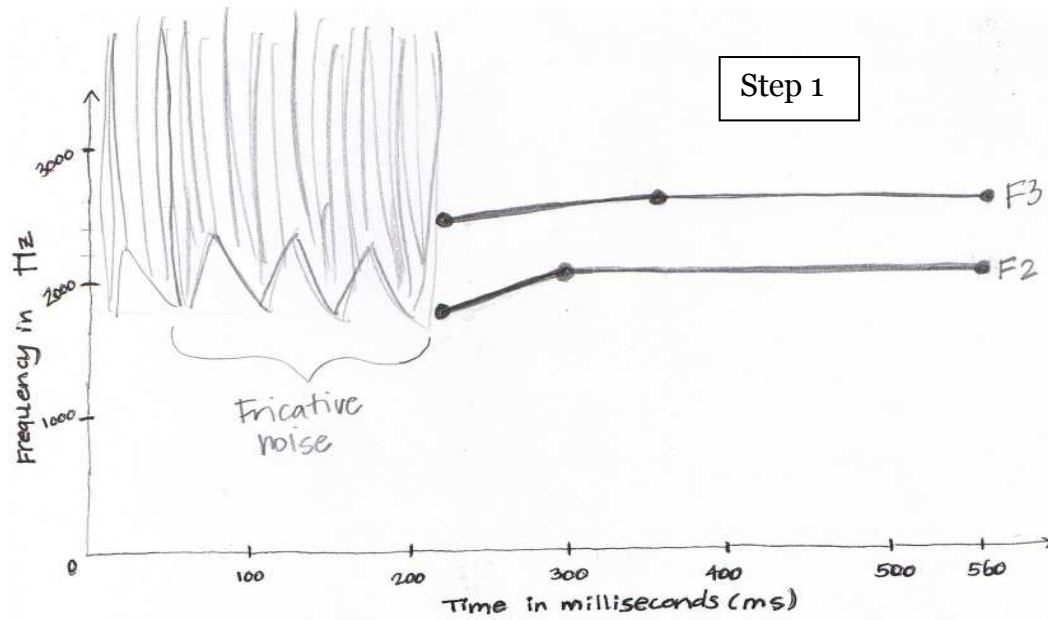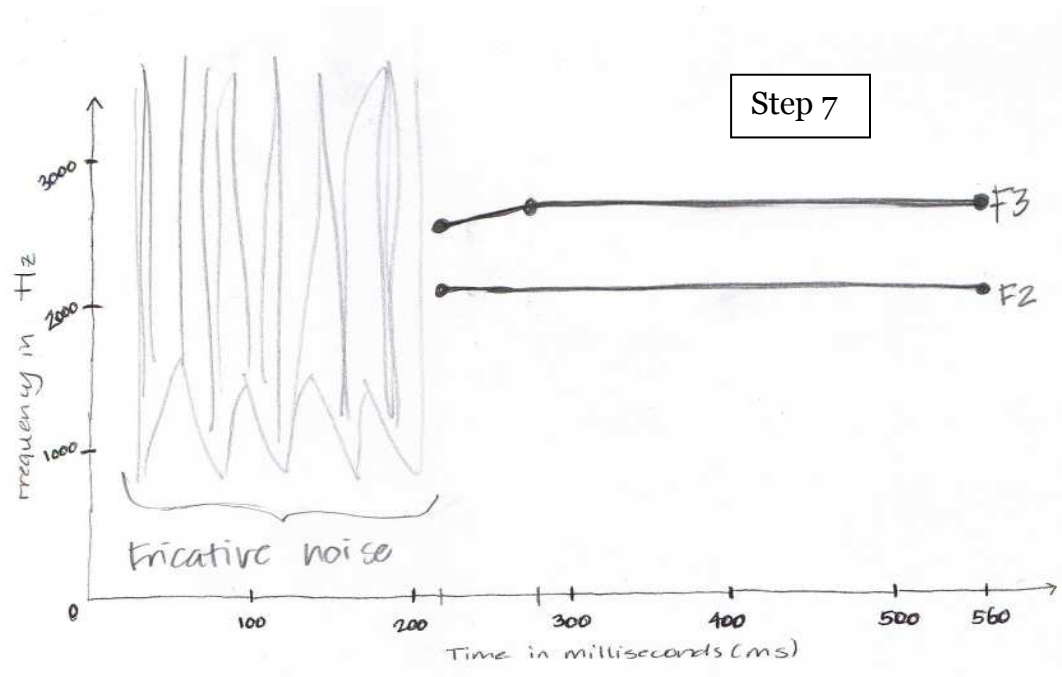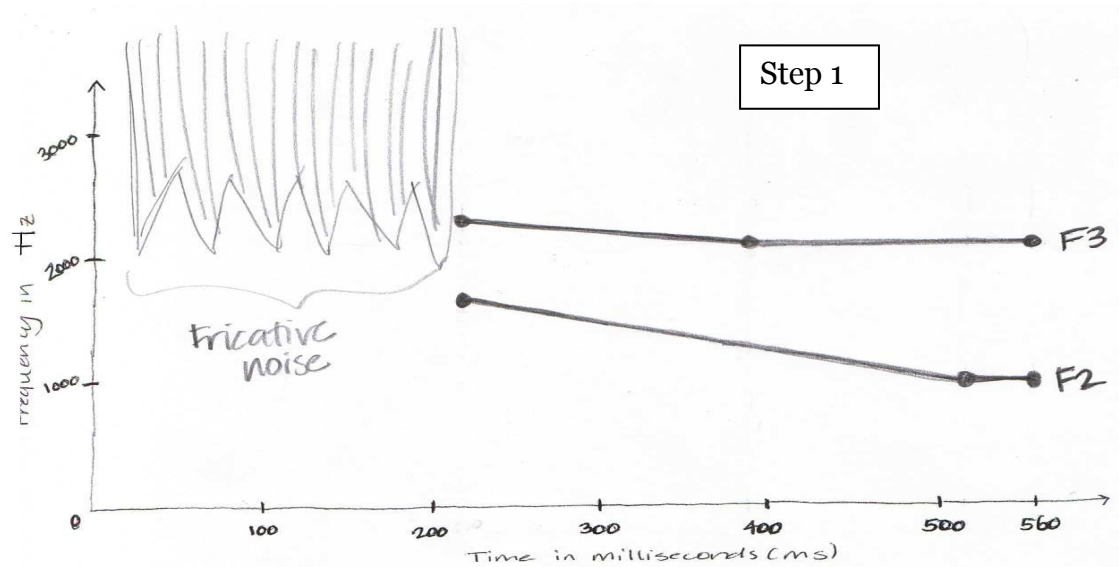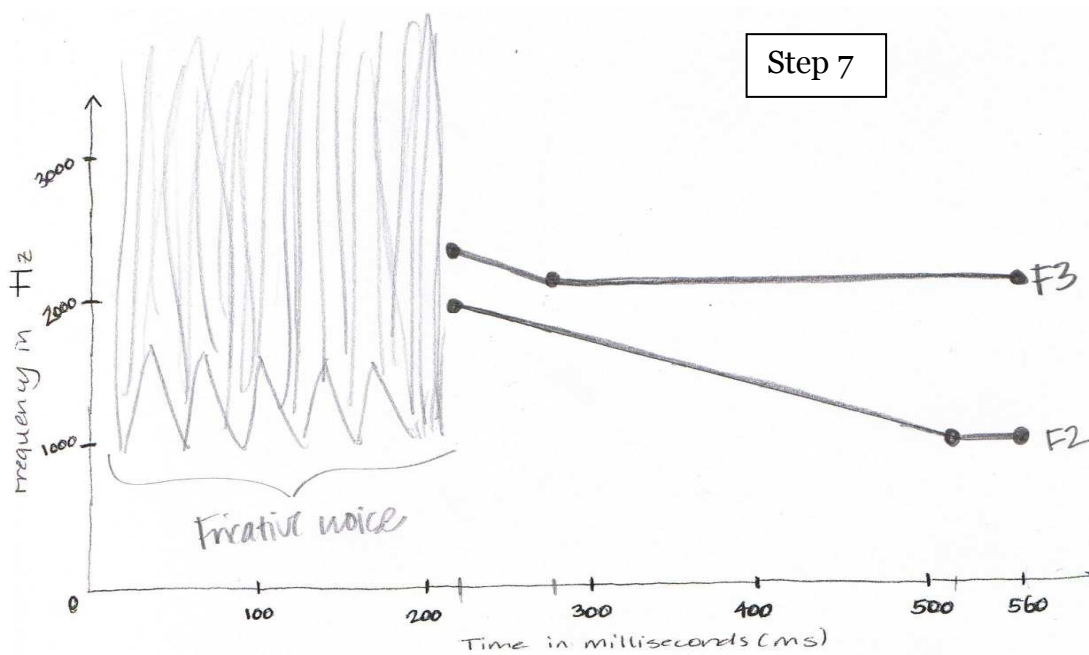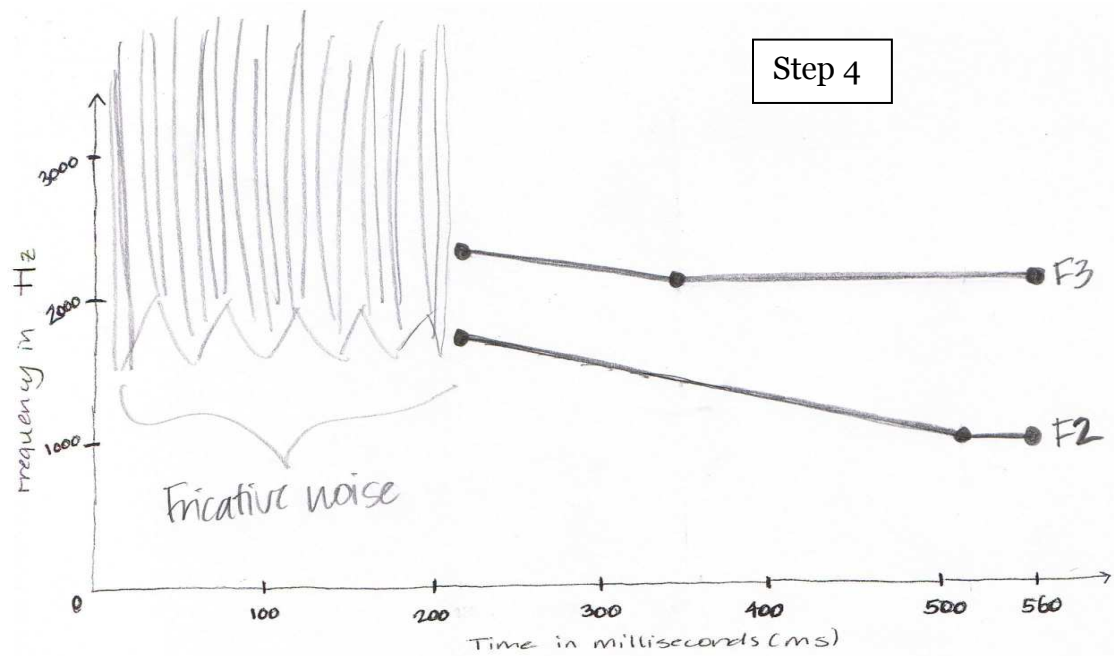

48

Figure 3, continued.

The bandwidths for F1, F2, and F3 were modeled after Hedrick & Ohde

(1993) and remained consistent over the duration of the vowel. The values for

49

the bandwidths were as follows: B1 = 60 Hz, B2 = 80 Hz, B3 = 350 Hz for [i], and B1 = 65 Hz, B2 = 110 Hz, B3 = 140 Hz for [u]. The amplitude of voicing for the vowel decreased linearly from 70 dB to 0 dB over the last 75 ms of the vowel portion.

Stimuli calibration. Peak calibration of the stimuli was conducted prior to initiation of data collection using the SpeechMeasures software program, external sound card, and headphones. Stimuli were calibrated at 80 dB SPL using a sound level meter calibration system (Larson Davis System 824). When stimuli were played through the setup configuration and headphones, the peak output of all tokens was verified to be 80 dB SPL (+/- 1 dB). All tokens were verified binaurally via this method. These settings were carefully replicated for all sessions of data collection. In addition, confirmation of stimuli playing correctly through the program, external sound card, and into the headphones, was conducted before each experimental session.

Procedures. The experimental identification task was administered after the production task for all participants. Identification testing is designed to test the ability to judge phonemic categories. In studies investigating perceptual abilities of children, identification testing has been argued to be a better test of a child's phonemic categorization abilities compared to discrimination testing (Locke, 1980; Monnin et al., 1974). The format of the identification task was similar to the screening task, but rather than the endpoints of the continuum being presented to participants, all 7 steps of the continuum were presented. Ten repetitions of the seven stimuli from each continuum were presented in a pre-ordered, pseudo-randomized manner (using a pseudo-random number

generator, www.randomizer.org), arranged so that no step of the continuum was presented adjacent to one another.   The same order of stimulus presentation was utilized for all participants, and is listed in Appendix 7. Participants were required to identify, for both continua, whether the stimuli they heard was either "see"/ "she" or "sue" / "shoe" by pointing to the picture and orthographic representation of the words. Responses were recorded by the investigator, as described in the identification screening task. Participants completed the identification task for the [si] - [ʃi] and [su] - [ʃu] continua, one at a time, representing responses to 140 tokens in total.   The perceptual identification task took approximately 6 – 10 minutes to complete for each continuum.

Data analysis. The dependent variable for the identification task was the percent of [s] responses for each step of the continuum. The Identification Task module of the SpeechMeasures (Haley, 2008) program generated an output file that reported each participant's response of either "see/she" or "sue/shoe" to each step of the respective continuum. The number of "see" or "sue" responses was calculated for each step, and expressed as a percentage.

Accuracy of data entry. To examine whether data were entered correctly on spreadsheets, 20% of the identification task output files were re-tabulated by a second observer/rater. Re-tabulation of data was 100% consistent with computations done initially.

Perceptual Discrimination Task

Equipment. The equipment for the perceptual discrimination task was the same for the perceptual identification task. This included the laptop computer (IBM ThinkPad R40) and the perceptual discrimination task module of a pilot

version of the SpeechMeasures (Haley, 2008) software program . The perceptual discrimination task was completed in the sound treated IAC booth. A participant wore headphones (Sennheiser, HD 250 Linear II) in order to block out distractions from the environment and to allow them to focus on the task.

Stimuli. The stimuli and calibration procedures used in the perceptual discrimination task were the same as described for the perceptual identification task.

Procedures. The administration of the perceptual discrimination task followed the perceptual identification task. Discrimination testing was used in order to investigate the participants' sensitivity to changes in the fricative spectrum and vowel transitions. In investigating perception of fricatives in children, discrimination testing may be useful in determining the locus of breakdown when a child demonstrates the inability to identify phonemic boundaries (Hoffman et al., 1985). The format of the discrimination task was modeled after Perkell et al. (2004b), by using an ABX paradigm (Liberman, Harris, Kinney, & Lane, 1951). Two and 3-step comparisons were utilized. There were five 2-step comparisons (1 vs. 3, 2 vs. 4, 3 vs. 5, 4 vs. 6, 5 vs. 7) and four 3-step comparisons (1 vs. 4, 2, vs. 5, 3 vs. 6, 4 vs. 7) for each sibilant + vowel contrast. Each of these comparisons was presented in 4 sequences (e.g., using stimuli 1 and 3: 131, 133, 311, 313). Presentation of the comparisons was randomly generated by the SpeechMeasures program.

A visual representation from the perceptual discrimination task is provided in Appendix 8. A participant would hear the first stimulus (using the above example, step 1) while the box with "1" flashed on the screen. After a one

52

second delay, the second stimulus (step 3) would play while the box with "2" flashed on the screen. Finally, after another one second delay, the third stimulus (step 1) would play while the box with "?" flashed on the screen. After each sequence was presented, the participant stated either "1" or "2" to indicate whether the first (1) or second (2) stimulus was the same as the third stimulus (?). The principal investigator recorded each response by clicking the number on the screen with the mouse that corresponded with the participant's response. Each participant completed two sessions of the discrimination task for each continuum (four sessions total), representing 144 tokens (36 tokens in each session; 20 tokens for 2-step discrimination and 16 tokens for 3-step discrimination). Each session of the discrimination task lasted approximately 5 minutes.

In previous research, using relatively similar stimuli and procedures, (Perkell, Matthies et al., 2004), adult participants performed at chance levels for a 1-step discrimination task and demonstrated performance at ceiling (100% accuracy) for a 3-step discrimination task. For this reason, the discrimination task for child participants included 2-step and 3-step discrimination, assuming that child participants would not perform better than adult participants on this type of task.

Data analysis. The dependent variables for the perceptual discrimination task included proportion correct for both the 2-step and 3-step discrimination trials. The discrimination task module of the SpeechMeasures software generated an output file that recorded responses as correct or incorrect for each comparison in the discrimination task. The proportion correct was calculated for

each comparison (i.e. 1 vs. 3...5 vs. 7 for 2-step discrimination, 1 vs. 4...4 vs. 7 for 3-step discrimination) and for each continuum. An overall proportion correct was calculated for 2-step and 3-step discrimination trials for each continuum. The proportion correct from the two sessions of the [si] - [ʃi] and [su] - [ʃu] continua were averaged across continua to obtain the overall proportion correct for both 2-step and 3-step discrimination trials.

Accuracy of data entry. To determine whether the data were entered correctly on spreadsheets, 20% of the discrimination output files were re-tabulated by a second observer/rater. Re-tabulation of data was 100% consistent with computations made initially.

IRB Approval

The entire experimental protocol was approved by the University of North Carolina at Chapel Hill Biomedical Institutional Review Board. Participants were recruited via UNC broadcast email and via word-of-mouth. Parents provided consent for their child's participation, and all participants provided assent. Participants were paid for their participation.

Planned Analyses

Both quantitative and qualitative analyses were utilized for data analysis purposes. The SPSS software ("SPSS 14 for Windows"; SPSS, Inc., 2005) was used to conduct statistical analyses. For all statistical tests, an a priori significance level of 0.05 was adopted. Planned analyses for each of the five research questions, here expressed as hypotheses, are detailed in the following section.

Hypothesis One

     Typically developing children age 7 and 11 years of age will demonstrate a reliable distinction between [s] and [ʃ], which would reflect the speech production pattern found in adults. Visual inspection of time history plots was used to examine spectral overlap between [s] and [ʃ] production qualitatively and to evaluate inter-speaker variability. This qualitative analysis approached allowed for visual examination of the entire fricative segment. In addition, the values of the FSM at midpoint were examined for overlap to determine if categorical distinction existed between the two phonemes for each speaker.

Hypothesis Two

     The 11-year-old children will produce the fricatives [s] and [ʃ] with greater magnitude of distinction compared to the 7-year-old children, approaching adult-like productions. Dependent variables for testing developmental differences in production for each speaker for both [s] and [ʃ] included the average FSM at midpoint in order to estimate the magnitude of distinction (i.e., difference in FSM between [s] and [ʃ]) between the fricatives. In addition, fricative duration and intra-speaker variation, estimated using the coefficient of variation, were also measures of interest that were examined for the two age groups. Independent sample student's t-tests were used to determine whether the two age groups differed on these measures.

Hypothesis Three

     The older children will demonstrate steeper identification function curves and higher levels of accuracy in fricative discrimination than the younger children. The dependent variables for perception measures included percent of

[s] responses on the 7 steps of the continuum for the identification function, percent accuracy on 2- and 3-step discrimination trials, and the coefficient of variation of discrimination responses. In addition, graphs of the identification functions for each group were created by overlaying the curves of individual participant's percent [s] responses for the TD-7 and TD-11 group. These graphs were used to estimate the steepness of the slope of the identification function qualitatively, and to compare the two age groups. Student's independent sample t-tests were used to evaluate if differences existed between the two age groups on the discrimination task.

Hypothesis Four

Typically developing children who demonstrate greater distinction between [s] and [ʃ] and less variability in production will have higher discrimination abilities than children with a decreased magnitude of distinction in production. The plan for addressing this hypothesis was to use Pearson product moment correlations to determine if a linear relationship existed between production and perception measures. The production measures included the magnitude of distinction at midpoint (i.e. the difference between FSM at midpoint for [s] and [ʃ]) and accuracy on 2-step or 3-step discrimination tasks.

Hypothesis Five

Children with repaired cleft palate, who demonstrate distorted fricative production judged by the examiner, will show spectral overlap measured by spectral moment analysis. We may further predict that these same children will also show shallower identification curves and less accurate discrimination of [s]

versus [/] on the perception tasks. Perceptual and production measures from the children with repaired cleft palate will be inspected qualitatively on an individual speaker basis and will be discussed in conjunction with the findings from the typically developing child group.

CHAPTER 4

Results

Fricative Production

It was hypothesized that typically developing children, age 7 and 11 years,

produce a reliable distinction between the voiceless sibilant fricatives [s] and [ʃ]

measured by spectral moment analysis. It was also hypothesized that the

magnitude of distinction would be larger in the older children, representing more

adult-like production.

Distinction in Fricative Production for Individual Speakers

The mean FSM at midpoint and range of FSM at midpoint, the difference

in FSM between [s] and [ʃ] at midpoint, and the percent of non-overlapping

distributions for each participant are reported in Table 6. In order to evaluate if

the participants in the TD-7 and TD-11 group produced a reliable distinction

between [s] and [ʃ], the values for the range of FSM for [s] and [ʃ] at midpoint

were examined for overlap. If there was overlap, the lowest FSM value in the

data range for [s], and then the highest FSM value in the data range for [ʃ] were

temporarily removed, alternately, from the entire distribution until a non-

overlapping distribution was achieved. The number of temporarily removed data

points was counted, and then the proportion of the data points that remained

compared to the total number of data points was calculated to reach a percentage

of non-overlapping productions.

Table 6. FSM (expressed in kHz) at midpoint for the TD-7 and TD-11 groups. Means and ranges are reported for individual speakers in the TD-7 group (G701 – B711) and TD-11 group (G1101 – B1110) in addition to the means and ranges for each group. Female speakers are represented by the initial "G," and male speakers with the initial "B." The magnitude difference, [s] - [ʃ], is the difference in mean values for the speaker or speaker group. The percentage of productions that were non-overlapping was estimated based on midpoint values.

Speaker	[ʃ] mean (SD)	[ʃ] Range	[s] mean (SD)	[s] Range	[s]-[ʃ]	% of non-overlapping productions
G701	8.9 (.57)	7.8 – 10.0	11.9 (.92)	9.9 – 15.1	3.0	99%
G702	6.5 (.46)	5.8 – 7.3	9.8 (1.00)	7.8 – 12.5	3.3	100%
G703	7.2 (.55)	5.9 – 8.5	10.0 (.62)	8.8 – 11.4	2.8	100%
G704	7.3 (.67)	5.7 – 9.0	11.5 (1.19)	9.2 – 14.0	4.2	100%
G705	7.3 (.56)	6.4 – 8.6	11.7 (.64)	10.5 – 13.0	4.4	100%
B706	6.9 (.57)	5.7 – 8.2	10.5 (.97)	8.5 – 12.5	3.7	100%
B708	6.5 (.53)	5.3 – 7.4	7.2 (.87)	5.1 – 9.1	0.7	72%
B709	6.2 (.57)	4.4 – 7.4	9.5 (.67)	8.1 – 10.9	3.3	100%
B710	7.8 (.64)	6.4 – 9.3	9.3 (.62)	8.0 – 12.1	1.5	91%
B711	7.5 (1.24)	5.0 – 11.2	10.6 (1.15)	7.8 – 14.6	3.1	95%
TD-7 group	*7.2 (.64)*	*4.4 – 11.2*	*10.2 (.87)*	*5.1 – 15.1*	*3.0*	
G1101	7.3 (.36)	6.5 - 9.2	10.6 (.63)	9.1 - 12.0	3.4	99%
G1102	5.9 (.75)	4.7 - 8.1	8.8 (.78)	7.0 - 11.6	2.9	96%
G1103	6.3 (.31)	5.6 - 7.1	9.6 (.72)	8.3 - 11.1	3.3	100%
G1104	6.8 (.83)	4.5 - 8.4	10.7 (.52)	9.5 – 12.0	4.0	100%
G1105	8.0 (.78)	6.3 - 10.1	9.7 (.51)	8.8 - 10.9	1.7	94%
B1106	6.6 (.58)	5.4 - 8.4	9.1 (.80)	6.6 - 10.7	2.5	95%
B1107	6.9 (.64)	5.9 - 9.3	8.6 (.66)	7.4 - 11.1	1.5	90%
B1108	6.8 (.44)	6.6 - 7.7	8.7 (.31)	8.0 - 9.5	1.9	100%
B1109	6.3 (.53)	5.0 - 7.3	9.0 (.60)	7.5 - 10.5	2.7	100%
B1110	7.5 (.58)	6.3 - 8.9	11.3 (.71)	9.8 - 12.9	3.8	100%
TD-11 group	*6.8 (.58)*	*4.5 – 10.1*	*9.6 (.62)*	*7.0 – 12.9*	*2.8*	

Within the TD – 7 group, 8/10 (80%) of the participants showed non-overlapping distributions in at least 95% of productions, and 9/10 (90%) of the participants showed non-overlapping distributions in at least 90% of productions. In the TD – 11 group, 8/10 (80%) of the participants showed non-overlapping distributions in at least 95% of productions, and 10/10 (100%) of the

participants showed non-overlapping distributions in at least 90% of productions.

Qualitative analysis of overlap in fricative production was completed by visually inspecting time history plots of the FSM for [s] and [ʃ] for each participant. Time history plots displaying the FSM for [s] and [ʃ] for each participant within each age group are shown in Figures 4 and 5. Within both groups, some degree of distinction was apparent in all speakers, with the exception of B708 (the speaker who demonstrated 28% of overlapping productions). Speakers B710, B711, G1105 and B1107 had [s] and [ʃ] distributions that appeared to possibly overlap; however, when overlapping FSM values at midpoint were temporarily removed, the range of [s] and [ʃ] frequencies were non-overlapping. Differences across participants in magnitude of distinction between the fricatives can be appreciated, such as when comparing speakers G703 and G705 (magnitude of distinction was 2.8 kHz and 4.4 kHz, respectively), both female speakers from the TD-7 group. Variability in dynamic patterns throughout the time course of the fricatives can be appreciated as well. By visual inspection alone, greater variability in [s] production compared to [ʃ] production can be seen in several speakers, including G702, G704, B706, and G1103. Dynamic patterns that differ between speakers include consistent, steady fricative spectral energy (for example, in speaker G703 and B1108) throughout the fricative segment versus speakers who displayed varied spectral energy from onset to offset of frication (for example, speakers G1104 and B1110). In summary, a wide variety of individual patterns in fricative production existed among participants in both age groups.

Figure 4. First spectral moment (FSM) time history plots for TD-7 participants' repeated productions of [s] and [ʃ]. G701 – G705 are female speakers, and B706 – B711 are male speakers. Time histories across the repeated utterances are superimposed on each other to display categorical distinction visually. The higher frequency clusters are [s] targets and the lower frequency clusters are [ʃ] targets. Zero ms on the time axis represents the first analysis window positioned 10 ms into the fricative segment. Please note the different time scale on the x-axis for participant B711 (800 ms) compared to the other participants (300 ms).
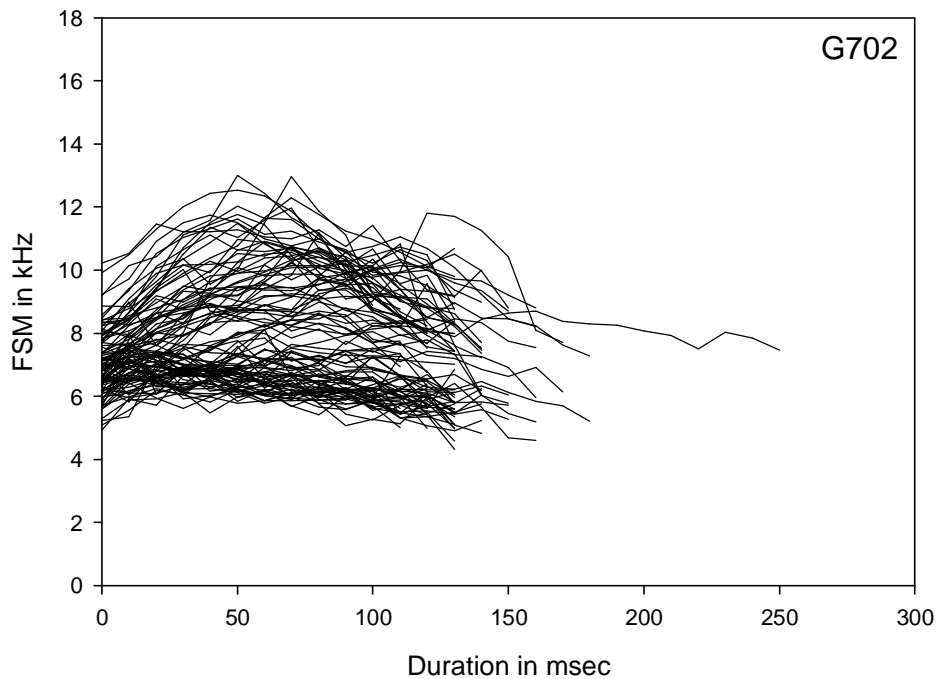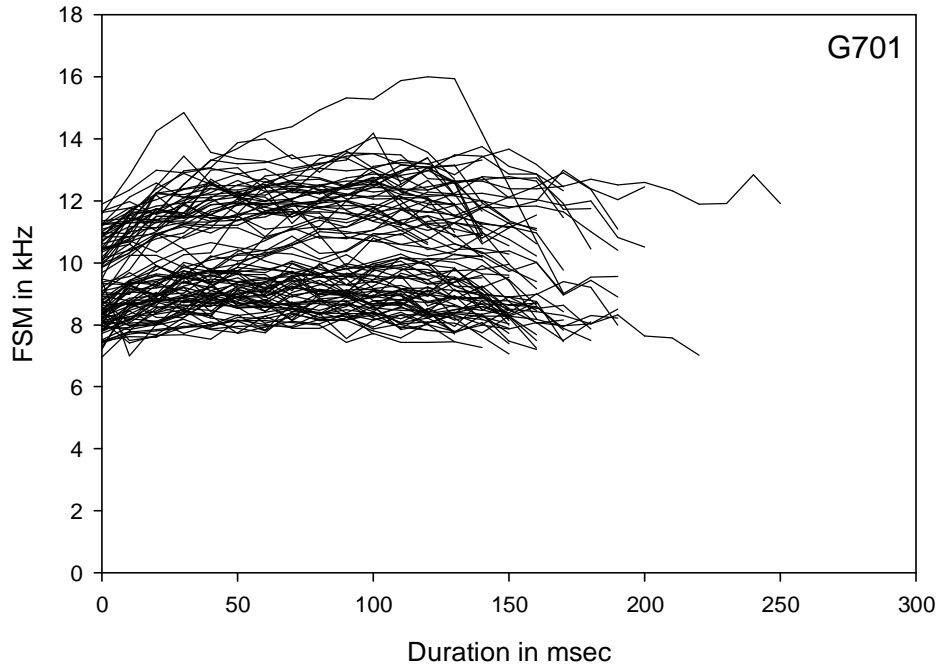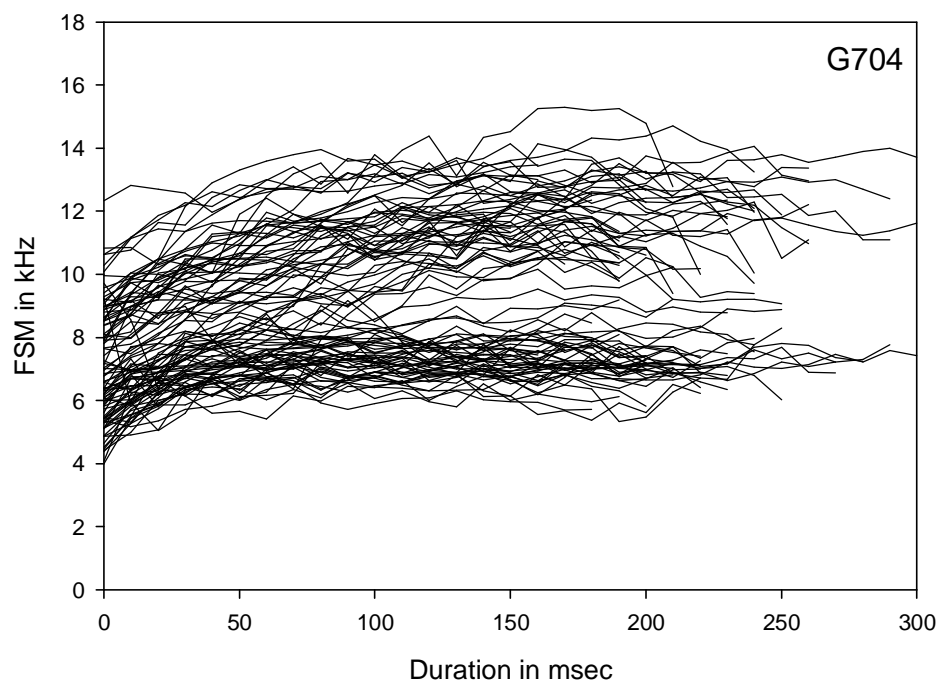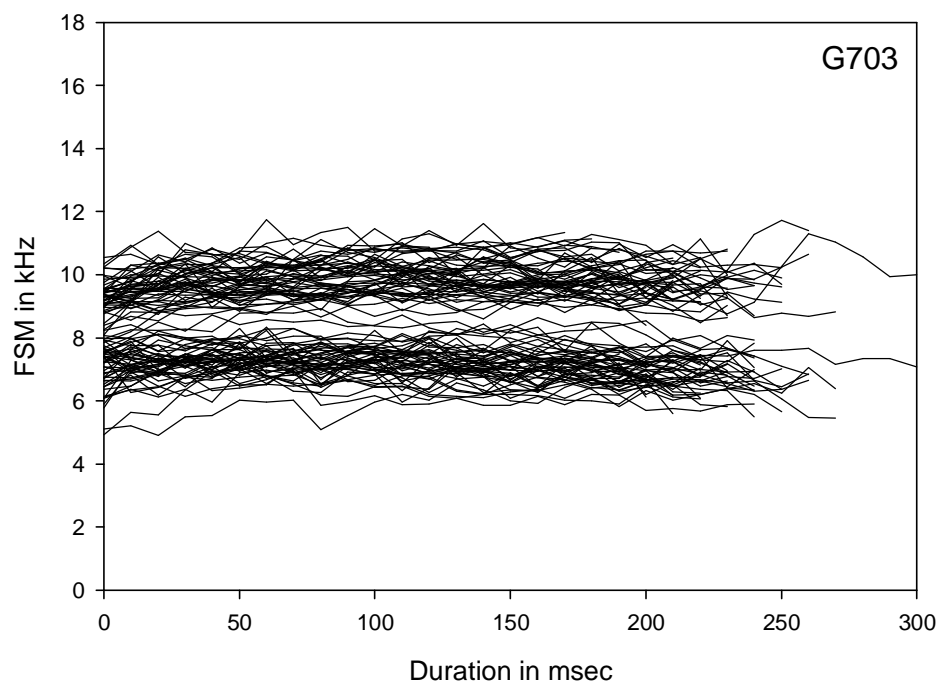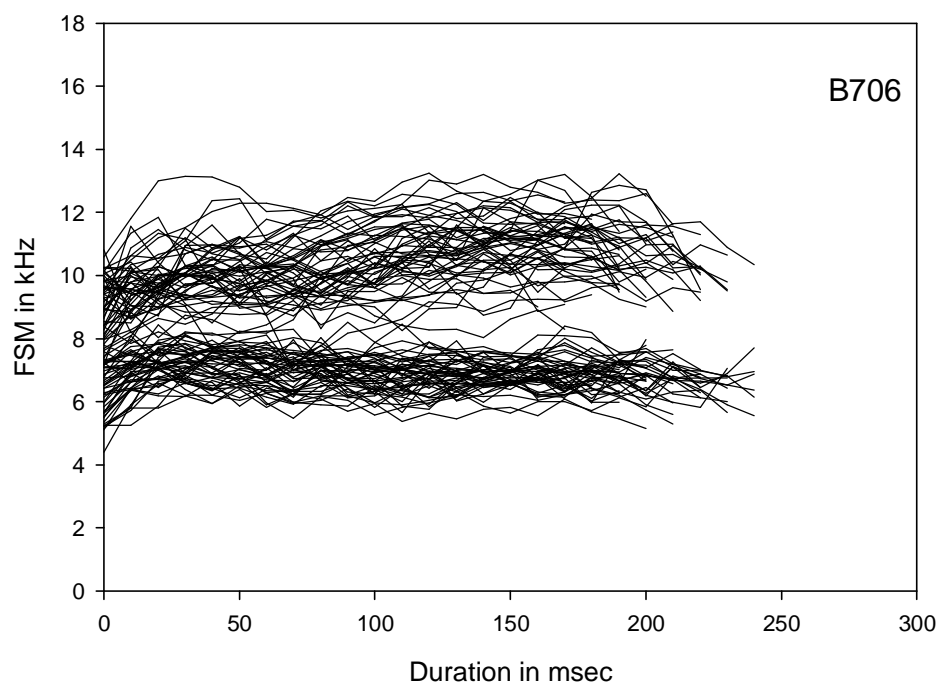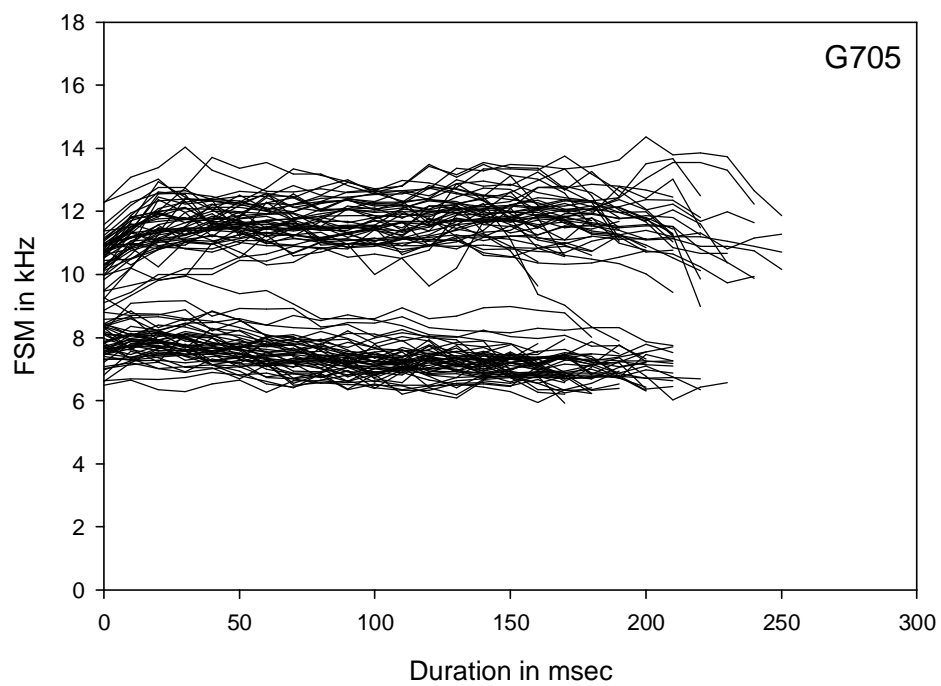
Figure 4, continued.
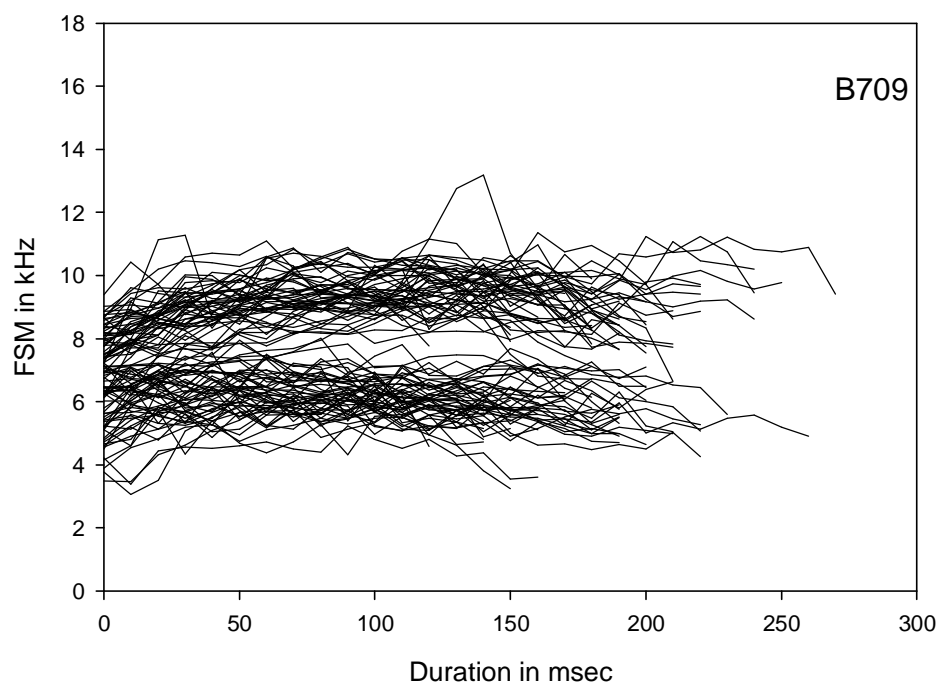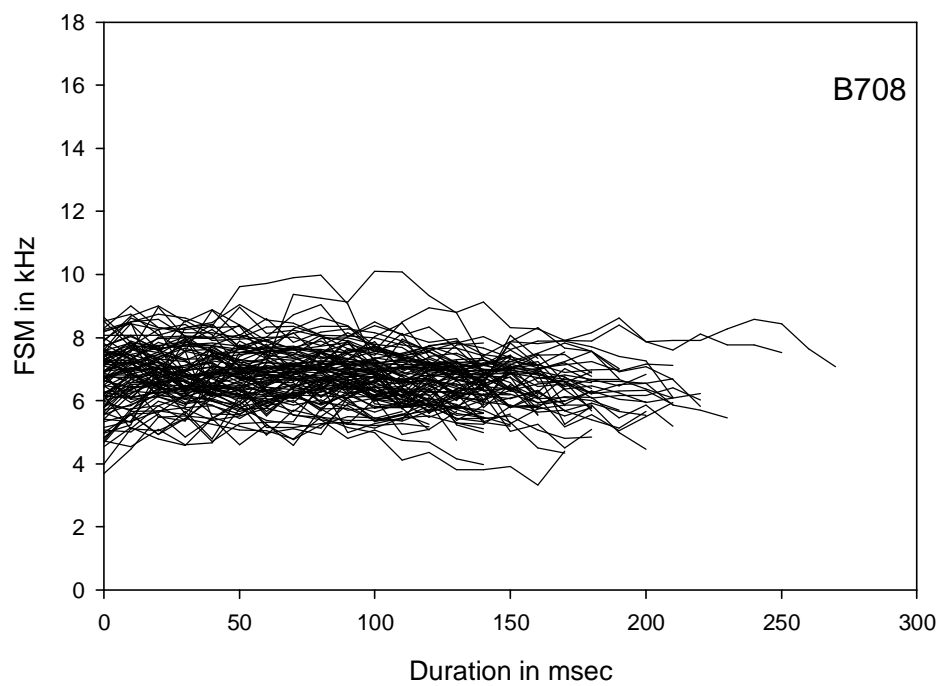
Figure 4, continued.

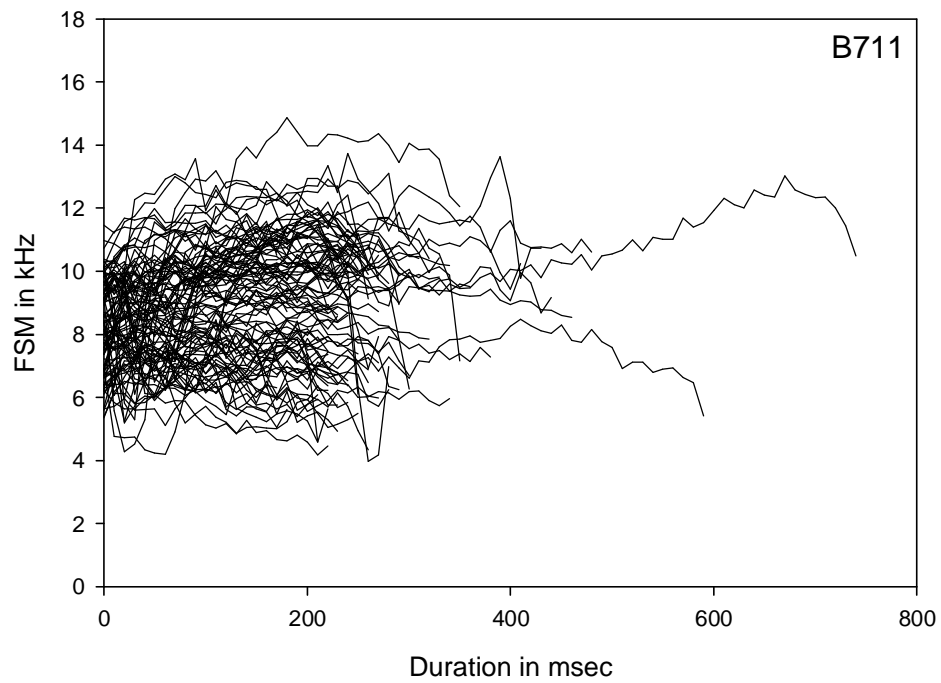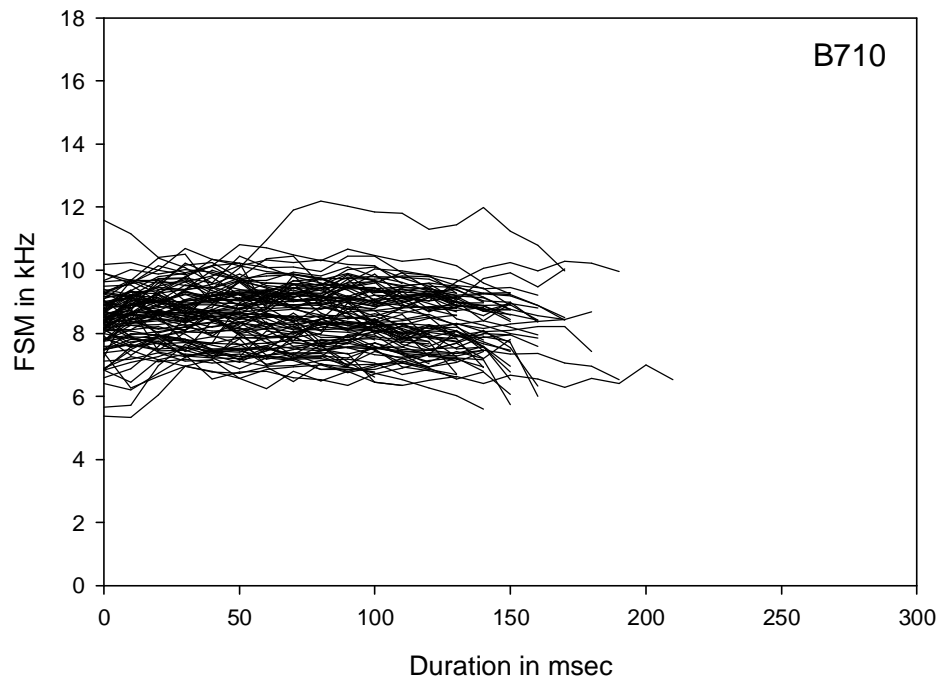Figure 4, continued.

Figure 4, continued.

Figure 5. First spectral moment (FSM) time history plots for TD-11 participants'
repeated productions of [s] and [ʃ]. G1101 – G1105 are female speakers, and B1106 –
B1110 are male speakers. Time histories across repeated utterances are superimposed on
each other to display categorical distinction visually. The higher frequency clusters are
[s] targets and the lower frequency clusters are [ʃ] targets. Zero ms on the time axis
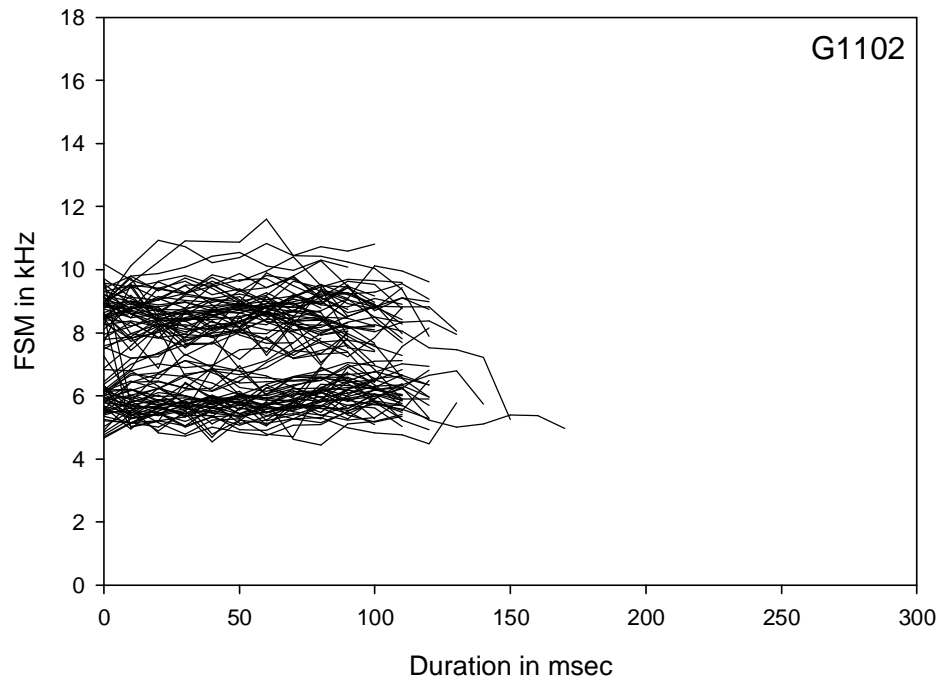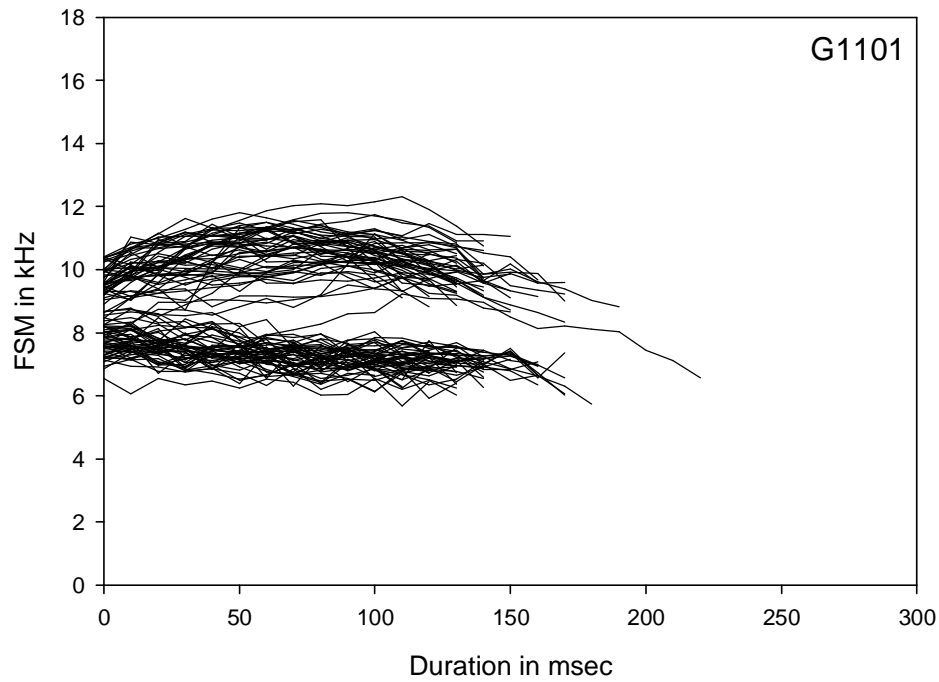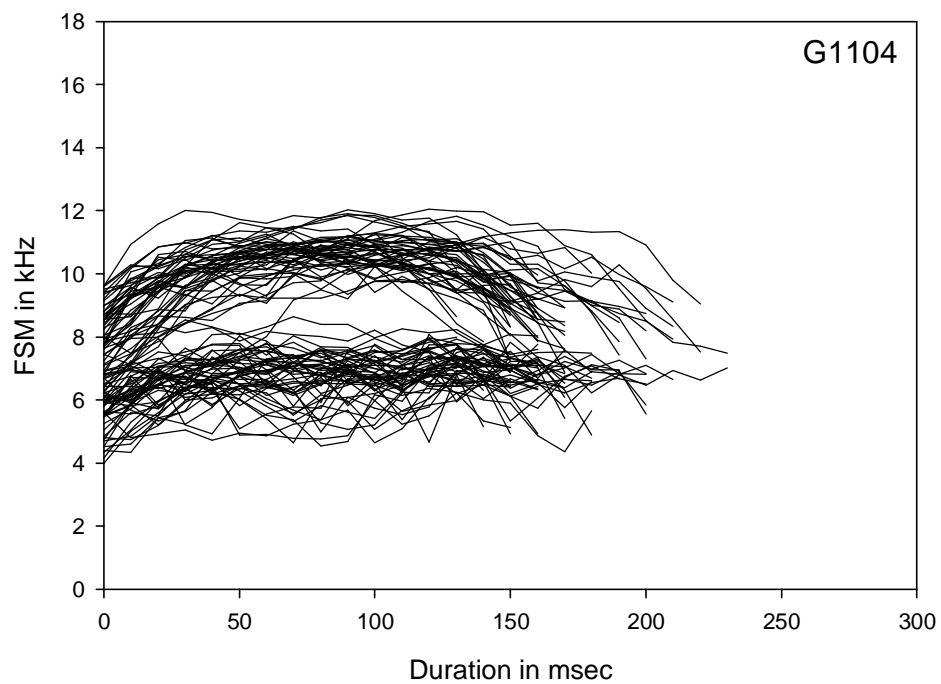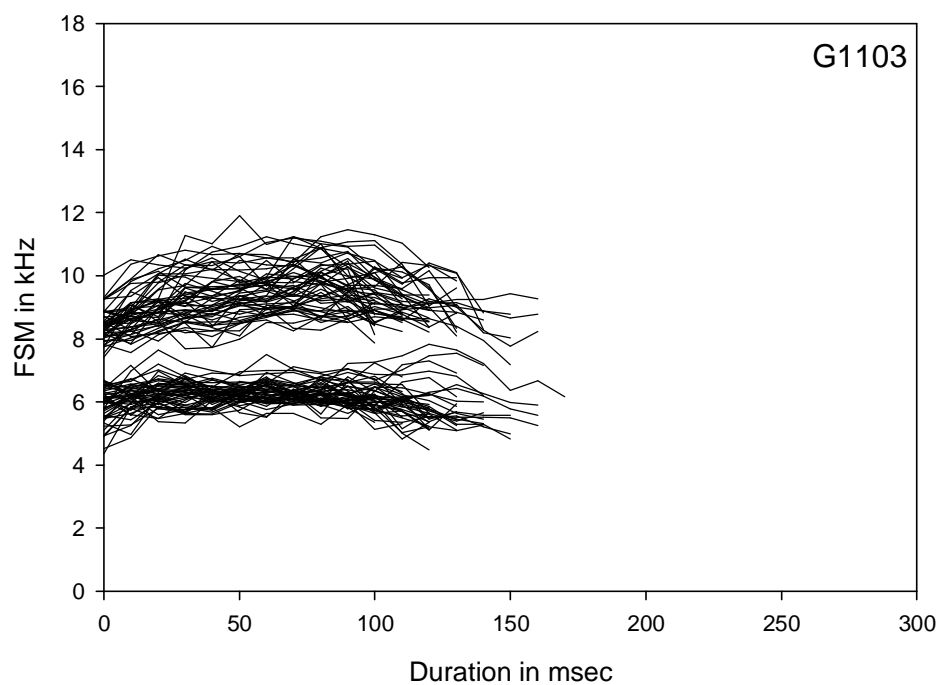represents the first analysis window positioned 10 ms into the fricative segment.

66

Figure 5, continued.
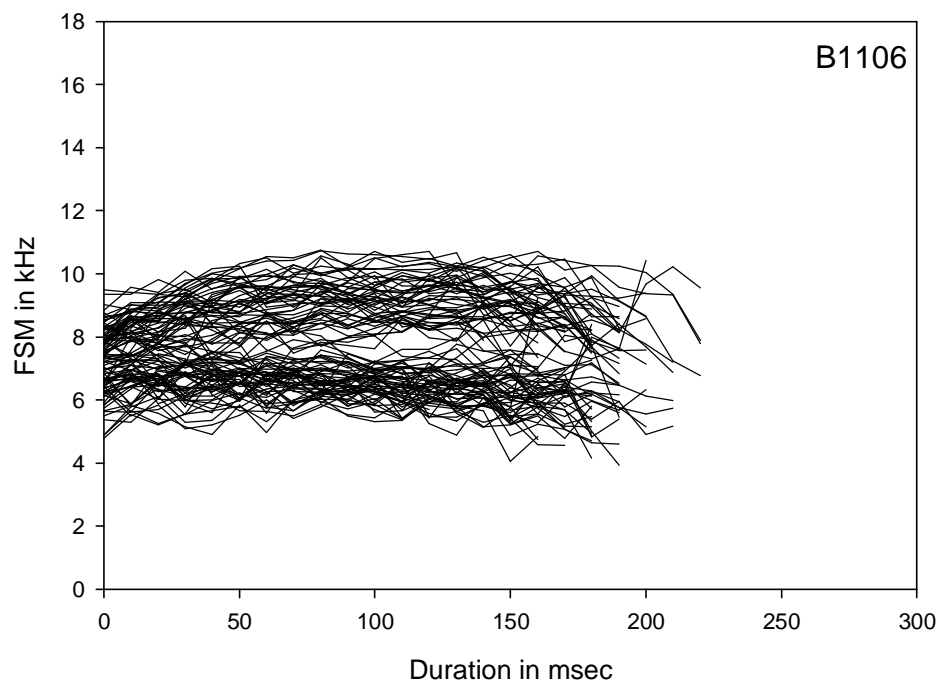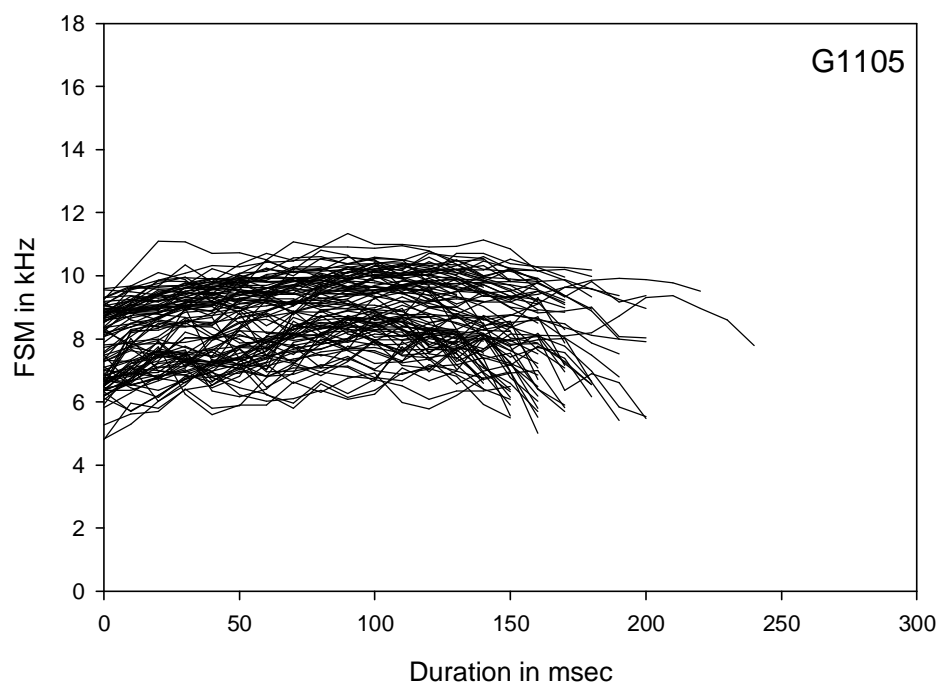
Figure 5, continued.
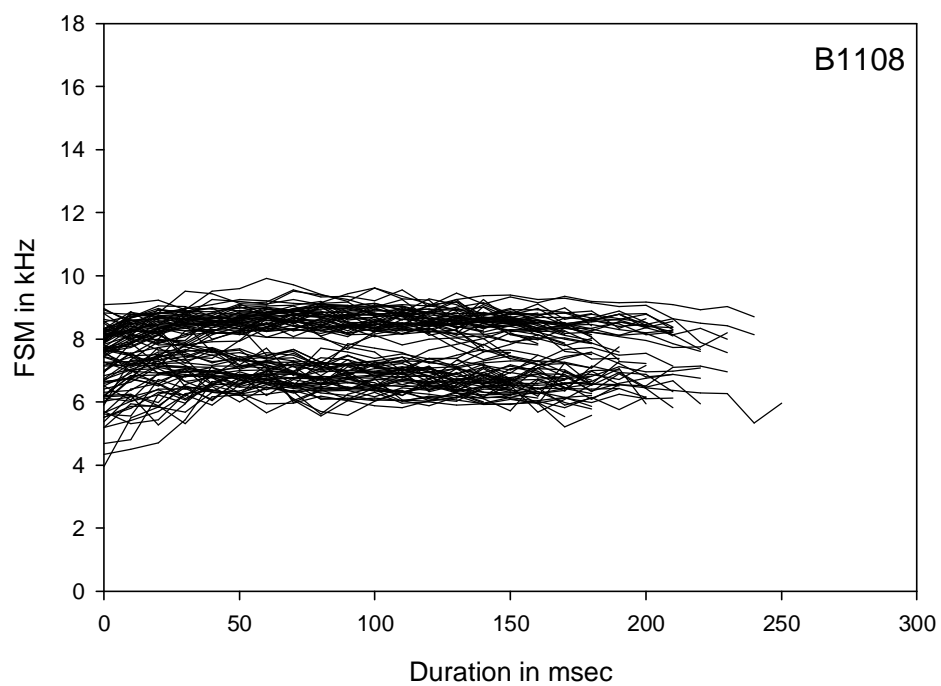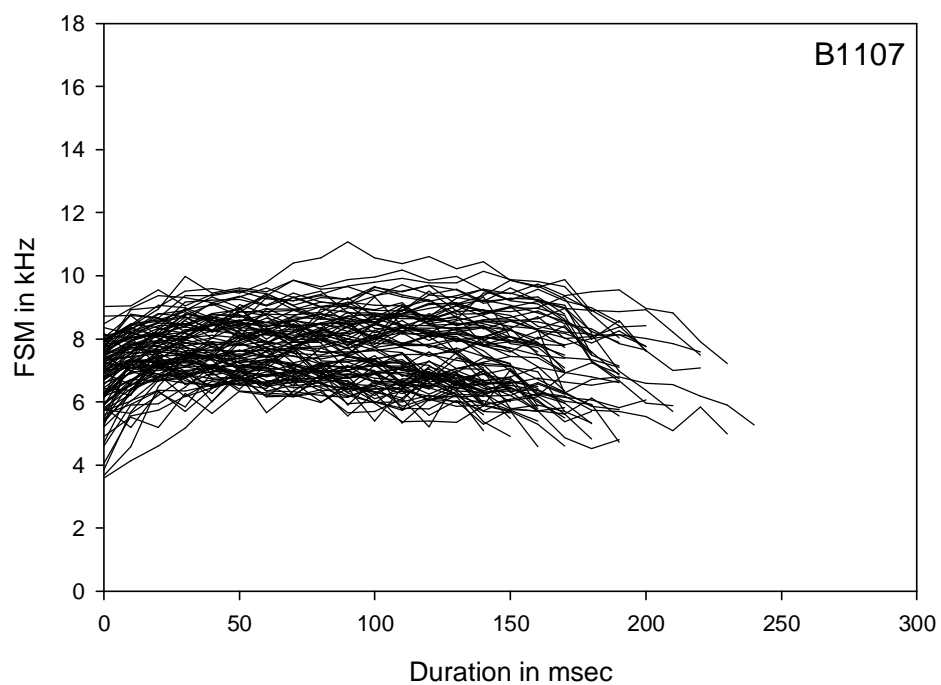
Figure 5, continued.

Figure 5, continued.
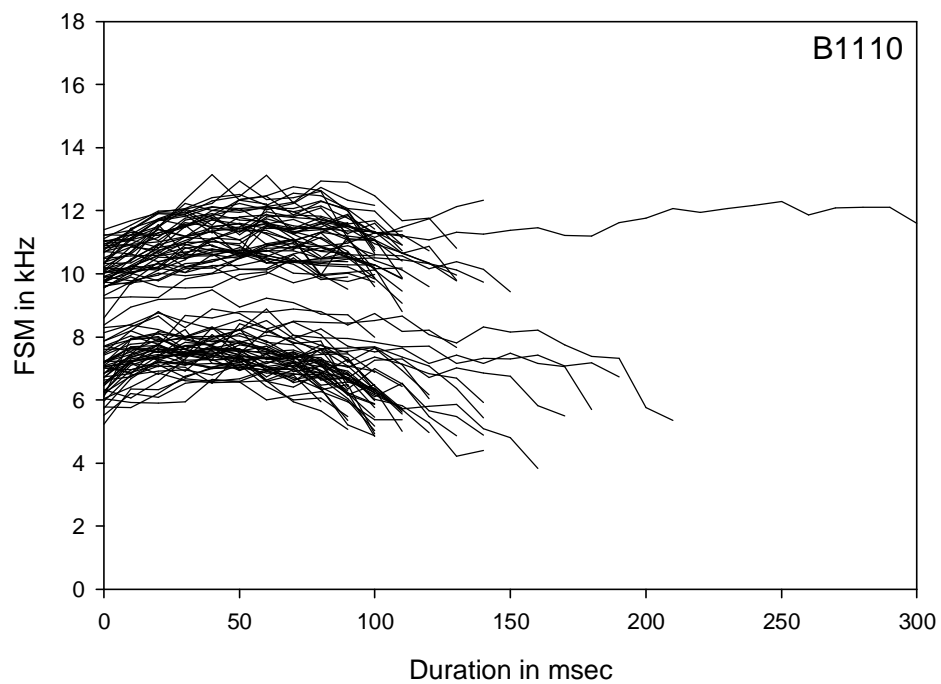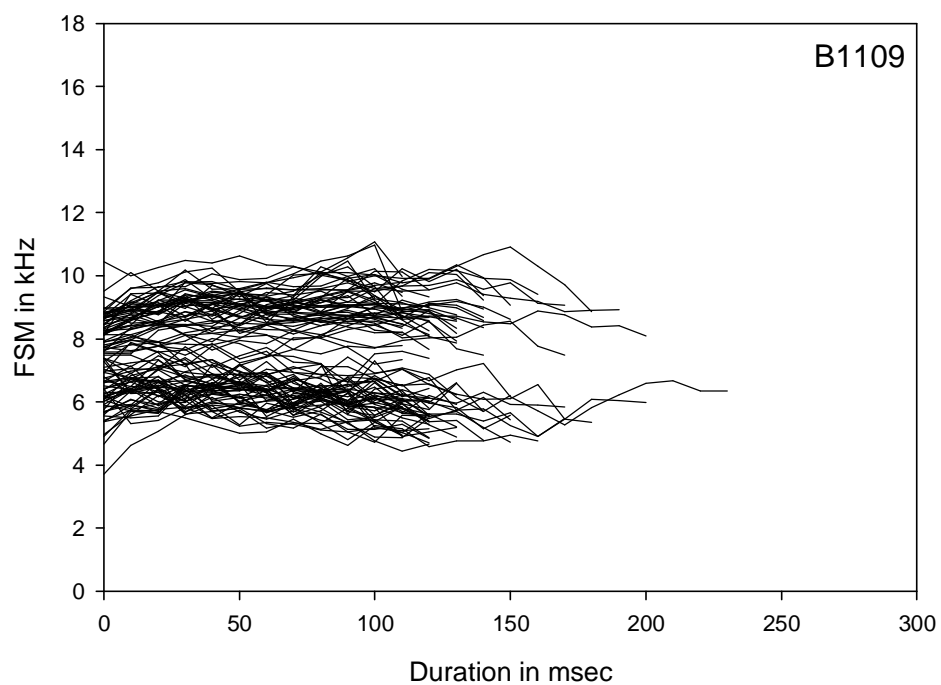
Only one participant (B708), a 7-year-old boy, showed a magnitude of distinction less than 1 kHz in production between [s] and [ʃ] measured by spectral moment analysis and substantial overlap (28%) between [s] and [ʃ] productions. Figure 6 shows the time history plots for [s] and [ʃ] separately for speaker B708. On the time history plots, the [s] productions demonstrate a decreasing FSM throughout the duration of the fricative segment. Unlike the other participants in the TD-7 and TD-11 groups, this participant's fricatives demonstrated considerable overlap yet were judged to be produced correctly on the speech assessment (GFTA-2) during the screening procedures. When comparing the mean FSM for this speaker's [s] and [ʃ], his [s] productions were much lower, on average, than the other participants in the TD-7 group (mean FSM for [s] for B708 = 7.2 kHz, range of FSM for [s] of other participants in the group = 9.3 kHz – 11.9 kHz). This may indicate that B708 was producing the [s] targets with a tongue position slightly backed in the oral cavity, closer to the point of constriction for [ʃ] targets. Although B708's [s] productions were distinct from his [ʃ] productions, they were observed to be somewhat "noisier" than what might be expected for [s] productions. No overt production error, such as a frontal or lateral lisp, was observed by the examiner.

Figure 6. First spectral moment (FSM) time history plots for participant B708's repeated productions of [s] and [ʃ], separated by consonant. Zero ms on the time axis represents the first analysis window positioned 10 ms into the fricative segment.
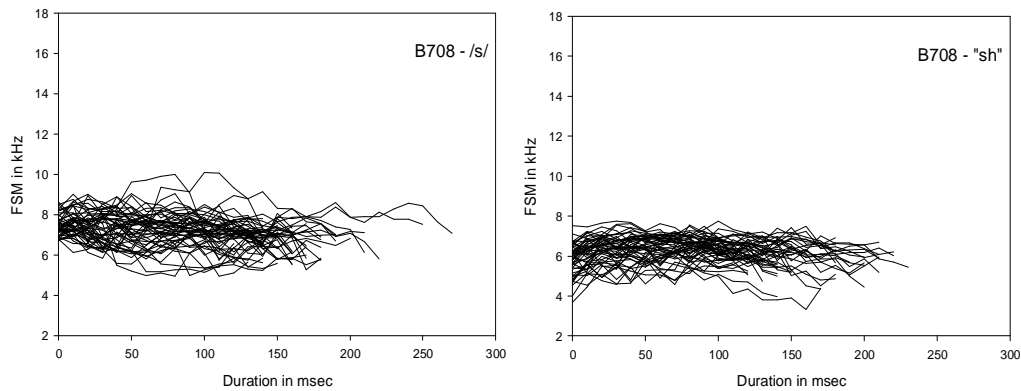
Vowel Effects on Distinction of Fricative Production for Individual Speakers

The fricative + [i] or fricative + [u] represented two different coarticulatory configurations, as the resonating frequencies of the fricatives [s] and [ʃ] could be differently affected by the lip rounding required for articulation of [u]. The effect of the vowel, if any, on the magnitude of distinction between the fricatives was also examined in this study. Time history plots of FSM for [s] and [ʃ] separated by the vowels [i] and [u] are displayed in Figures 7 and 8.

Figure 7. First spectral moment (FSM) time history plots for TD-7 participants' repeated productions of [s] and [ʃ] separated by vowel. G701 – G705 are female speakers and B706 – B711 are male speakers. Time histories across the repeated utterances are superimposed on each other to display categorical distinction visually. The higher frequency clusters are [s] targets and the lower frequency clusters are [ʃ] targets. Zero ms on the time axis represents the first analysis window positioned 10 ms into the fricative segment.

Figure 7, continued.
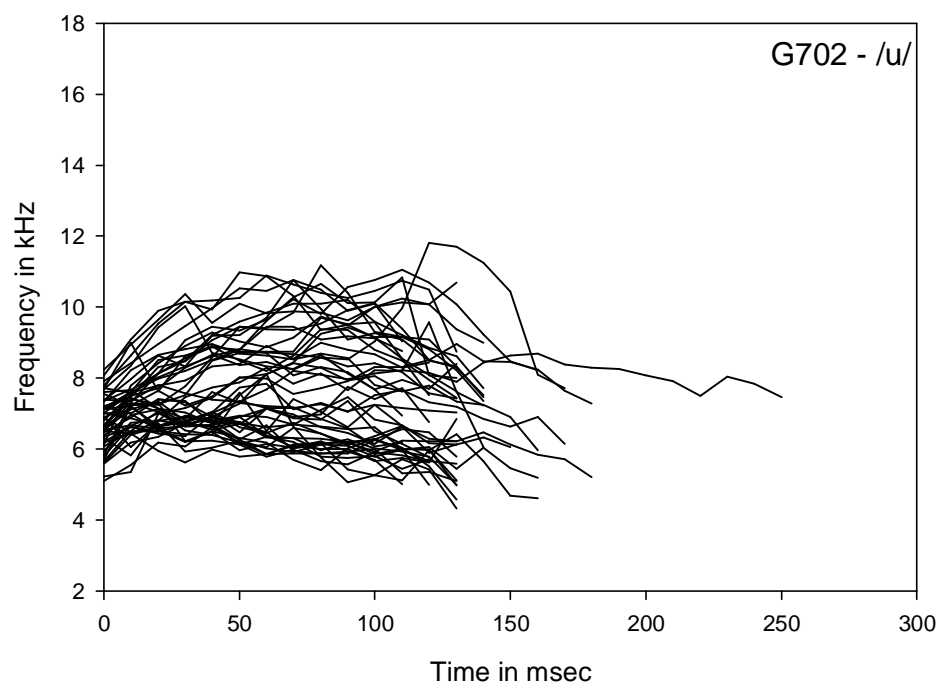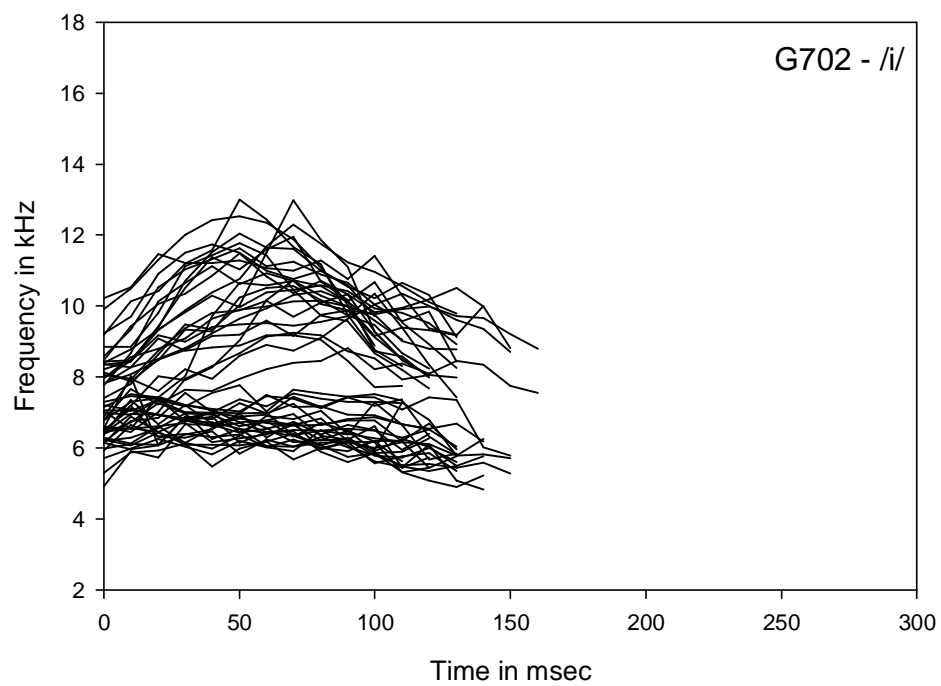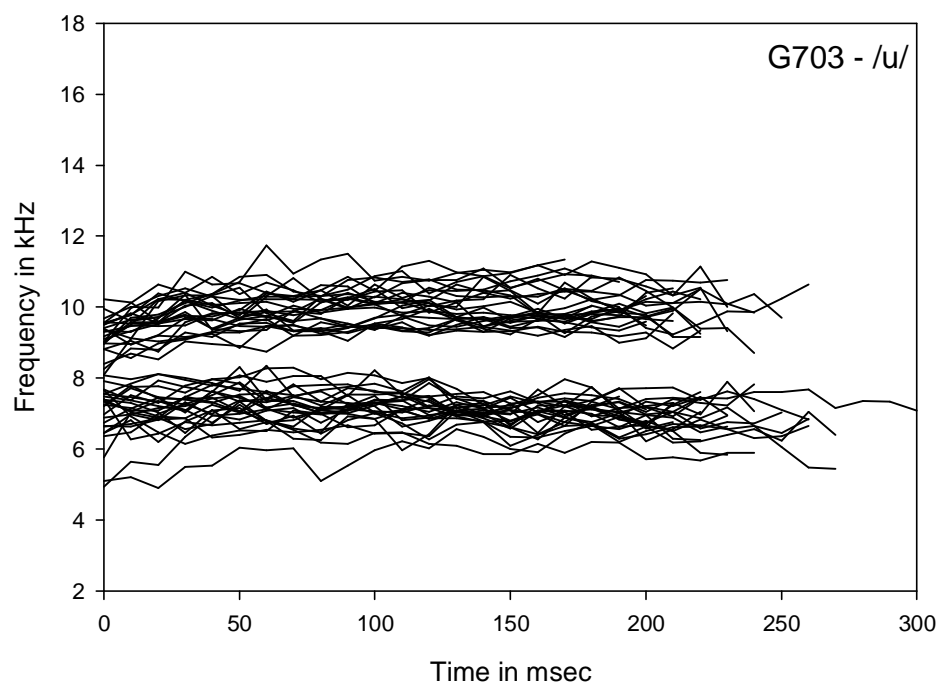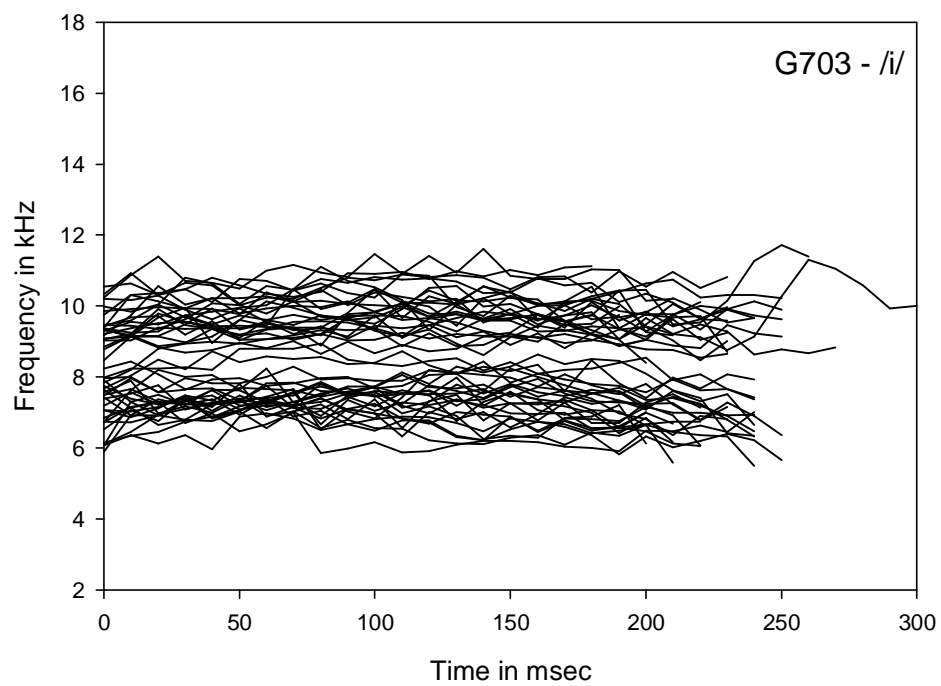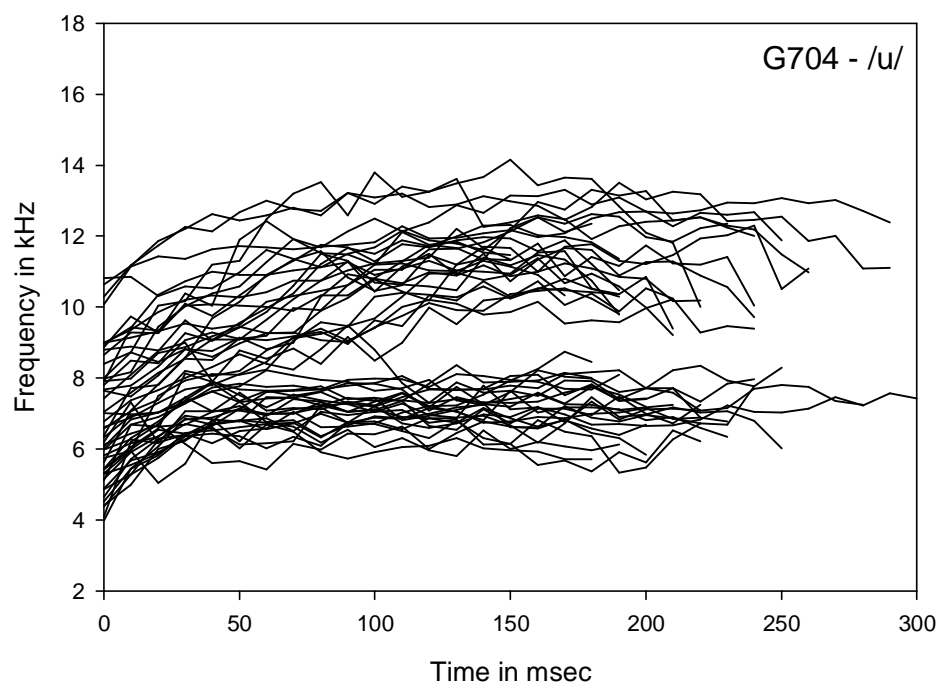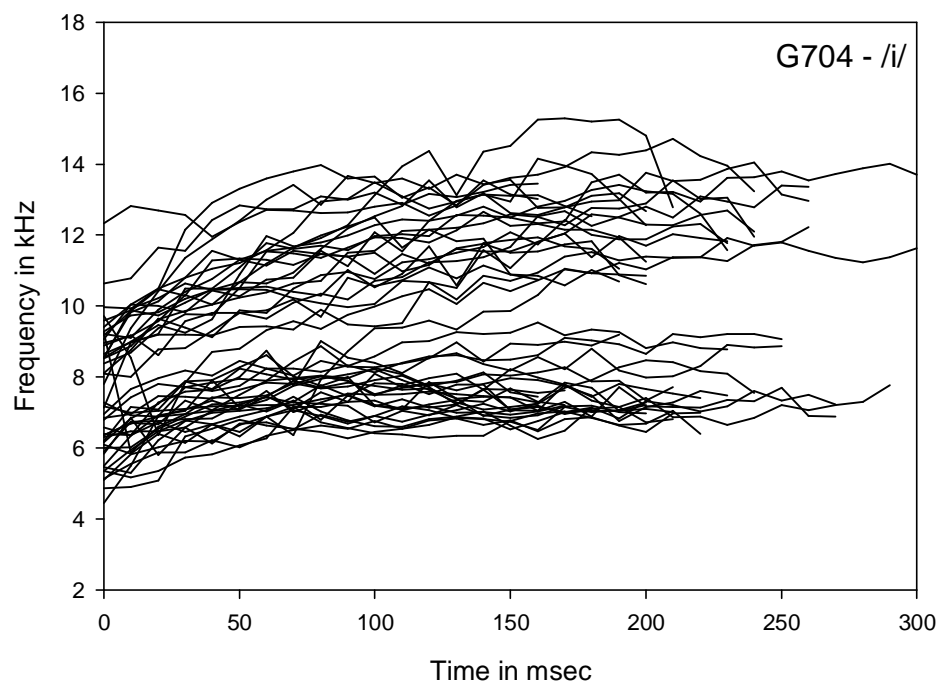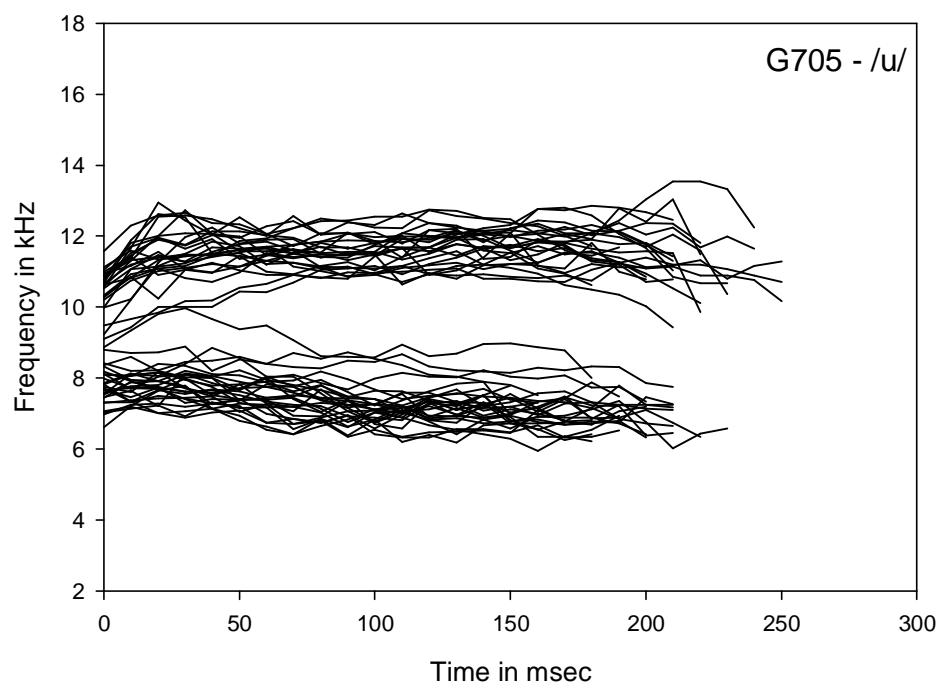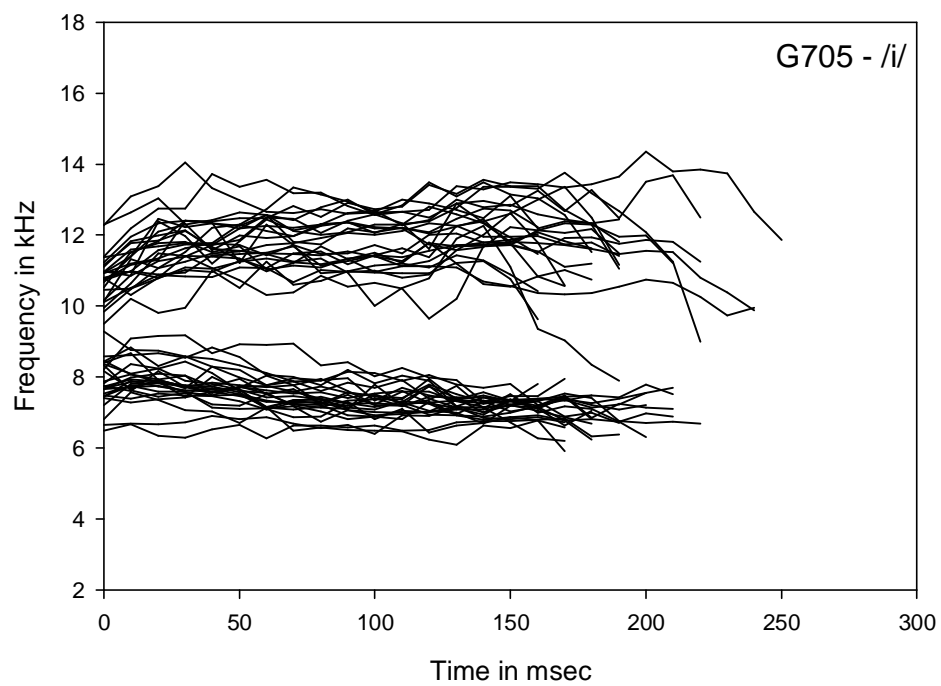
Figure 7, continued.

Figure 7, continued.

Figure 7, continued.
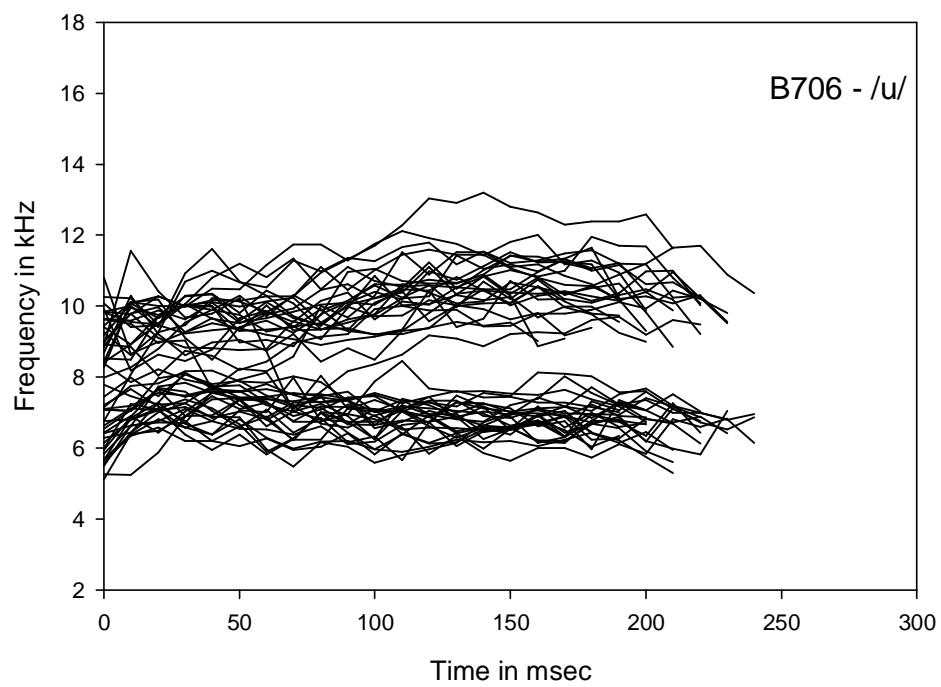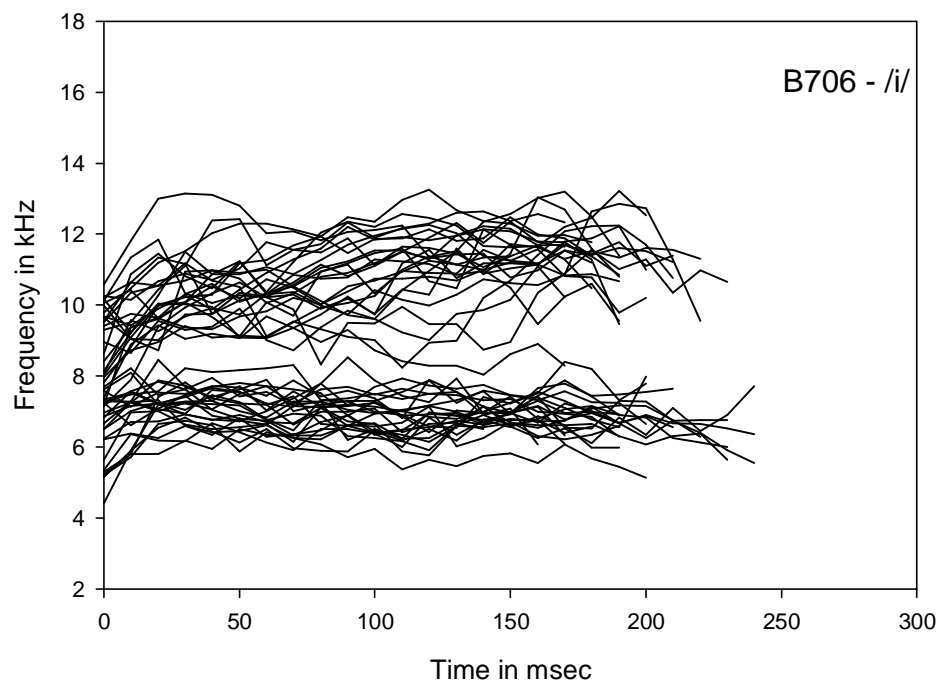
Figure 7, continued.
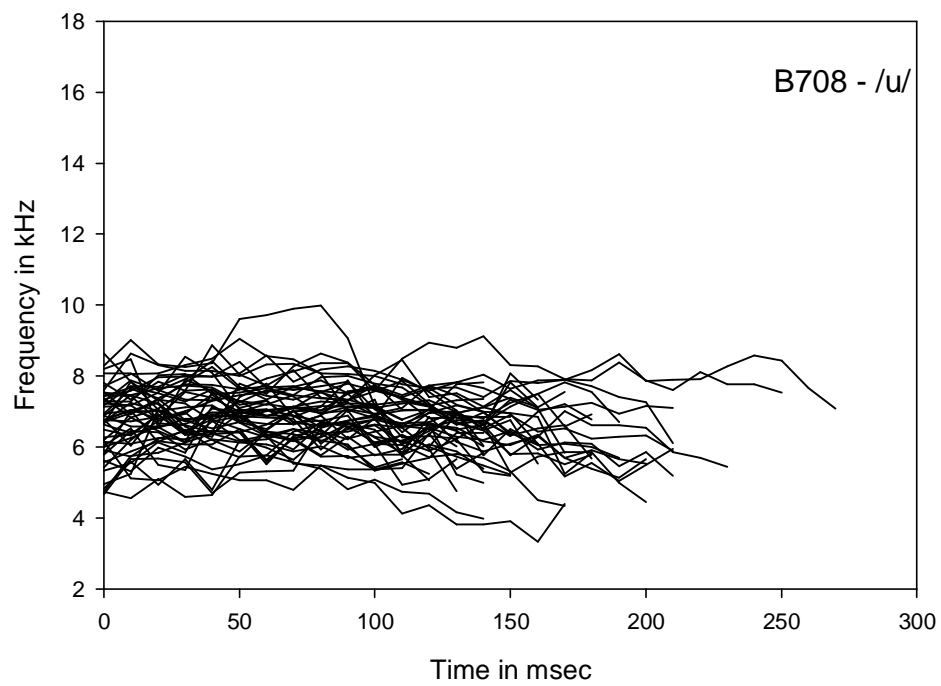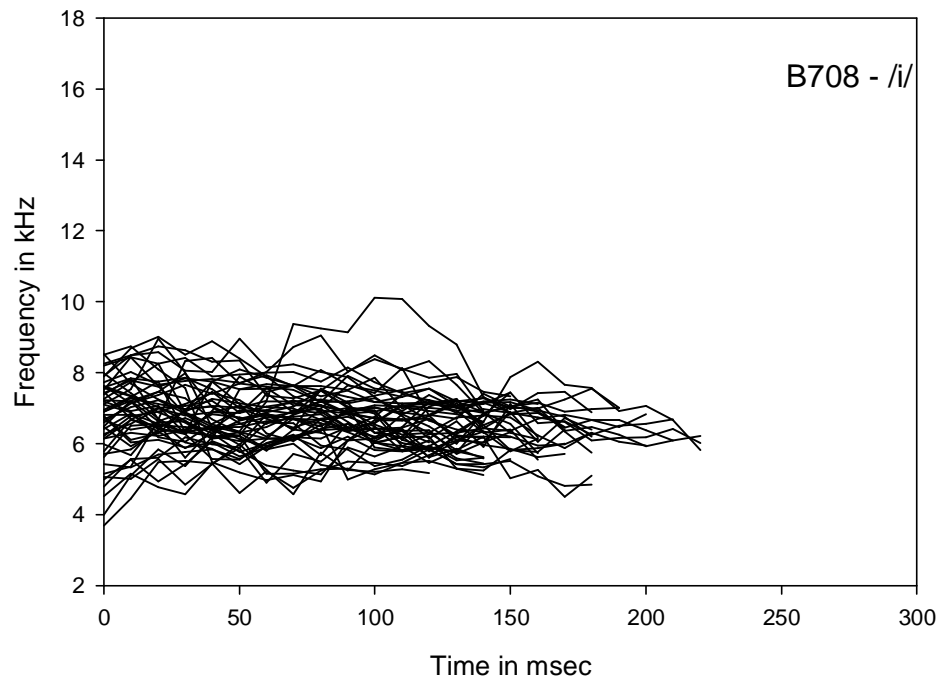
Figure 7, continued.
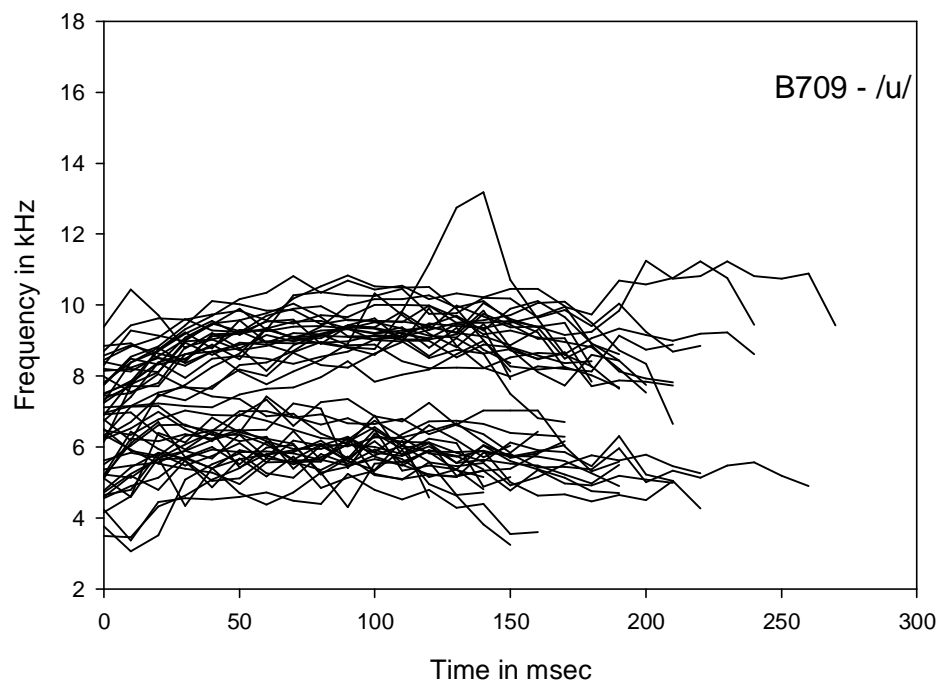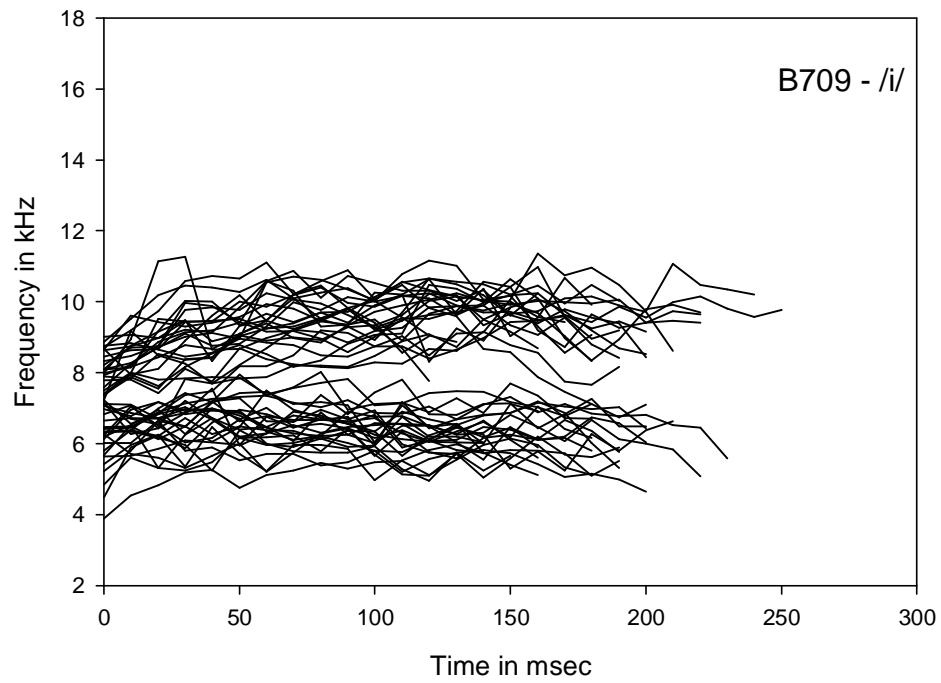
Figure 7, continued.
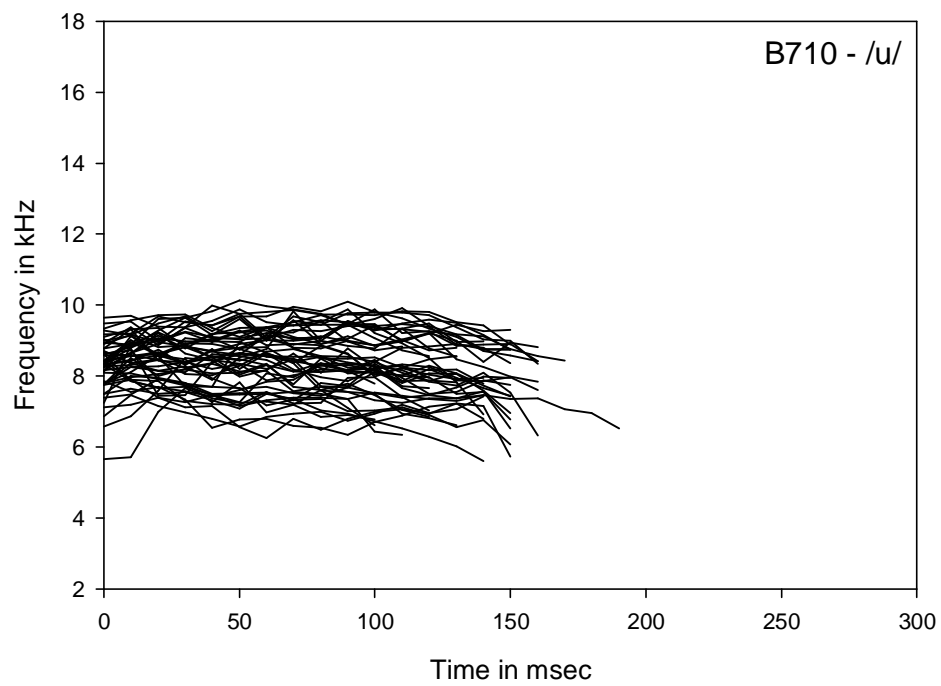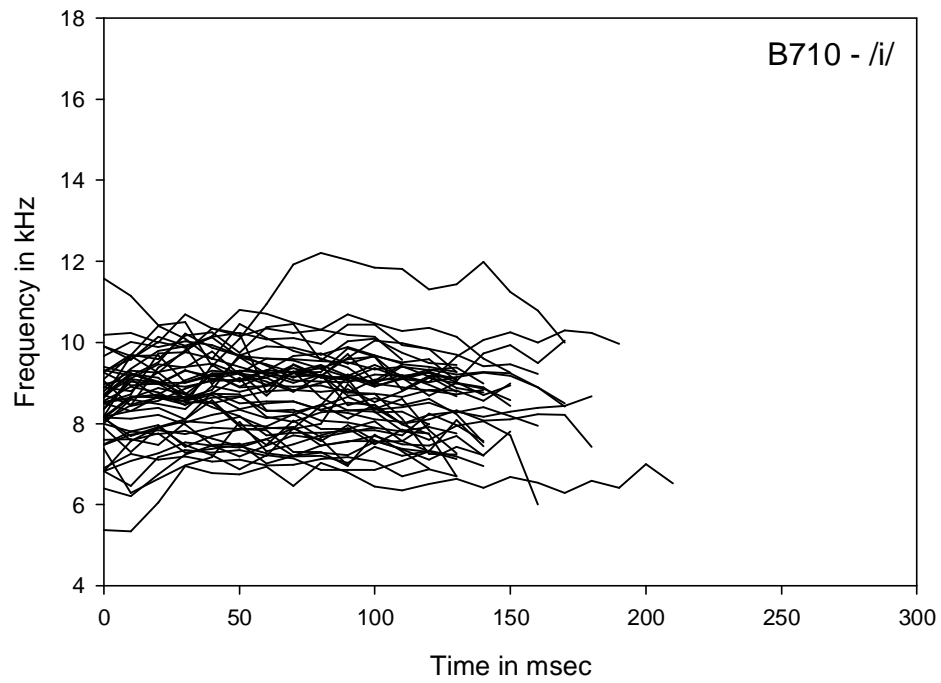
Figure 7, continued.

Figure 7, continued.
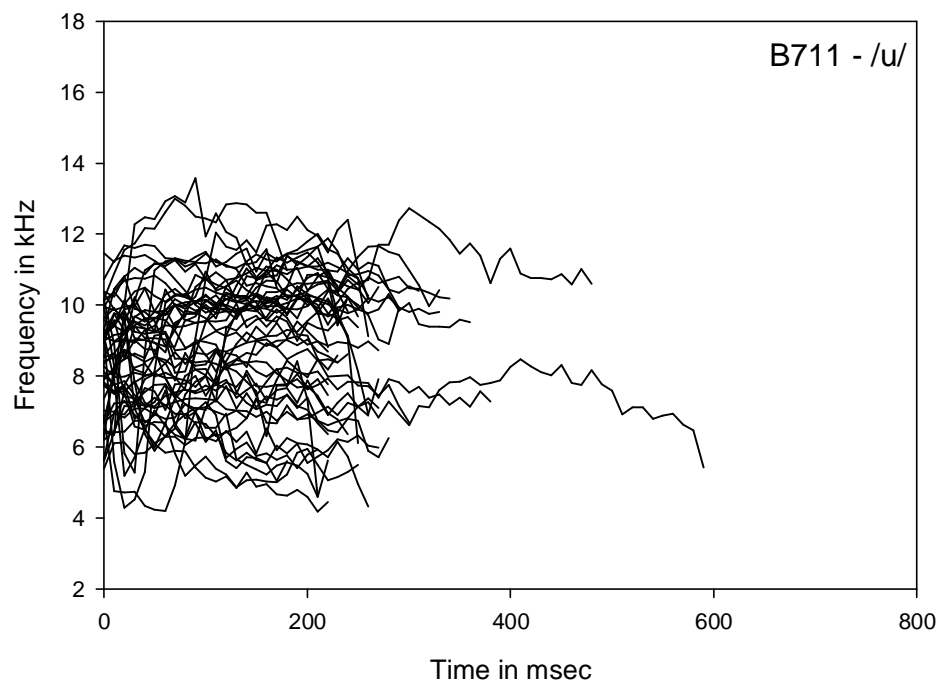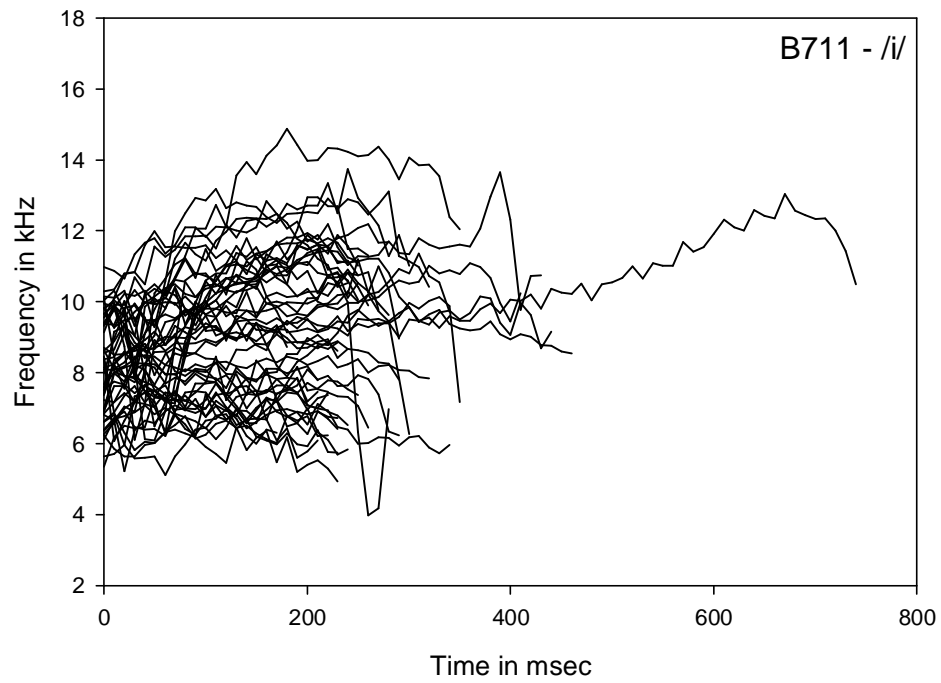
Figure 8. First spectral moment (FSM) time history plots for TD-11 participants' repeated productions of [s] and [ʃ] separated by vowel. G1101 – G1105 are female speakers and B1106 – B1110 are male speakers. Time histories across the repeated utterances are superimposed on each other to display categorical distinction visually. The higher frequency clusters are [s] targets and the lower frequency clusters are [ʃ] targets. Zero ms on the time axis represents the first analysis window positioned 10 ms into the fricative segment.

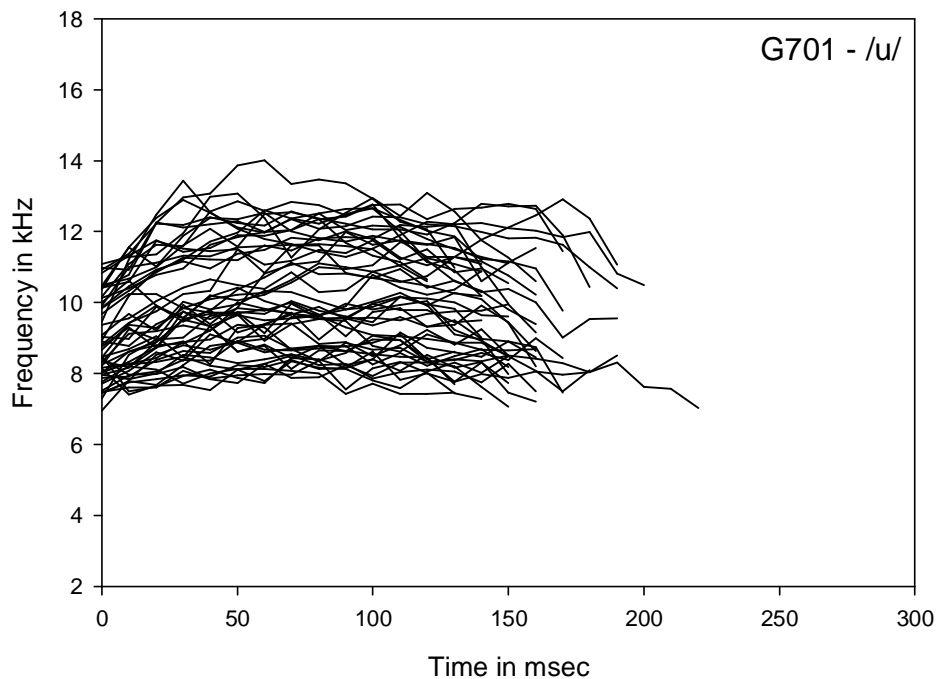Figure 8, continued.
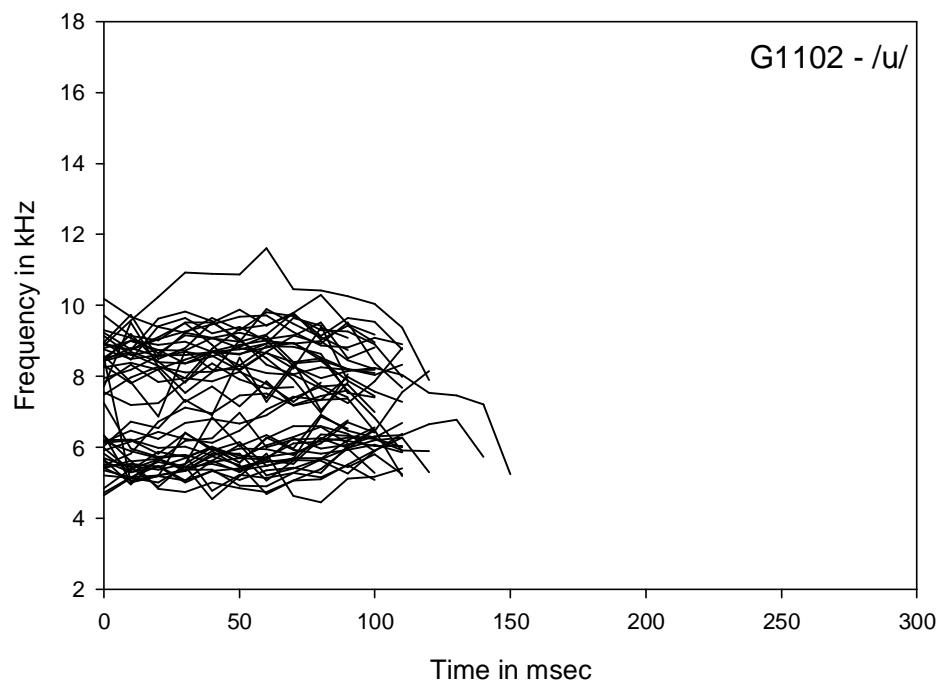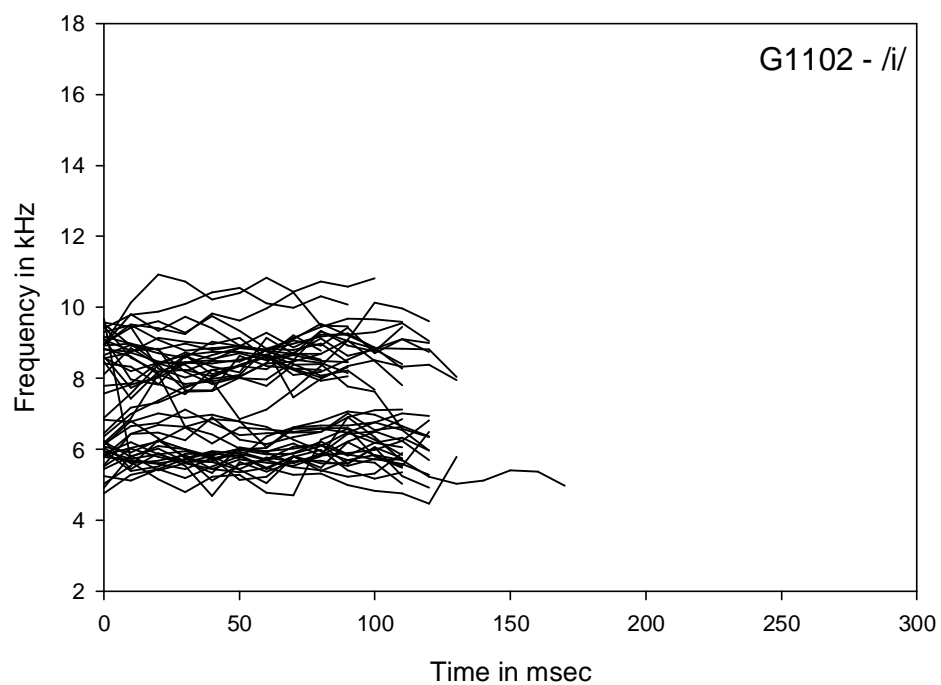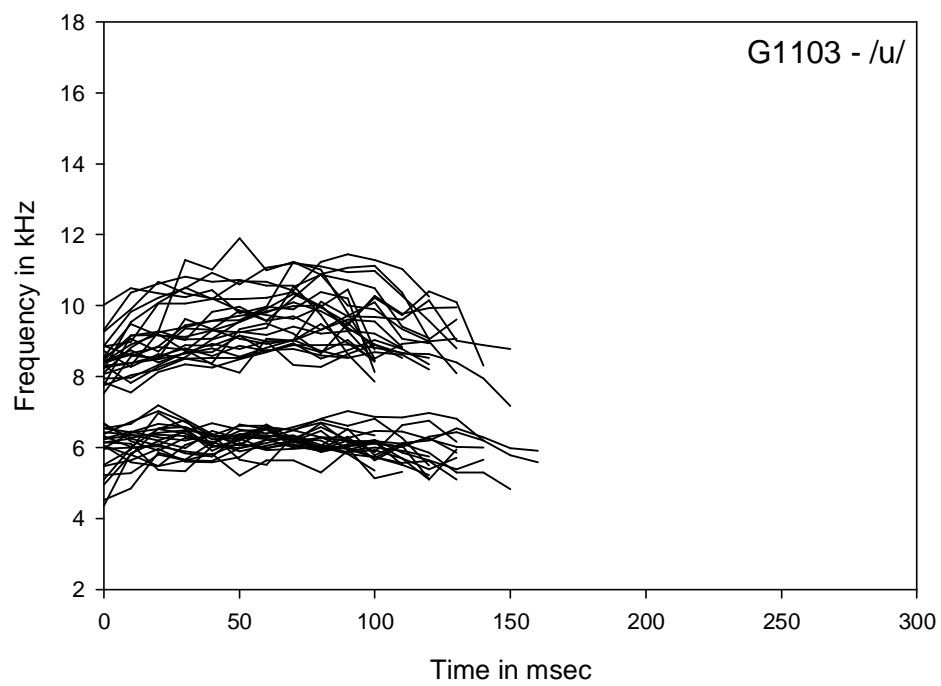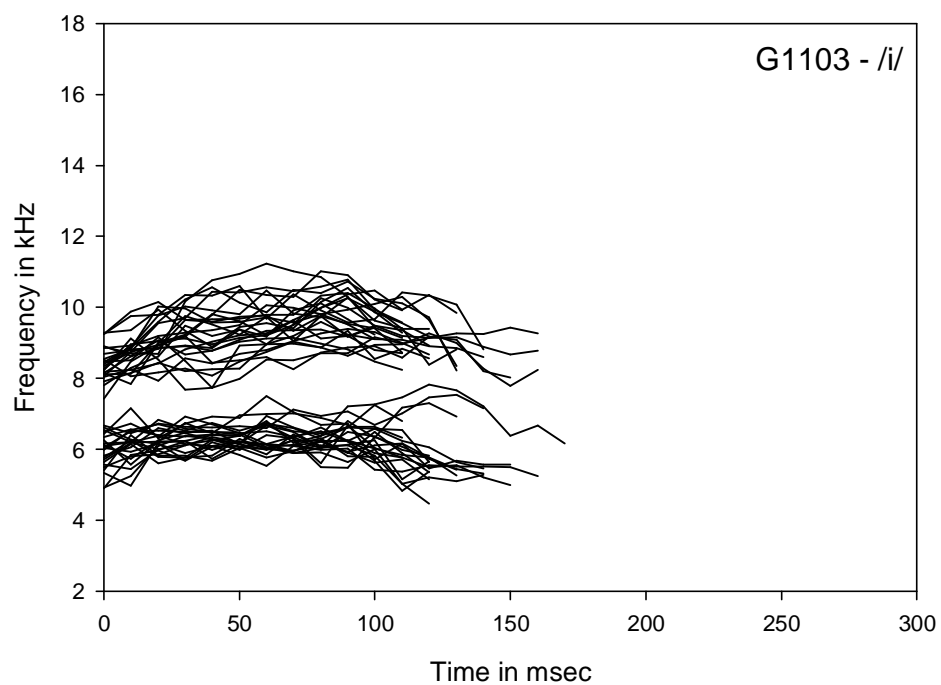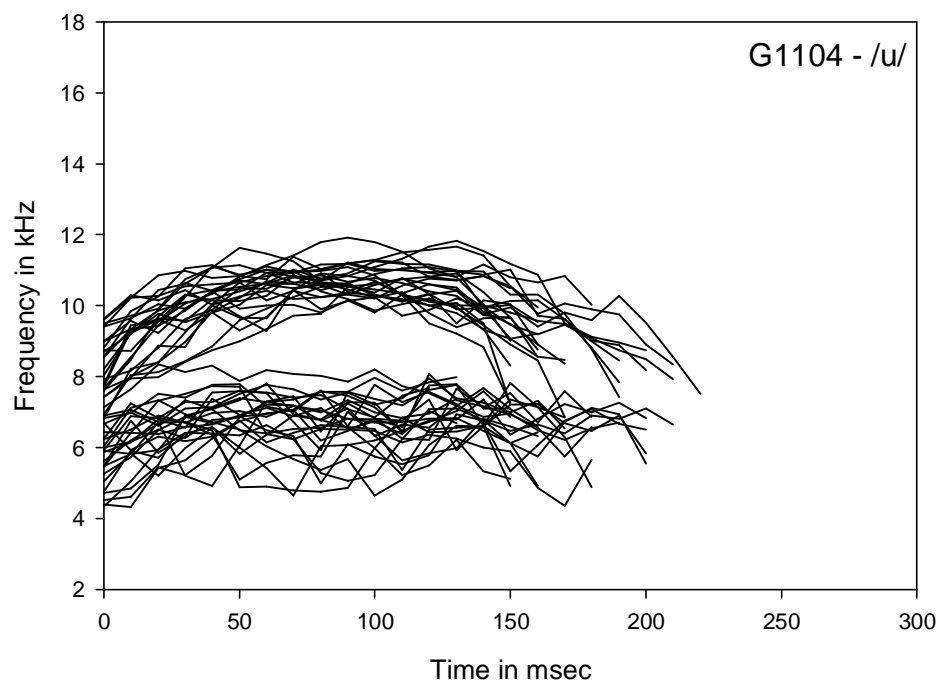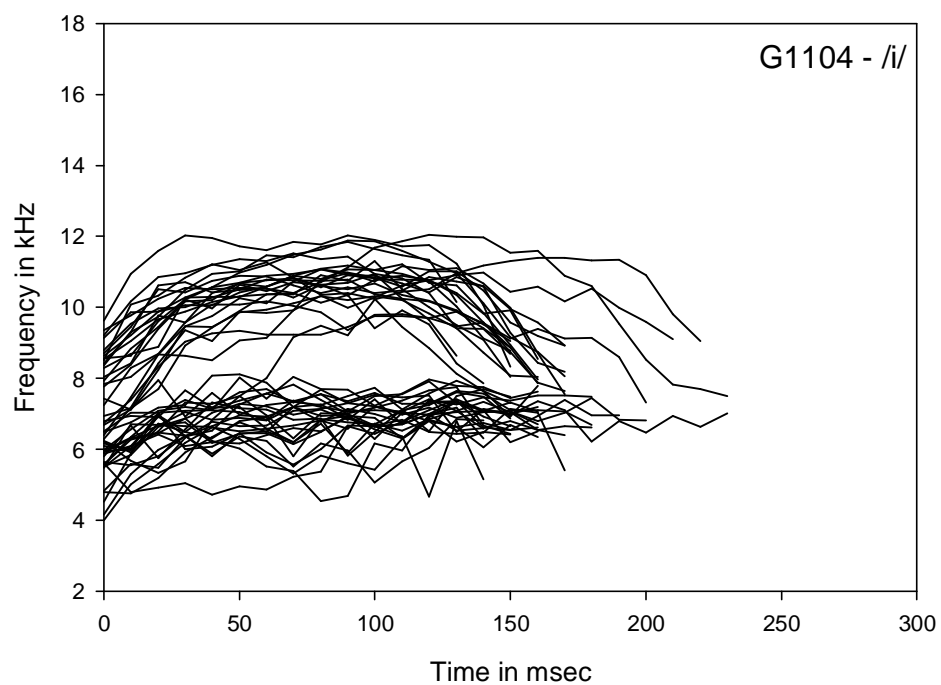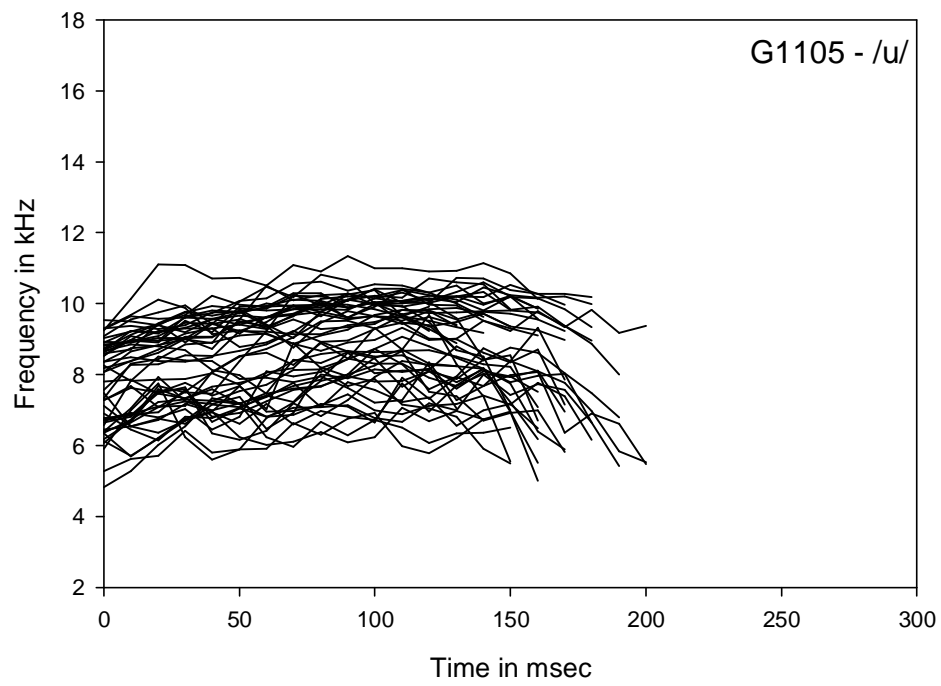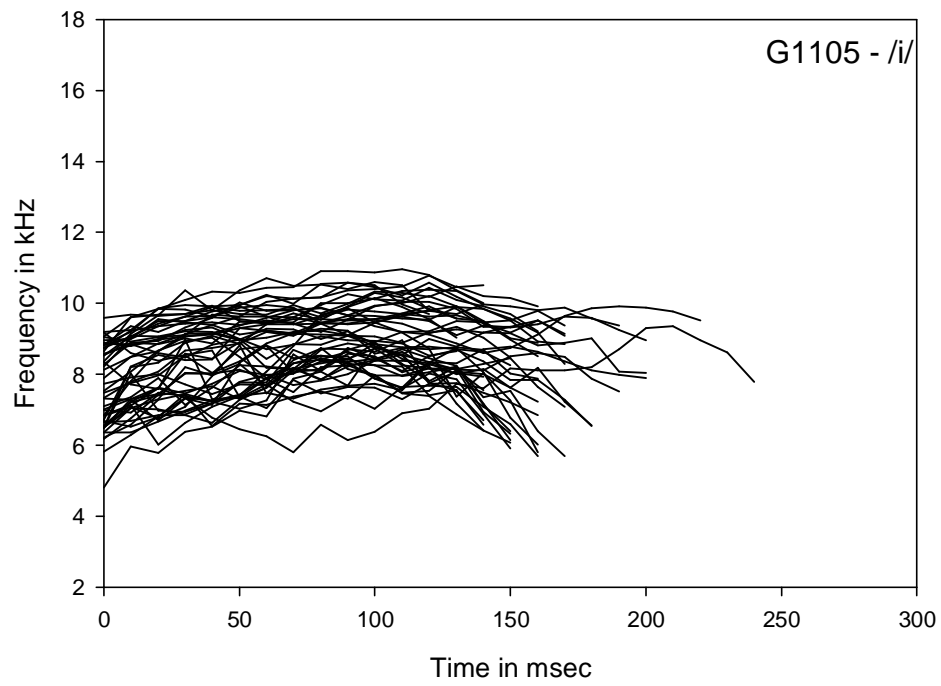
Figure 8, continued.

G1103 - /i/

G1103 - /u/

Figure 8, continued.

Figure 8, continued.

Figure 8, continued.

Figure 8, continued.

Figure 8, continued.

B1108 - /i/

B1108 - /u/
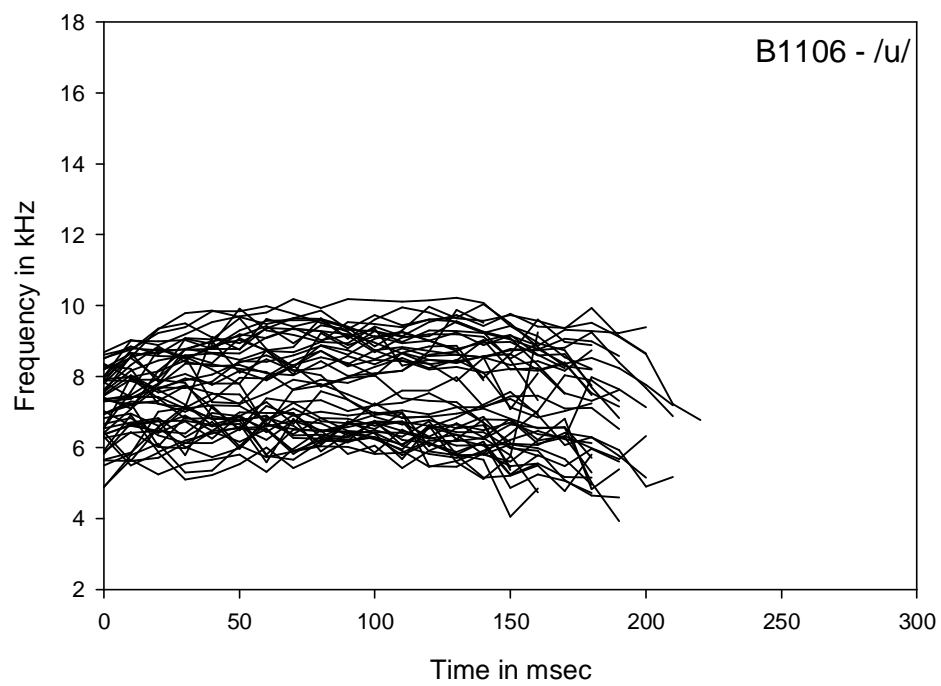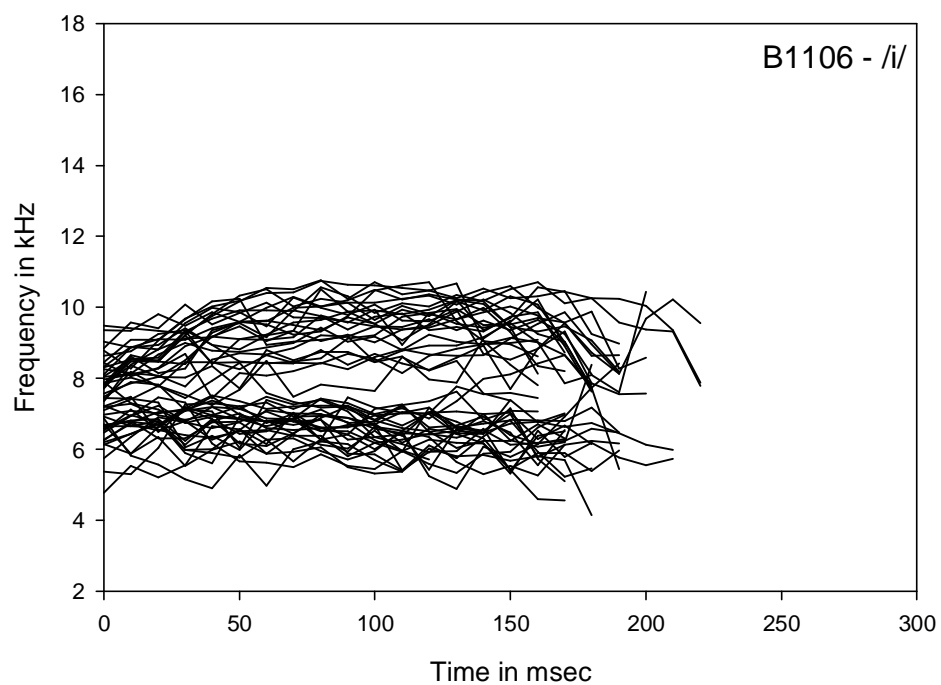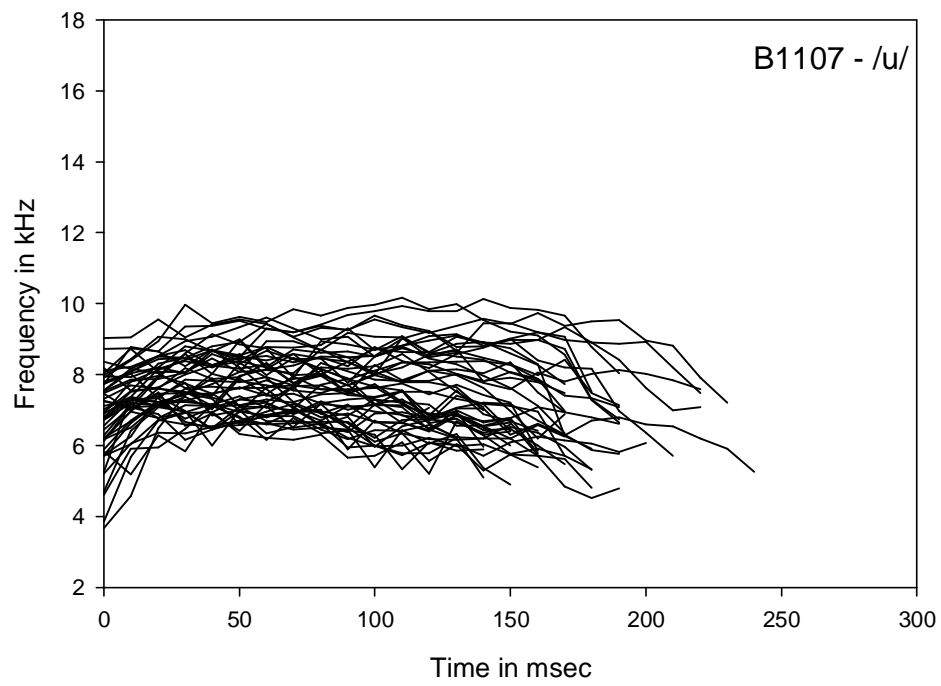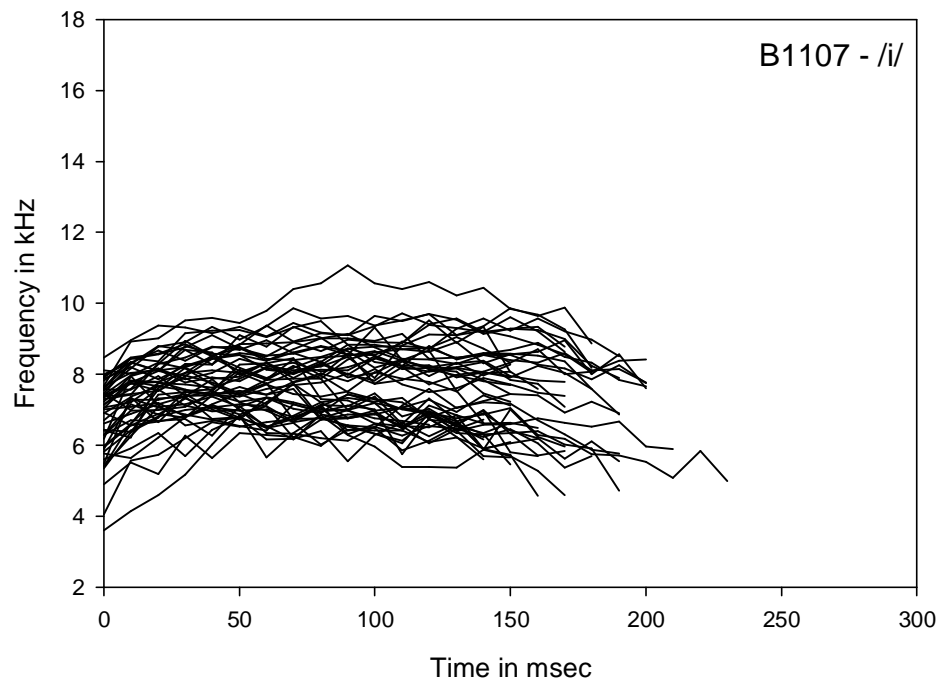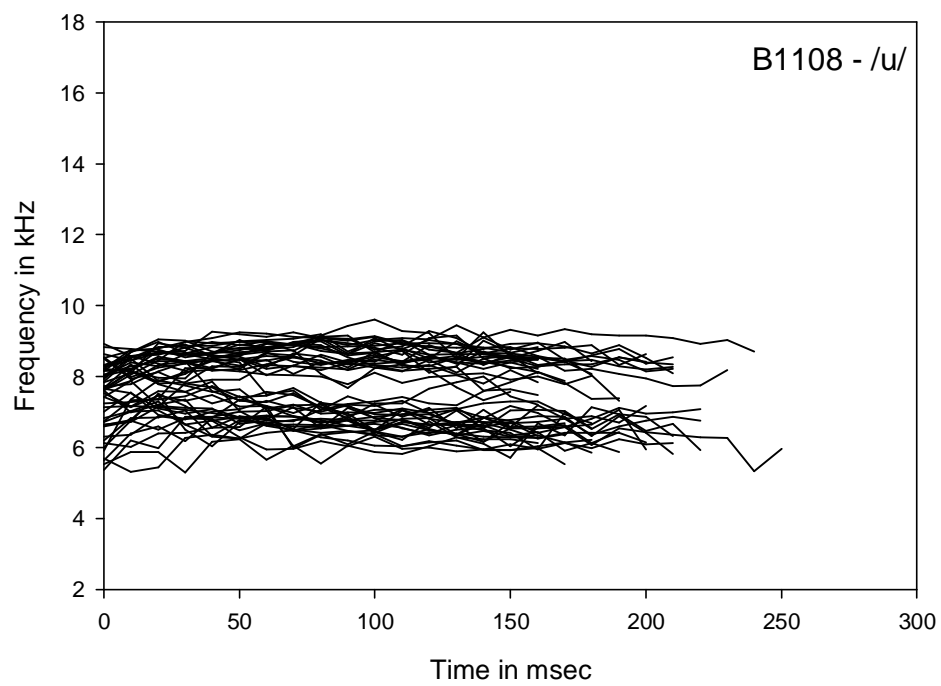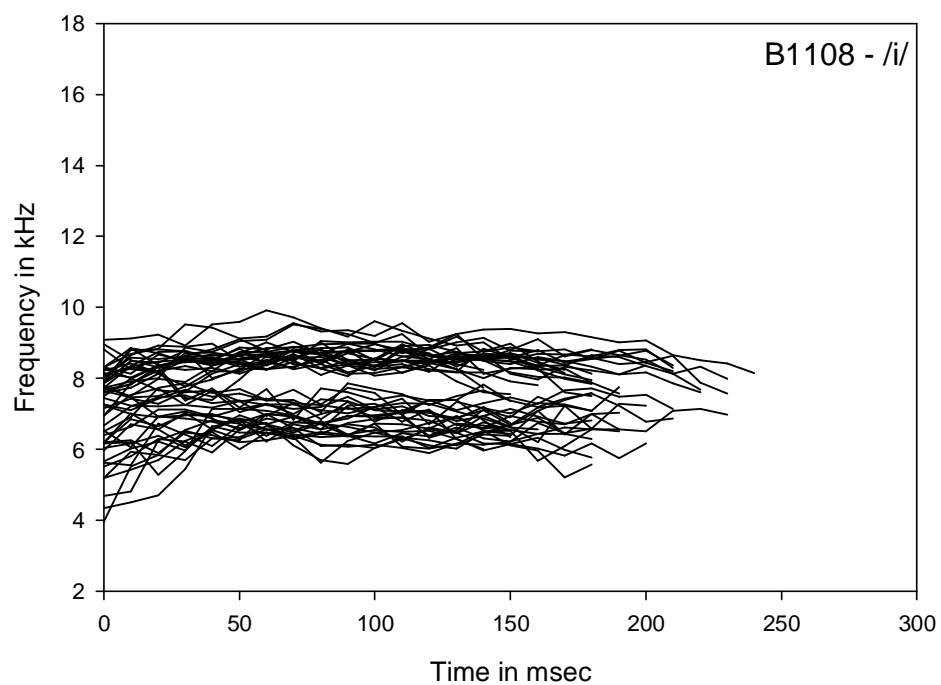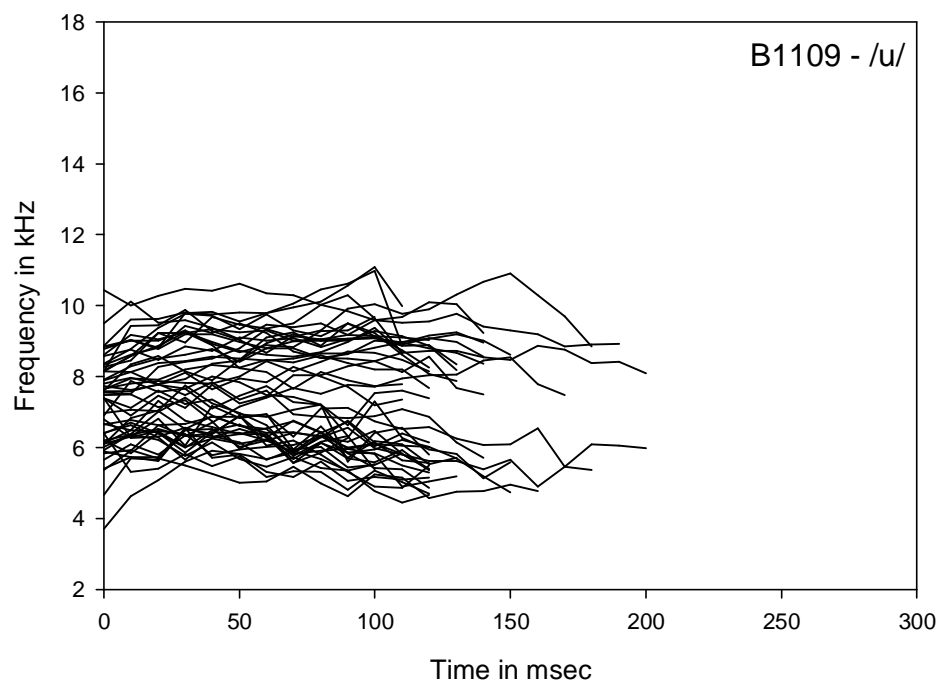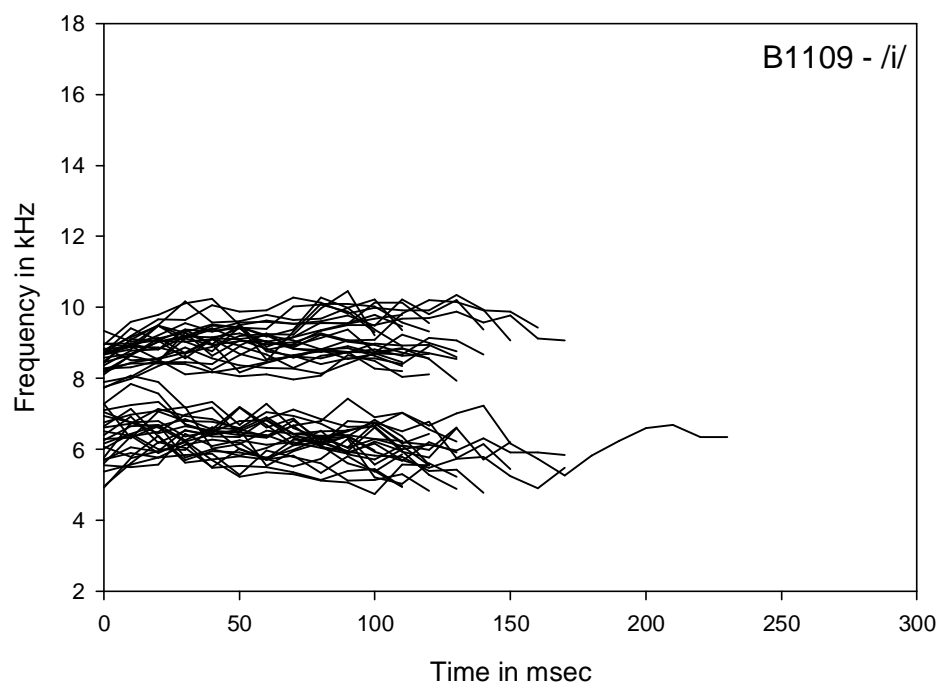
Figure 8, continued.

Figure 8, continued.
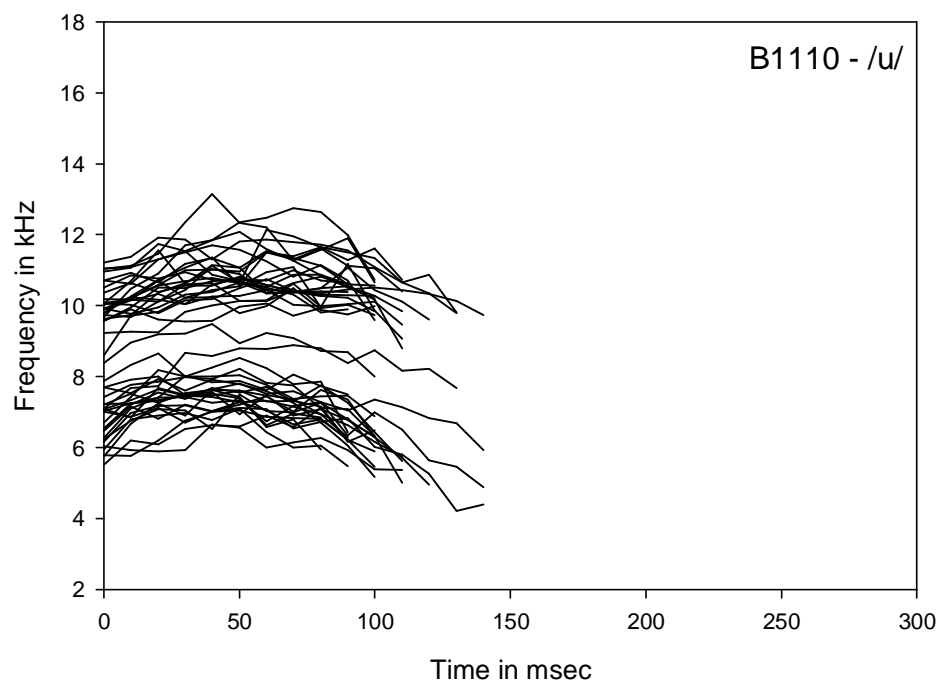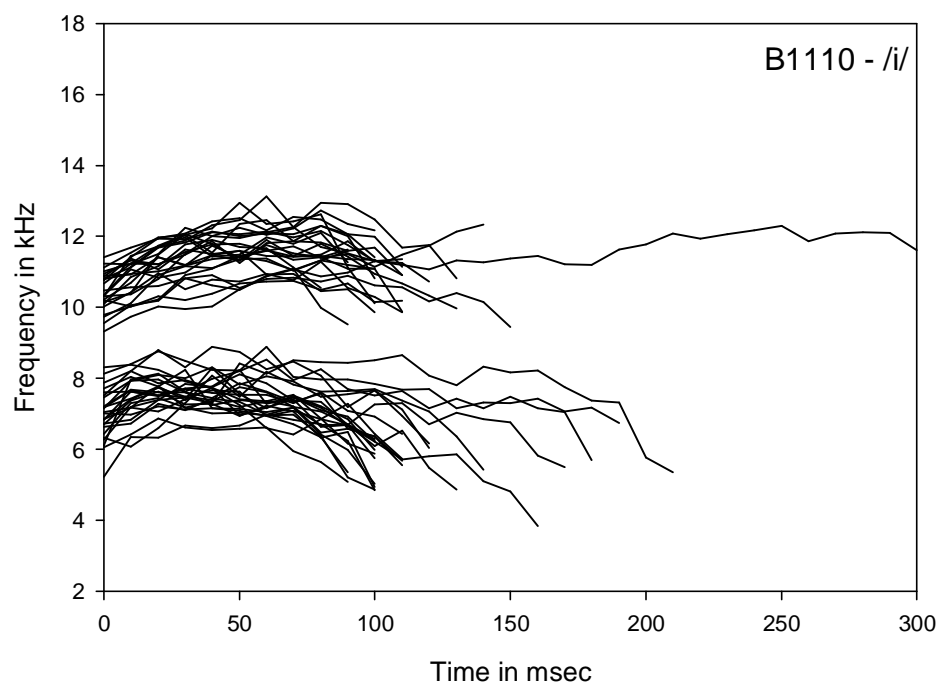
When inspecting Figures 7 and 8, the co-articulatory influence of the vowel [u] can account for some production overlap in a subset of speakers (G701, G702, B710, G1101, G1102, B1106, B1107, B1109, B1110). In these speakers, the magnitude of distinction between [s] and [ʃ] is decreased in the [u] vowel context compared to the [i] context.

The values for FSM at midpoint for [si], [su], [ʃi], and [ʃu] were averaged among participants in both the TD-7 and TD-11 groups and then plotted. Figure 9 shows the mean FSM at midpoint separated by vowel. For [s] production, with the [i] and [u] vowel, both age groups showed similar patterns of producing [s] with higher spectral energy before the [i] vowel compared to the [u] vowel. This is an expected effect, as [u] is produced with lip rounding, effectively lengthening the vocal tract and subsequently lowering the resonating frequency of the preceding fricative. This effect is not as apparent for both age groups for [ʃ] production. This may also be an expected outcome, as production of [ʃ] may include lip rounding for both [ʃi] and [ʃu] syllables.

Figure 9. Mean FSM at midpoint, separated by vowels and by age groups. The x-axis displays the two age-groups, and the y-axis shows the mean FSM in kHz.
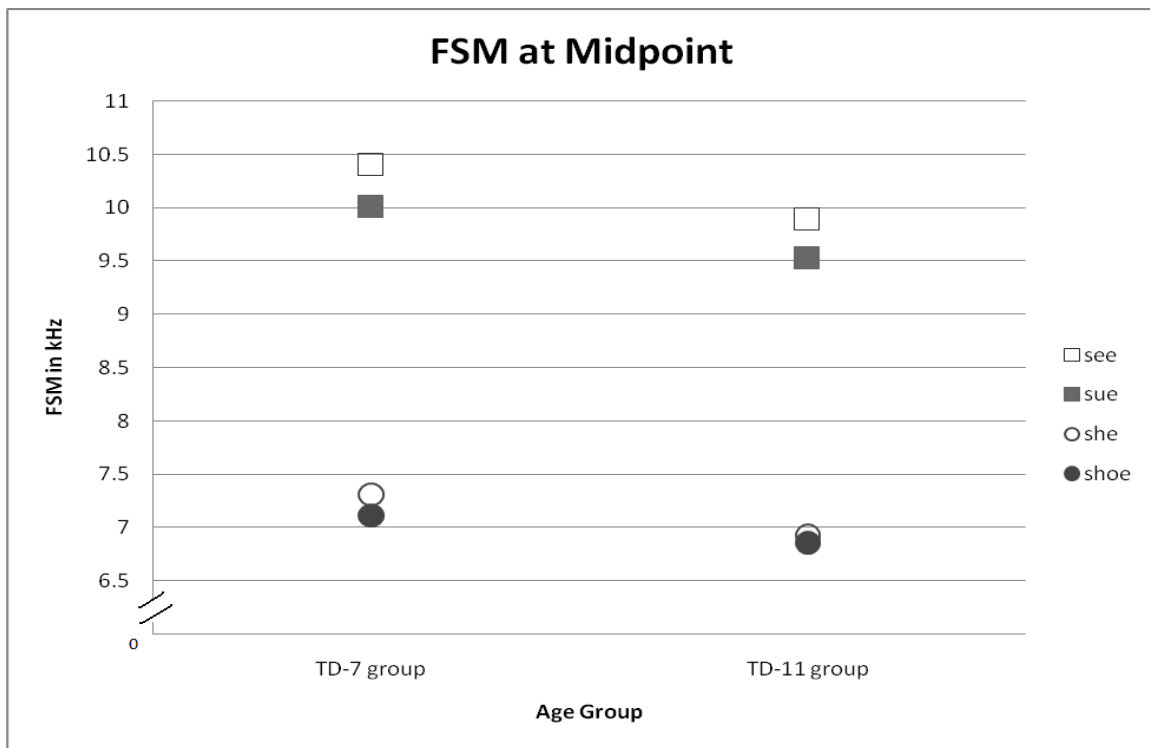
Developmental Differences in Production

Univariate independent sample t-tests were used to investigate group differences for production measures. At midpoint, the TD-11 group showed lower mean FSM for both [s] (9.6 kHz) and [ʃ] (6.8 kHz) production compared to the TD-7 group ([s] = 10.2 kHz, [ʃ] = 7.2 kHz); however, this difference was found to be non-significant [[s] (t = 1.115, *ns*); [ʃ] (t = 1.214, *ns*). The magnitude of the difference between [s] and [ʃ] at midpoint was slightly smaller for the TD-11 group (2.8 kHz) compared to the TD-7 group (3.0 kHz) by approximately 200 Hz. This difference was found to be non-significant (t = 0.487, *ns*). In the event these results reflect real differences between the two groups, they may reflect the differences in vocal tract size between the older and younger children, therefore

94

lowering resonating frequencies for the fricatives in the TD-11 group. Figures 10

and 11 display the magnitude of the [s] - [ʃ] difference for the individual

participants in the TD-7 and TD-11 group.

Figure 10. Mean FSM for [s] and [ʃ] showing the magnitude of difference between the two consonants for the TD-7 group. Mean FSM is shown in kHz and the error bars display variability of production for both consonants.
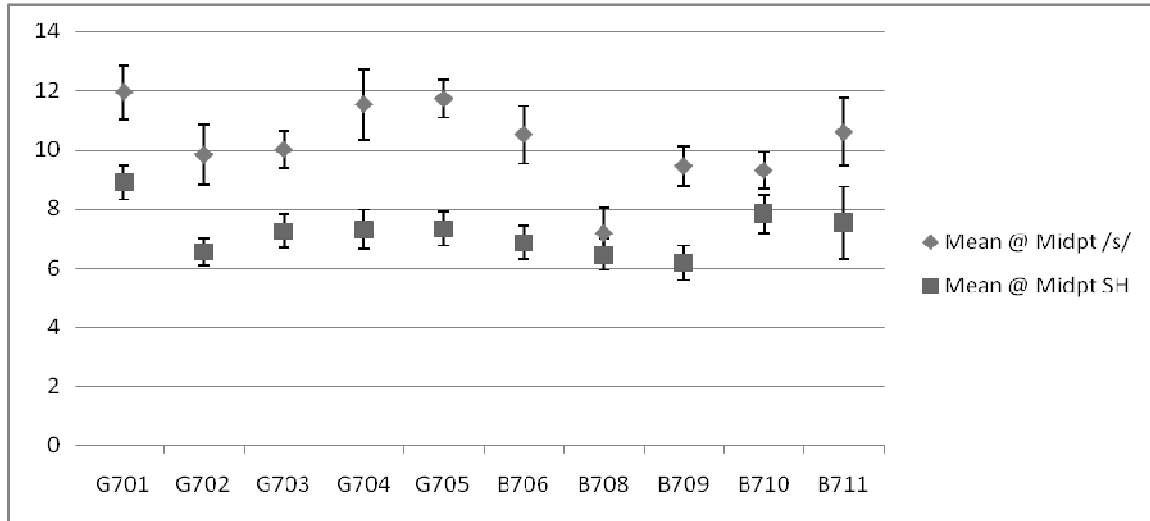
Figure 11. Mean FSM for [s] and [ʃ] showing the magnitude of difference between the two consonants for the TD-11 group. Mean FSM is shown in kHz and the error bars display variability of production for both consonants.
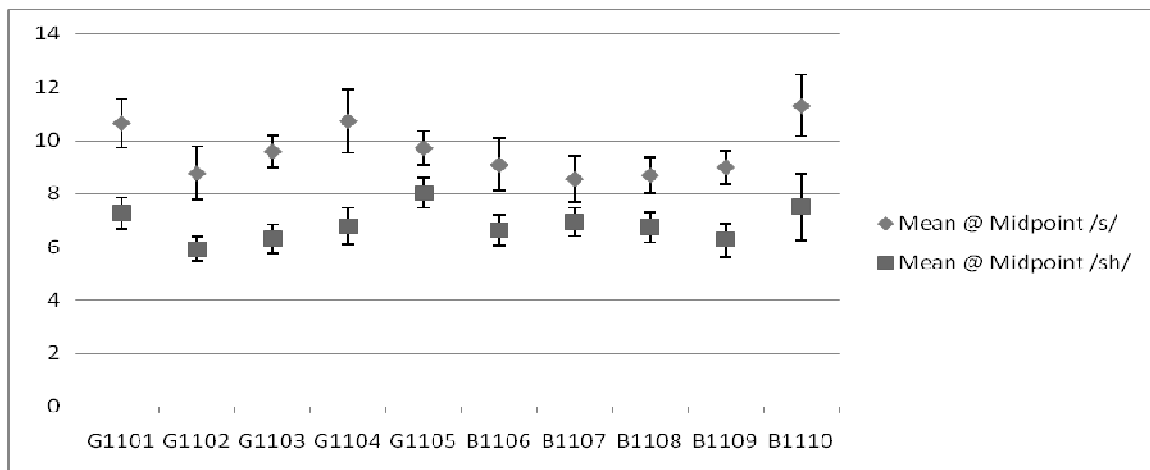
When examining Figures 10 and 11, a possible gender difference in

magnitude of distinction in both age groups was observed. In both age groups,

the female speakers are on the left and the male speakers on the right. Upon

visual inspection of both figures, it appeared that the female speakers made a larger overall magnitude of distinction than male speakers. These differences were not specifically tested statistically due to the small number of subjects in each group.

The COV measure was used to examine differences in variability of production in [s] and [ʃ] within speakers and was averaged across each group. The COV results are displayed in Tables 7 and 8.

Table 7. Coefficient of Variation for production of [s] and [ʃ] at midpoint and across the entire fricative for the TD-7 group. Female participants are represented by G701 – G705, and male participants are represented by B706 – B711. The magnitude of productive distinction in Hz between [s] and [ʃ] is included for comparative purposes.

Subject	COV: [s] at midpoint	COV: [s] overall	COV: [ʃ] at midpoint	COV: [ʃ] overall	Magnitude of Production Distinction
G701	.08	.09	.06	.11	3.0
G702	.10	.13	.07	.09	3.3
G703	.06	.06	.08	.08	2.8
G704	.10	.15	.09	.11	4.2
G705	.05	.07	.08	.08	4.4
B706	.09	.10	.08	.09	3.7
B708	.12	.11	.08	.11	0.7
B709	.07	.10	.09	.12	3.3
B710	.07	.07	.08	.09	1.5
B711	.11	.14	.16	.18	3.1
Average	*.09*	*.10*	*.09*	*.11*	*3.0*

Table 8. Coefficient of Variation for production of [s] and [ʃ] at midpoint and across the entire fricative for the TD-11 group. Female participants are represented by G1101 – G1105, and male participants are represented by B1106 – 1110. The magnitude of productive distinction in Hz between [s] and [ʃ] is included for comparative purposes.

Subject	Accuracy: 2-step	Accuracy: 3-step	COV: 2-step	COV: 3-step	Magnitude of Production Distinction
G1101	.06	.07	.05	.07	3.4
G1102	.09	.08	.13	.13	2.9
G1103	.08	.09	.05	.08	3.3
G1104	.05	.11	.12	.12	4.0

Speaker				
G1105	.05	.07	.10	.11	1.7
B1106	.09	.10	.09	.11	2.5
B1107	.08	.09	.09	.11	1.5
B1108	.04	.05	.07	.08	1.9
B1109	.07	.08	.08	.11	2.7
B1110	.06	.06	.08	.12	3.8
Average	*.07*	*.08*	*.09*	*.10*	*2.8*

The TD-11 group had a smaller COV for [s] at midpoint compared to the TD-7 group and this difference was found to be statistically significant (t = 2.263, p = 0.03). There was not a statistically significant difference between groups for the COV estimated across the entire fricative for [s] (t = 1.948, p = 0.067). In addition, no statistically significant differences were found between the two groups for the COV estimated across the entire fricative for [ʃ] or at [ʃ] midpoint (t = 0.227, *ns* and t = 0.241, *ns*, respectively). Overall, the TD – 11 group demonstrated a smaller COV for both [s] and [ʃ] when estimated across the entire fricative and at midpoint compared to the TD – 7 group.

Developmental Differences in Fricative Duration

The duration for each fricative token ([si], [su], [ʃi], and [ʃu]) for each speaker was estimated to the nearest 10 ms (described in the Methods section), and then averaged across token. The results from this analysis are detailed in Table 9.

Table 9. Mean (SD) duration measures in milliseconds for the tokens [si], [su], [ʃi], [ʃu] for speakers in the TD-7 (G701 – B711) and TD-11 (G1101 – B1110) age groups. Female speakers are indicated by G701 – G705 and G1101 – G1105, and male speakers are indicated by B706 – B711 and B1106 – B1110.

Speaker	[si] (SD)	[su] (SD)	[ʃi] (SD)	[ʃu] SD
G701	160 (31)	166 (22)	163 (23)	166 (20)
G702	137 (16)	149 (29)	139 (13)	142 (18)
G703	240 (27)	224 (22)	237 (17)	243 (26)
G704	232 (47)	226 (39)	223 (36)	228 (39)
G705	202 (27)	224 (22)	193 (21)	201 (22)

B706	200 (20)	214 (22)	211 (26)	221 (19)
B708	172 (24)	178 (36)	171 (33)	174 (36)
B709	195 (33)	200 (33)	189 (27)	182 (34)
B710	152 (23)	150 (18)	148 (26)	148 (22)
B711	325 (107)	272 (46)	258 (60)	280 (92)
TD-7 group	*201 (36)*	*200 (29)*	*193 (28)*	*199 (33)*
G1101	144 (17)	150 (19)	154 (26)	145 (17)
G1102	111 (17)	109 (12)	126 (14)	118 (14)
G1103	134 (19)	126 (18)	137 (19)	133 (19)
G1104	174 (26)	183 (22)	174 (21)	177 (25)
G1105	168 (26)	166 (23)	170 (25)	174 (19)
B1106	192 (21)	192 (19)	183 (23)	187 (18)
B1107	184 (19)	188 (29)	174 (26)	174 (25)
B1108	195 (30)	192 (27)	188 (24)	199 (25)
B1109	133 (18)	144 (27)	140 (29)	138 (24)
B1110	128 (53)	114 (14)	132 (34)	115 (18)
TD-11 group	*156 (25)*	*156 (21)*	*158 (24)*	*156 (20)*

Statistical analyses using student's independent t-test for overall [s] and [ʃ] durations revealed statistically significant differences between age groups. The TD-11 group had shorter fricative durations for both [s] ($t = 2.521$, $p < .05$) and [ʃ] ($t = 2.512$, $p < .05$) compared to the TD-7 group.

Fricative Perception

It was hypothesized that the TD- 11 participants would show qualitatively steeper identification functions and quantitatively higher accuracy on discrimination tasks than the TD-7 participants.

Identification

The results of the identification test by group are shown in Tables 10 and 11. The identification functions for the TD – 7 and TD – 11 subjects, as well as the mean identification functions for the two groups, are plotted in Figures 12 and 13.

98

Table 10. Identification responses of the TD-7 group. The percentage of [s] responses for each step of the continuum is displayed. The percentages represent the mean responses on both [i] and [u] identification tasks for each speaker. Female speakers are G701 – G705, and male speakers are B706 – B711.

	Step 1 - [s]	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7 - [ʃ]
G701	100	95	90	75	5	5	0
G702	100	100	95	95	20	0	0
G703	100	100	90	95	15	0	5
G704	100	100	100	95	40	0	0
G705	95	100	95	100	0	0	0
B706	100	100	100	95	10	0	5
B708	100	100	100	100	35	0	0
B709	100	100	100	100	5	0	0
B710	100	100	95	85	0	0	0
B711	100	95	100	100	10	0	0
MEAN	*99.5*	*99*	*96.5*	*94*	*14*	*0.5*	*1*
(SD)	*(1.58)*	*(2.11)*	*(4.12)*	*(8.10)*	*(13.90)*	*(1.58)*	*(2.11)*
COV	*0.016*	*0.021*	*0.043*	*0.086*	*0.993*	*3.168*	*2.108*

Table 11. Identification responses of the TD-11 group. The percentage of [s] responses for each step of the continuum is displayed. The percentages represent the mean responses on both [i] and [u] identification tasks for each speaker. Female speakers are G1101 – G1105, and male speakers are B1106 - B1110.

	Step 1 - [s]	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7 - [ʃ]
G1101	100	100	100	100	10.5	0	0
G1102	100	100	100	100	0	0	0
G1103	100	100	100	65	0	0	0
G1104	100	100	100	95	0	0	0
G1105	100	95	95	95	45	0	0
B1106	100	100	100	90	5	0	0
B1107	100	100	100	100	10	0	0
B1108	100	100	100	85	0	0	0
B1109	100	100	100	100	0	0	0
B1110	100	100	100	100	50	0	0
MEAN	**100**	**99.44**	**99.44**	**92.22**	**7.83**	**0**	**0**
(SD)	**(0)**	**(1.67)**	**(1.67)**	**(11.49)**	**(14.61)**	**(0)**	**(0)**
COV	**0**	**0.0168**	**0.0168**	**0.125**	**1.865**	**0**	**0**

All participants in both groups showed monotonic crossovers in identification of the [s] - [ʃ] continuum. The TD-11 group was 100% accurate in

identifying the endpoints of the continuum, including step 1 (representing the [s] endpoint of the continuum) and steps 6 and 7 (representing the [ʃ] endpoint of the continuum). The TD – 11 group demonstrated less average variability (COV = .29) compared to the TD – 7 group (COV = .92) with regard to the percent of [s] responses across the entire continuum for both the [i] and [u] vowel. Visual inspection of the graphs in Figures 12 and 13 reveals a qualitatively steeper slope and less variability in responses for the TD-11 group compared to the TD-7 group, particularly at the [s] end of the continuum.

Figure 12. Identification function for TD-7 group. The x-axis depicts the step of the continuum, starting with Step 1 ([s]) and ending with Step 7 ([ʃ]). The y-axis shows the percentage of [s] responses corresponding to each step of the continuum.
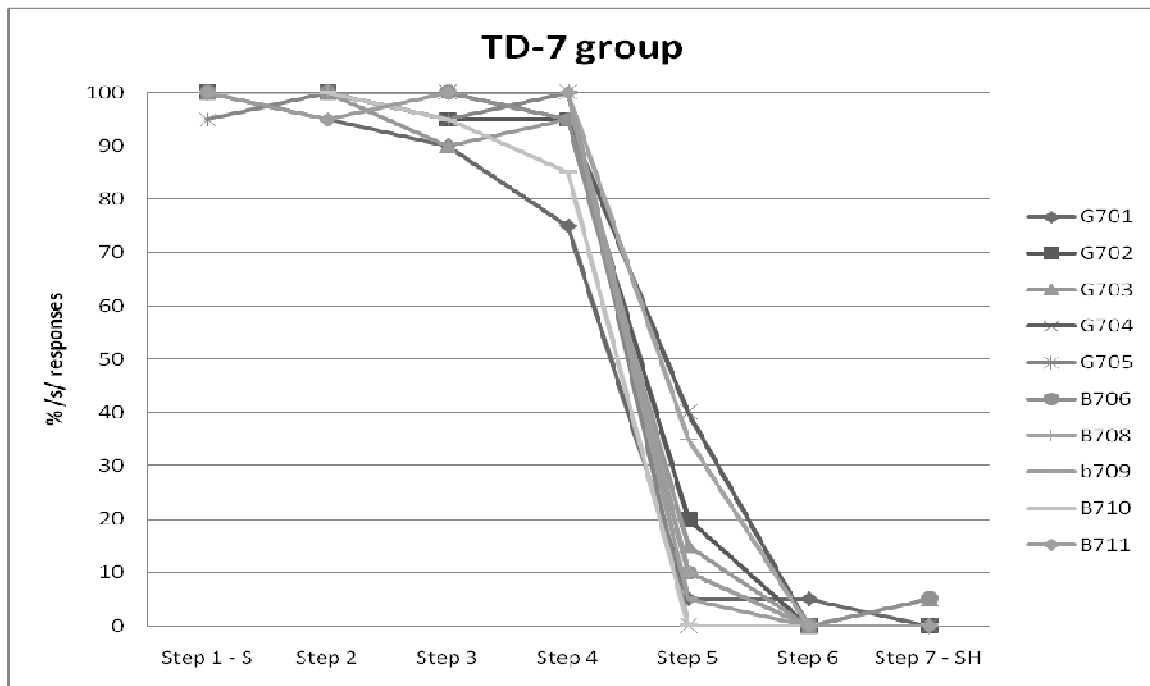
Figure13. Identification function for TD-11 group. The x-axis depicts the step of the continuum, starting with Step 1 ([s]) and ending with Step 7 ([ʃ]). The y-axis shows the percentage of [s] responses corresponding to each step of the continuum.
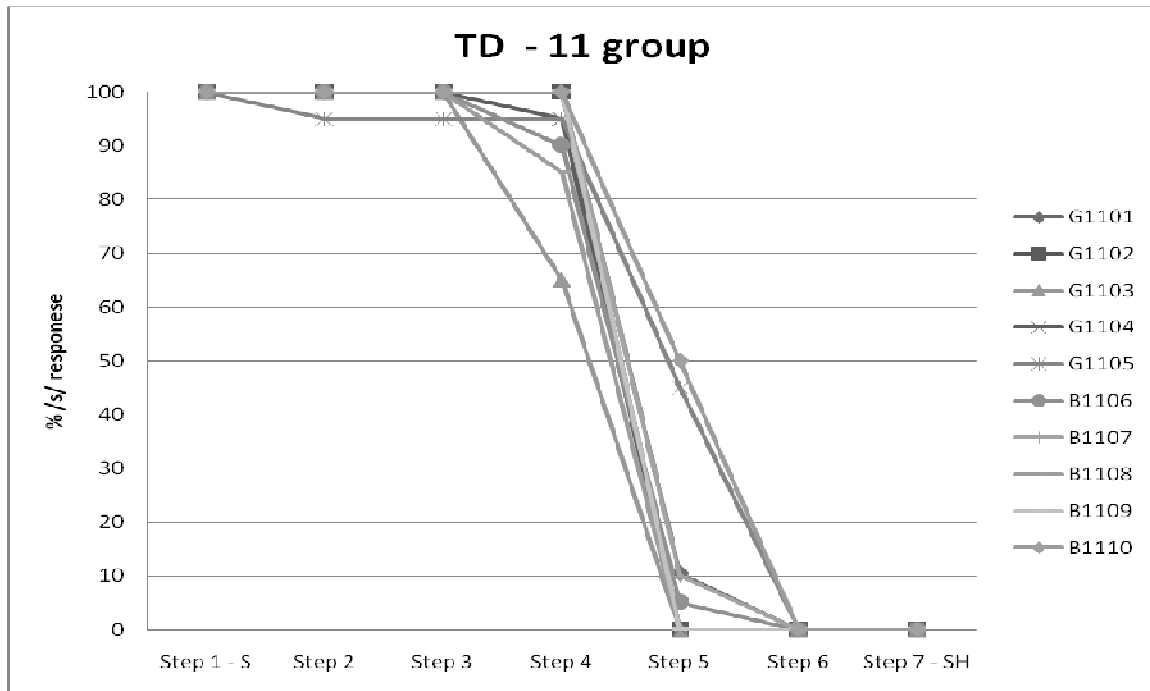
Discrimination

Differences in variability and accuracy of responses on the discrimination trials were found between the two age groups. Individual listener results for both age groups on the discrimination task are detailed in Tables 12 and 13.

Table 12. Discrimination performance for individual participants in the TD-7 group. Female participants are represented by G701 – G705, and male participants are represented by B706 – B711. Mean accuracy (SD) measures are expressed in proportion correct for 2-step and 3-step discrimination tasks. The magnitude of productive distinction in Hz between [s] and [ʃ] is included for comparative purposes.

Subject	Accuracy: 2-step	Accuracy: 3-step	COV: 2-step	COV: 3-step	Magnitude of Production Distinction
G701	.65 (.23)	.52 (.30)	.35	.58	3.0
G702	.59 (.21)	.67 (.13)	.35	.20	3.3
G703	.60 (.22)	.63 (.29)	.37	.46	2.8
G704	.64 (.24)	.72 (.21)	.37	.29	4.2
G705	.69 (.15)	.83 (.16)	.22	.19	4.4
B706	.63 (.19)	.78 (.15)	.30	.19	3.7
B708	.55 (.27)	.67 (.20)	.49	.29	0.7

B709	.66 (.20)	.80 (.12)	.30	.15	3.3
B710	.60 (.17)	.75 (.18)	.29	.23	1.5
B711	.55 (.30)	.52 (.19)	.54	.37	3.1
Average	*.62 (.22)*	*.69 (.19)*	*.36*	*.30*	*3.0*

Table 13. Discrimination performance for individual participants in the TD-11 group. Female participants are represented by G1101 – 1105, and male participants are represented by B1106 – B1110. Mean accuracy (SD) measures are expressed in proportion correct for 2-step and 3-step discrimination tasks.

Subject	Accuracy: 2-step	Accuracy: 3-step	COV: 2-step	COV: 3-step	Magnitude of Production Distinction
G1101	.65 (.23)	.67 (.21)	.35	.32	3.4
G1102	.73 (.17)	.86 (.12)	.23	.14	2.9
G1103	.79 (.16)	.95 (.06)	.21	.06	3.3
G1104	.83 (.19)	.91 (.04)	.23	.04	4.0
G1105	.68 (.18)	.91 (.15)	.27	.17	1.7
B1106	.69 (.21)	.89 (.14)	.30	.16	2.5
B1107	.86 (.08)	.89 (.09)	.09	.11	1.5
B1108	.71 (.24)	.80 (.13)	.33	.16	1.9
B1109	.66 (.21)	.83 (.21)	.32	.26	2.7
B1110	.54 (.19)	.63 (.23)	.36	.36	3.8
Average	*.71 (.19)*	*.83 (.14)*	*.27*	*.18*	*2.8*

The TD – 11 group achieved a greater mean accuracy on both the 2-step and 3-step discrimination tasks than the TD – 7 group. This difference was statistically significant for both 2-step (t = -2.94, p = .009) and 3-step (t = -2.994, p = .008) discrimination trials. The difference between groups on variability in performance, measured by the COV of response accuracy, was also statistically significant with the TD-11 group demonstrating lower COV for both 2-step (t = 2.267, p = .036) and 3-step (t = 2.166, p = .044) discrimination trials. There were also moderate to strong negative correlations between the variability in accuracy on the various comparisons (i.e. 1 vs. 3…5 vs. 7) and the overall accuracy measure, for both age groups and on both 2- and 3-step discrimination trials (TD – 7 group, 2-step: r = -0.719, p = .019, 3-step: r = -0.642, p = .030; TD-11 group,

2-step: r = -0.594, p = 0.070, 3-step: r = -0.777, p = .008). These results indicate

that when participants in either the TD-7 or TD-11 group achieved higher

accuracy scores on the discrimination task, the participants were also less

variable in their responses.

Evidence for Categorical Perception

Evidence for improved discrimination abilities at the phonetic boundary of

the [s] - [ʃ] continuum compared to within phonetic categories was examined in

the TD-7 and TD-11 group. In order to investigate this, the mean results of the 2-

step discrimination task and the average identification function were plotted in

the same figure for each group. This analysis was completed in order to examine

if the participants demonstrated evidence of discriminating better across the

phonetic boundary than within each phonetic category. Figures 14 and 15 display

the results.

Figure 14. Identification function and discrimination accuracy results by comparison graphed simultaneously for the TD-7 group. For the discrimination results, the comparisons are represented by the steps of the continuum for the identification function (2 = 1 vs. 3, 3 = 2 vs. 4, 4 = 3 vs. 5, 5 = 4 vs. 6, 6 = 5 vs. 7).
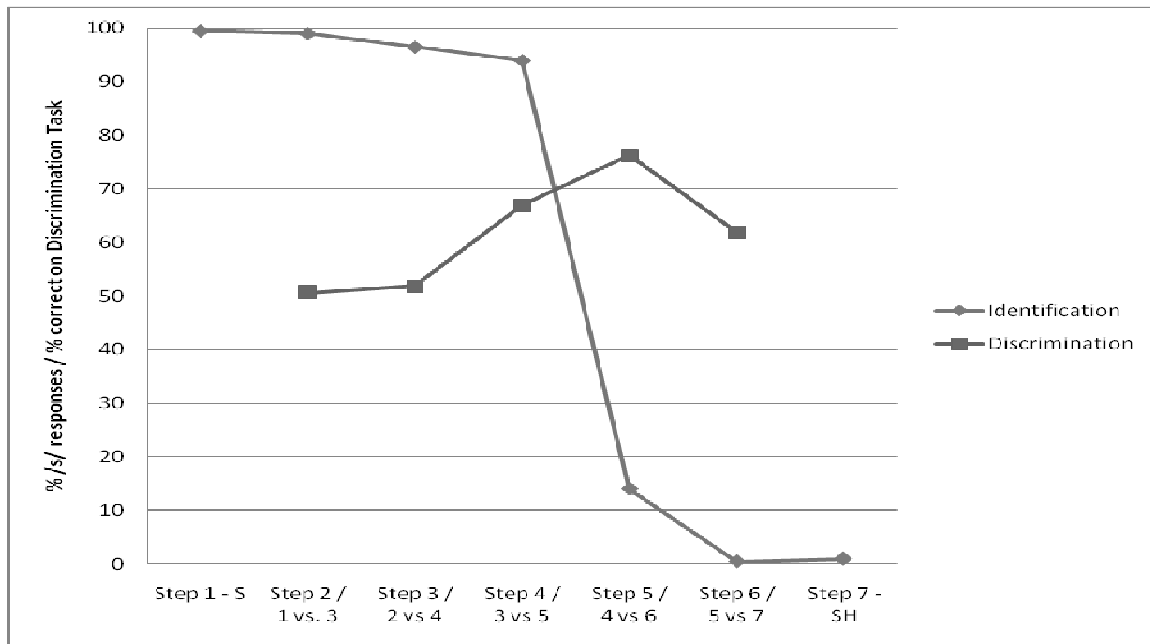
Figure 15. Identification function and discrimination accuracy results by comparison graphed simultaneously for the TD-11 group. For the discrimination results, the comparisons are represented by the steps of the continuum for the identification function (2 = 1 vs. 3, 3 = 2 vs. 4, 4 = 3 vs. 5, 5 = 4 vs. 6, 6 = 5 vs. 7).
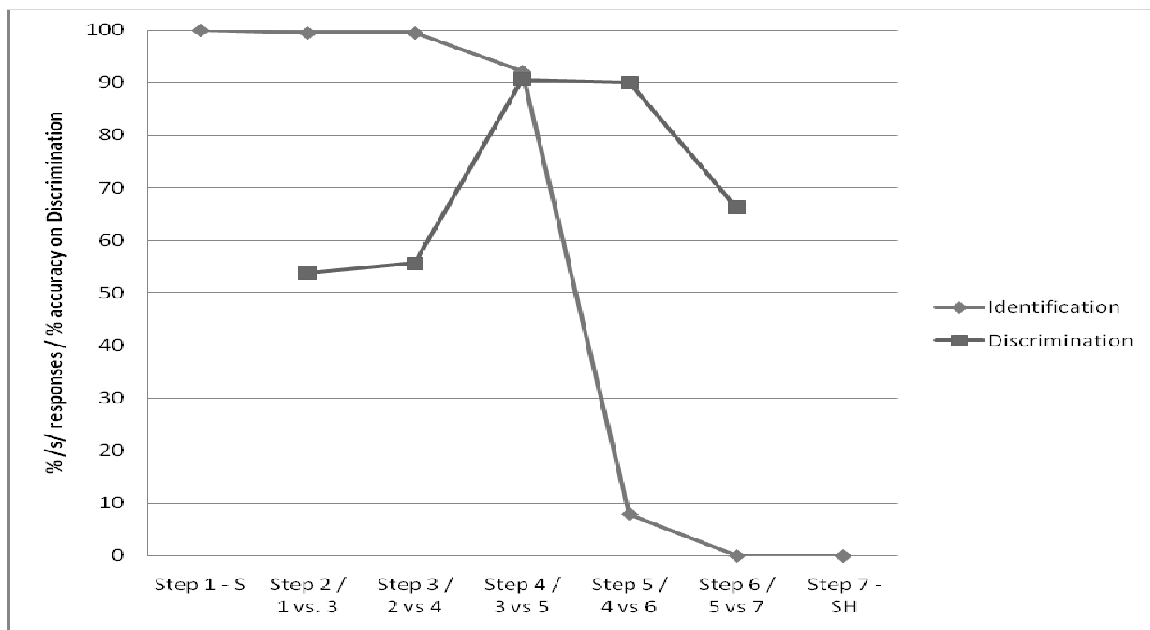
Participants in the TD-7 and TD-11 group both showed the effect of more accurate discrimination at the phonetic boundary than within each category. The TD-11 group showed higher levels of accuracy, however, at the phonetic boundary.

Relationship between Production and Perception

It was hypothesized that typically developing children who have higher discrimination performance will demonstrate greater productive distinction between [s] and [ʃ] than children with lower discrimination performance. The relationship between production and perception was examined using Pearson product moment correlations. Table 14 shows the correlation matrix for this analysis.

Table 14. Correlation matrix showing the correlations between outcomes measures for production (Magnitude of FSM Difference at Midpoint; Overall Difference in FSM) and perception (Mean Accuracy for 2-step Discrimination; Mean Accuracy for 3-step Discrimination).

	FSM Difference at Midpoint	Overall FSM Difference	Mean Accuracy for 2-step Discrimination
Overall FSM Difference	0.972 (p = .000)**		
Mean Accuracy for 2-step Discrimination	0.024 (p = .921)	-0.004 (p = .987)	
Mean Accuracy for 3-step Discrimination	-0.080 (p = .739)	-0.043 (p = .857)	0.757 (p = .000)**

**p < .01

The correlation between perception and production was found to be non-significant, when examining the FSM at midpoint and either the 2-step or 3-step discrimination trials outcome measures and including all typically developing

participants (TD-7 and TD-11 group). As expected, the two production measures and the two perception measures were both strongly correlated to each other.

High and Low Achievers

Because there appeared to be no linear relationship between production and perception in the typically developing children, the results of individual participants were examined to investigate if there were children who demonstrated either high or low achievement on both production and perception measures. Within the TD-7 group, one participant (G705) was identified as a particularly high achiever on both the production measures (largest magnitude of distinction in production) and perception measures (greatest accuracy on both the 2-step and 3-step discrimination trials). The participant in the TD-7 group with the smallest magnitude of distinction (B708) was least accurate (with participant B711) on the 2-step discrimination, but was not the least accurate on the 3-step discrimination task. Therefore, within the TD-7 group, one participant could be identified as a high achiever; however, no participants particularly stood out as a "low" achiever.

Within the TD-11 group, no particularly high or low achieving participants emerged. The participant with the highest levels of accuracy on the 2-step discrimination task (B1107) was not the highest achiever on the 3-step discrimination task, nor did he demonstrate the largest magnitude of distinction of production. The participant with the second largest magnitude of distinction in production (B1110) demonstrated the lowest levels of accuracy on the discrimination task. However, the participant with the largest magnitude of

distinction on the production task (G1104) did have discrimination scores that were among the highest within the TD-11 group.

<center>*Participants with Repaired Cleft Lip and Palate*</center>

Data from the 5 participants with repaired cleft lip and palate were collected in order to compare production and perception of [s] and [ʃ] in these participants to the results from the children in the TD-7 and TD-11 groups. It was hypothesized that children with repaired cleft palate, who demonstrate distorted production of [s] and [ʃ], would demonstrate spectral overlap of fricative targets measured by spectral moment analysis. If production and perception skills co-evolve for these targets, it was predicted that these same children would also show shallower identification curves and less accurate discrimination of [s] versus [ʃ] on perception tasks.

Distinction in Fricative Production for Individual Speakers

In order to evaluate if the participants in the CLP group produced a reliable distinction between [s] and [ʃ], the values for the range of FSM for [s] and [ʃ] at midpoint were examined for overlap. These procedures are similar to those outlined above for the TD-7 and TD-11 group. The mean FSM at midpoint and range of FSM at midpoint, the difference in FSM between [s] and [ʃ] at midpoint, and the percent of non-overlapping distributions for each participant are reported in Table 15. Three of the subjects had values for magnitude of distinction that were almost zero or negative; these distributions were considered to be fully overlapping and the percentage of non-overlapping productions was not calculated.

<center>107</center>

Table 15. FSM (expressed in kHz) at midpoint for the CLP group. Means and ranges are reported for individual speakers (noted by CP101 – 105) in the CLP group, in addition to overall means and ranges. The magnitude difference, [s] - [ʃ], is the difference in mean values for the speaker or speaker group. The percentage of productions that were non-overlapping was estimated based on midpoint values.

Speaker	[ʃ] mean (SD)	[ʃ] Range	[s] mean (SD)	[s] Range	[s]-[ʃ]	% of non-overlapping productions
CP101	8.3 (0.58)	7.0 - 9.4	8.2 (0.82)	6.8 - 10.4	-0.04	Overlapping
CP102	6.8 (0.37)	6.0 - 7.4	7.5 (0.53)	6.1 - 9.6	0.71	82%
CP103	8.4 (0.53)	7.2 - 9.3	10.1 (0.64)	8.8 - 11.5	1.73	93%
CP104	10.0 (0.48)	8.9 - 11.6	10.0 (0.62)	9.2 - 11.8	0.03	Overlapping
CP105	9.2 (0.86)	7.1 – 11.5	8.9 (0.50)	7.4 – 10.1	-0.28	Overlapping
CLP Group	*8.5 (.56)*	*6.0 – 11.6*	*9.0 (.62)*	*6.1 – 11.8*	*0.37*	

Compared to the typically developing children, the children in the CLP group showed a greater proportion of overlapping distributions, with 3 out of 5 children (CP101, CP104, CP105) demonstrating fricative production without spectral distinction, and one participant (CP102) demonstrating more than 10% of productions that were overlapping. Only 1 participant (CP103) showed non-overlapping distributions in at least 90% of productions, compared to 9/10 (90%) of the TD-7 group and 10/10 (100%) of the TD-11 group. Compared to the TD-7 and TD-11 groups, the CLP group showed a higher average spectral mean for [ʃ] and lower average spectral mean for [s], which is reflected in the smaller magnitude of distinction in the CLP group (approximately 0.4 kHz) compared to the TD-7 and TD-11 groups (3.0 Hz and 2.8 kHz, respectively).

For all of the children in the CLP group, fricative production was characterized as "distorted," although the type or degree of distortion varied across speakers. None of the participants produced clear, precise-sounding productions of either fricative. However, individual differences in fricative

production were observed. The information below is a summary of the following: (a) any known history of speech intervention and characteristics, and a description of oral structure; (b) current speech status and an auditory-perceptual description of fricative production; and (c) the participant's FSM results for the production task.

CP101 was the oldest participant in the CLP group, at 11 years of age. He presented with a repaired unilateral (one-sided) cleft lip and palate and had undergone alveolar bone grafting. His anterior bite pattern was the most affected among the participants in the CLP group. He displayed an underbite and missing teeth on the cleft side. CP101 had a longstanding history of phoneme specific nasal air emission on [s] production, although his velopharyngeal function for speech was otherwise normal. On his FSM results, he demonstrated a magnitude of distinction of approximately zero   (-0.04 kHz). As mentioned above, this participant continued to produce inconsistent nasal air emission on [s], which may have affected his mean FSM values for this consonant. Because no research on spectral moment analysis of fricatives in children with repaired cleft palate has been completed, it is unknown how the presence (or absence) of nasal air emission can affect FSM values. This participant demonstrated more variability in [s] production compared to [ʃ], displayed on the time history plots. This is consistent with his variable productions of [s] with or without nasal air emission, compared to his relatively stable production of [ʃ]. Regardless of his specific production characteristics for [s] and [ʃ], both sounds were perceived by the examiner as distorted, likely related to occlusal status and its effect on this

participant's tongue position for these sounds. It should be noted, however, that this participant's [s] and [ʃ] did sound distinct from one another.

CP102 presented with a repaired bilateral cleft lip and palate and alveolar bone grafting. His anterior bite pattern was characterized by an underbite, and one of his central incisors was rotated 90 degrees. His overall speech production patterns were devoid of any overt errors, although his [s] and [ʃ] sounded moderately distorted and imprecise. Qualitatively, his [s] production sounded more accurate than his [ʃ] production, yet the mean FSM for [s] was shifted toward values for [ʃ] production. The FSM results for CP102 indicated a small magnitude of distinction (approximately 700 Hz).   This magnitude of distinction was similar to participant B708 from the TD-7 group, yet CP102 demonstrated 82% of productions that were non-overlapping (compared to 72% by participant B708).

CP103 was an 8-year-old male who presented with a history of repaired bilateral cleft lip and palate and alveolar bone grafting. This participant had the most normal anterior bite pattern compared to the other participants in this group. His bite was considered to be end-to-end, meaning that the incisal edges of his upper and lower approximated each other. He had a history of early enrollment in speech therapy at age 19 months, for approximately 6 months, for compensatory articulation patterns (characterized by backing of bilabial and lingua-alveolar sounds). His speech production skills demonstrated no overt errors, but he did produce a slight distortion of [s] and [ʃ]. The mean FSM values at midpoint for this speaker showed the largest magnitude of distinction compared to the other participants in the CLP group, and within the range of the

children in the typically developing group. It is likely that his relatively normal occlusal status contributed to his ability to produce a measurable degree of distinction between [s] and [ʃ].

Participant CP104 was a 7-year-old male with a history of repaired bilateral cleft lip and palate who had not yet undergone alveolar bone grafting. He presented with a protruding premaxillary segment and his alveolar clefts could be observed. CP104 had a history of a developmental articulation disorder, which was now characterized by vocalic [r] productions, and most relevant to this study, a consistent auditory-perceptual substitution of [s] for [ʃ] productions. His productions of [s] and [ʃ] were merged, with both sounds being produced as [s]. His mean FSM values reflected this pattern of articulation, as the values for [s] and [ʃ] both fell into the range of [s] productions, when compared to the participants in the TD-7 group. In addition, his magnitude of distinction between [s] and [ʃ] was approximately zero, reflecting overlapping fricative distributions. This participant's mother indicated that he was "just beginning to hear the difference" between [s] and [ʃ] in both his own productions and his parent's production of the sounds.

Participant CP105 was a 9-year-old male with a history of repaired unilateral cleft lip and palate and alveolar bone grafting. Despite the alveolar bone graft, this participant presented with a persistent large gap in his gum ridge at the site of the cleft, with permanent teeth yet to erupt. This participant also had a history of speech therapy for approximately 3 years, starting around age 3, in order to remedy compensatory articulation patterns. CP105 demonstrated distorted fricative production, similar to other participants in the CLP group. His

mean FSM values for [s] and [ʃ] demonstrated similar ranges (approximately 7 –

11 kHz for both sounds) and his magnitude of distinction was approximately zero.

Although his productions were not perceived to be merged, the range of values

for each was clearly overlapping. The type of distortion produced by this

participant may have created a large range of values for the FSM for each sound,

and was likely affected by the presence of the large gap in his alveolar ridge.

Qualitative analysis of overlap in fricative production was completed by

visually inspecting time history plots of the FSM for [s] and [ʃ] for each

participant. Time history plots displaying the FSM for [s] and [ʃ] for each

speaker in the CLP group are shown in Figure 16. Only 1 speaker in the CLP

group showed productive distinction (CP103) on the time history plots.

However, similar patterns in production, compared to the TD-7 and TD-11 group,

can be appreciated. For example, increased token-to-token variability on [s]

production compared to [ʃ] production was observed in 2 of the speakers (CP101

and CP102). Speaker CP104 showed an unusual pattern of production for the [ʃ]

fricative, with fricative onset starting at a much lower frequency (in the range of 4

– 6 kHz) and gradually sloping up to approximately 10 kHz by midpoint through

offset. This participant has a history of [s] for [ʃ] substitution, so this pattern of

production could reflect a possible attempt at correct placement for [ʃ] that

becomes more similar to [s] production. The initial portion of some of this

participant's [ʃ] productions sounded noisier compared to his [s] productions.

Overall, similar to what was found in the TD-7 and TD-11 group, individual

variability in fricative production patterns was evident among all 5 speakers.

Figure 16. First spectral moment (FSM) time history plots for CLP group participants'
repeated productions of [s] and [ʃ]. Time histories across the repeated utterances are
superimposed on each other to display categorical distinction visually. The higher
frequency clusters are [s] targets and the lower frequency clusters are [ʃ] targets. Zero
ms on the time axis represents the first analysis window positioned 10 ms into the
fricative segment. The time scale on the x-axis is not the same for each participant.
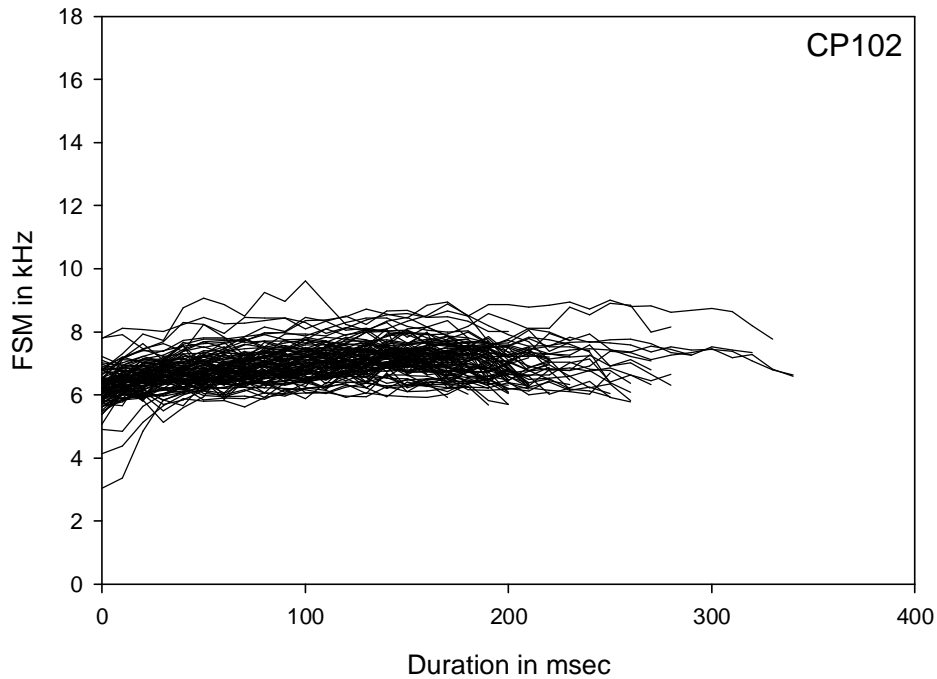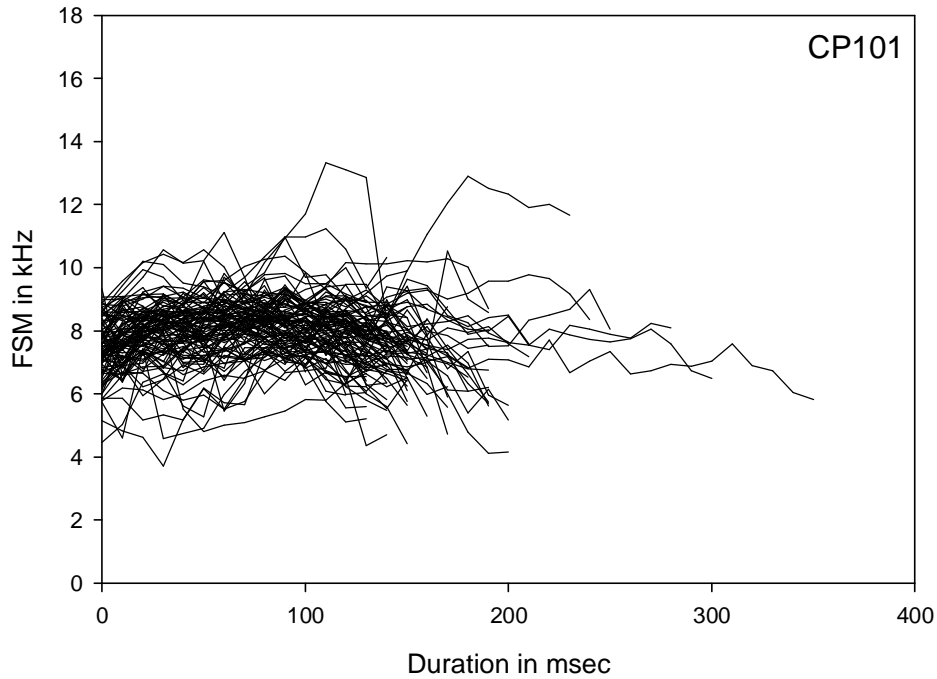
Figure 16, continued.
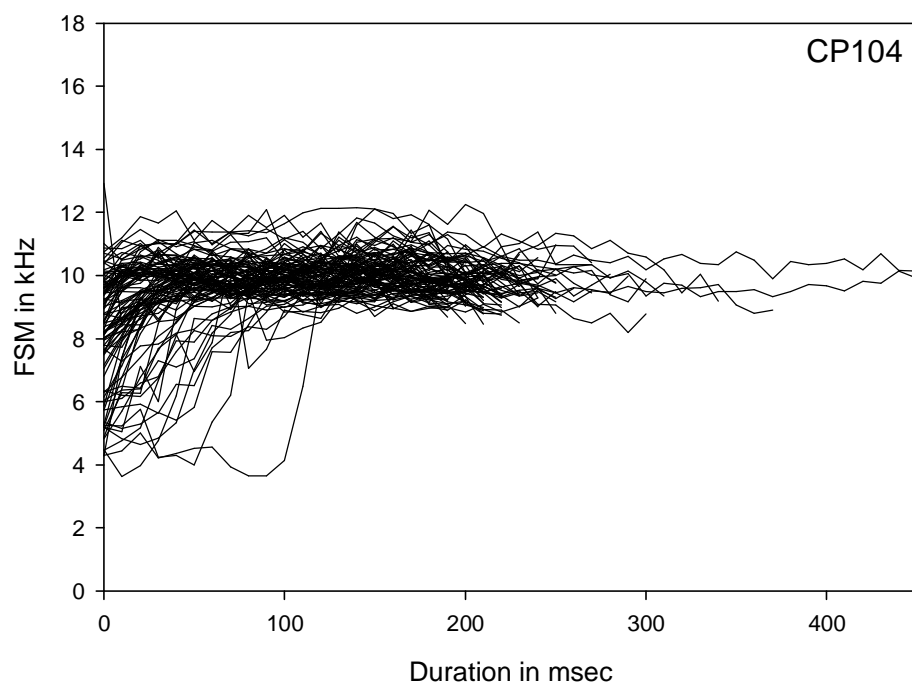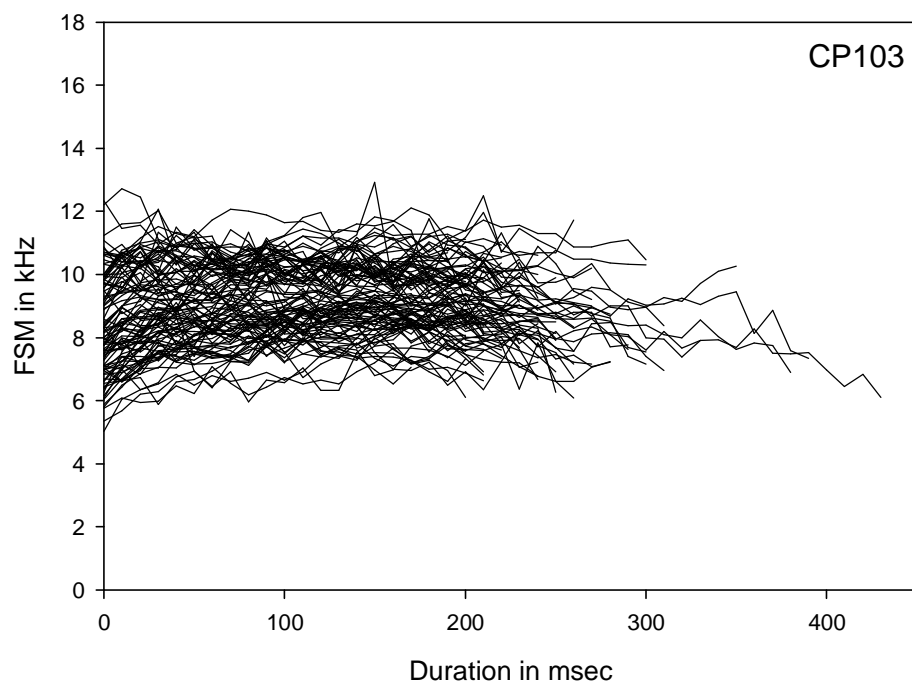
Figure 16, continued.
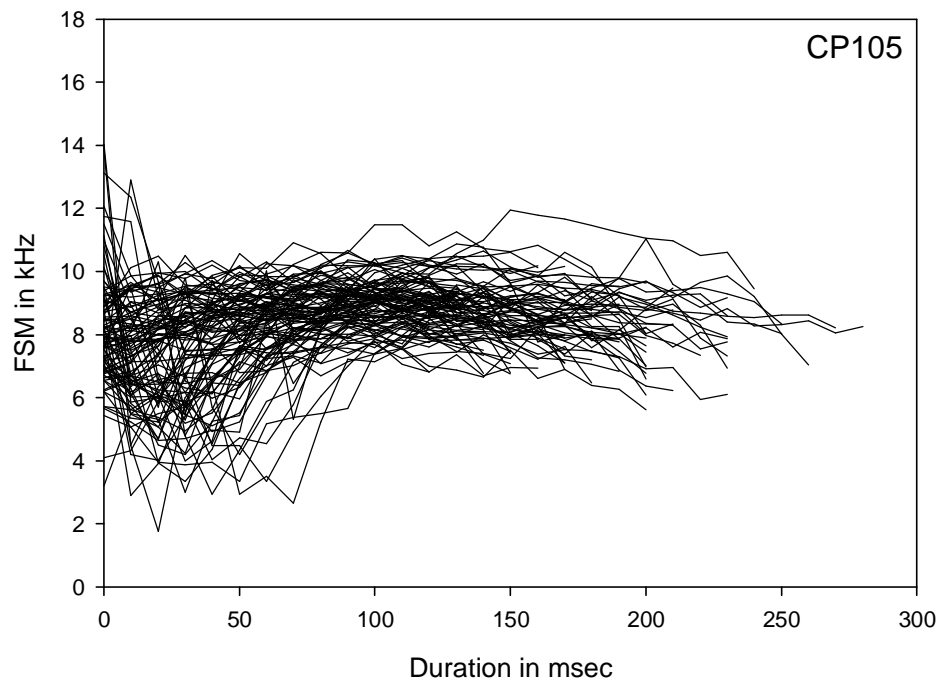
Effect of Vowel Context on Fricative Distinction for Individual Speakers

　　The effect of the vowel [i] or [u] on the magnitude of distinction between the fricatives was also examined for speakers in the CLP group. The time history plots of FSM for [s] and [ʃ] separated by vowel context [i] and [u] are displayed in Figure 17. There did not appear to be an influence of the following vowel context on magnitude of fricative distinction, likely due to the lack of distinction overall between [s] and [ʃ] for the majority of the participants in the CLP group.

Figure 17. First spectral moment (FSM) time history plots for CLP group participants'
repeated productions of [s] and [ʃ] separated by vowel. Time histories across the
repeated utterances are superimposed on each other to display categorical distinction
visually. The higher frequency clusters are [s] targets and the lower frequency clusters
are [ʃ] targets. Zero ms on the time axis represents the first analysis window positioned
10 ms into the fricative segment.
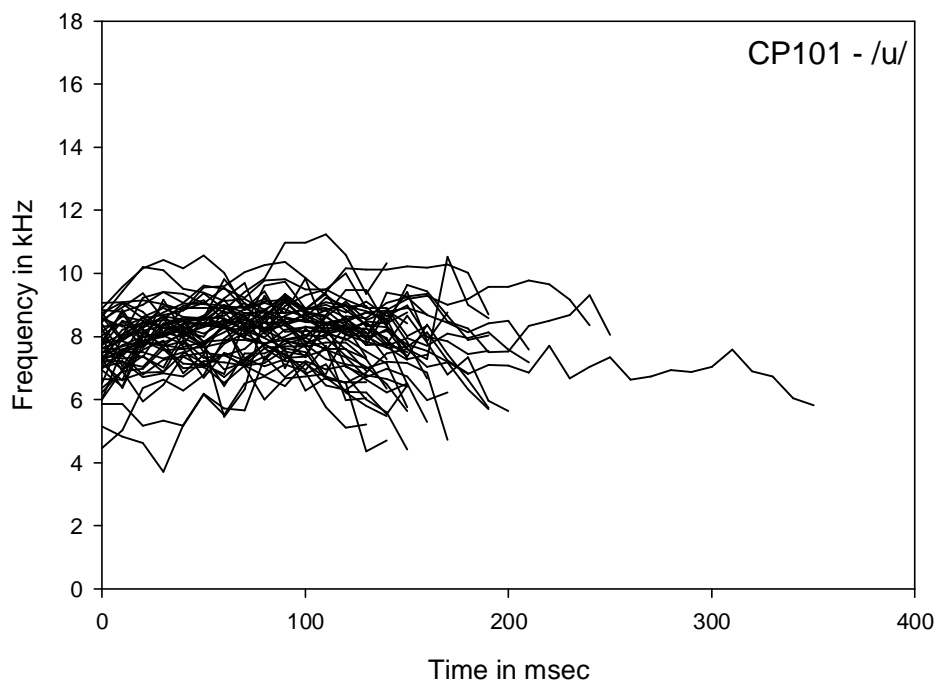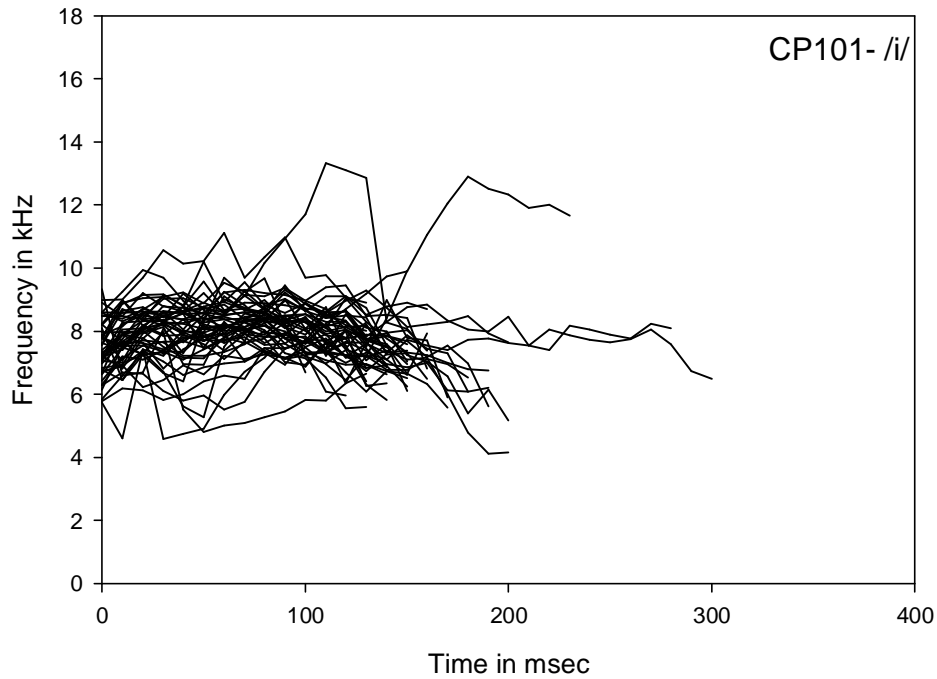
Figure 17, continued.
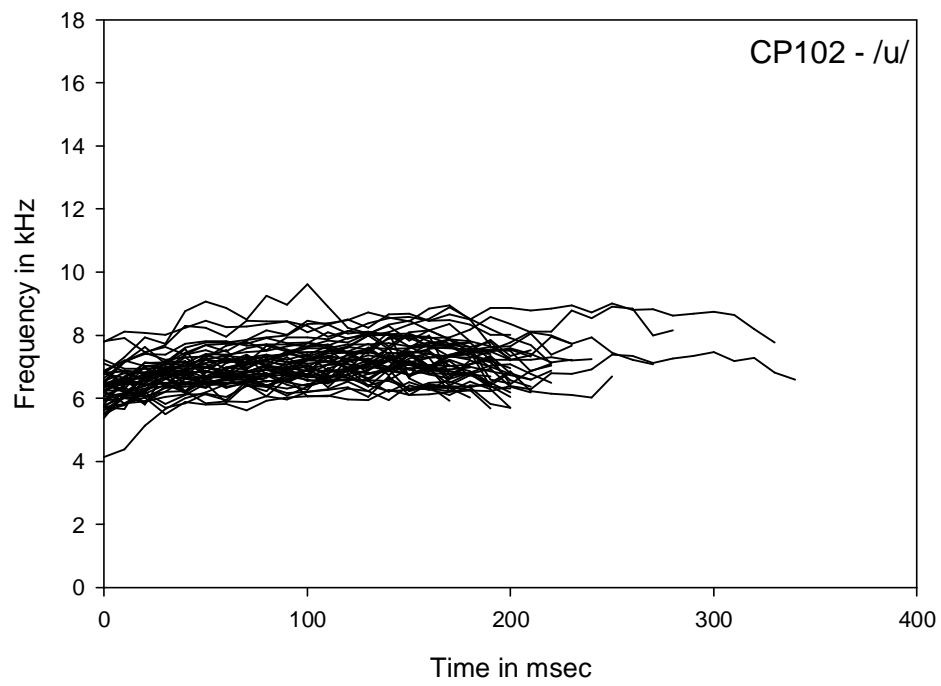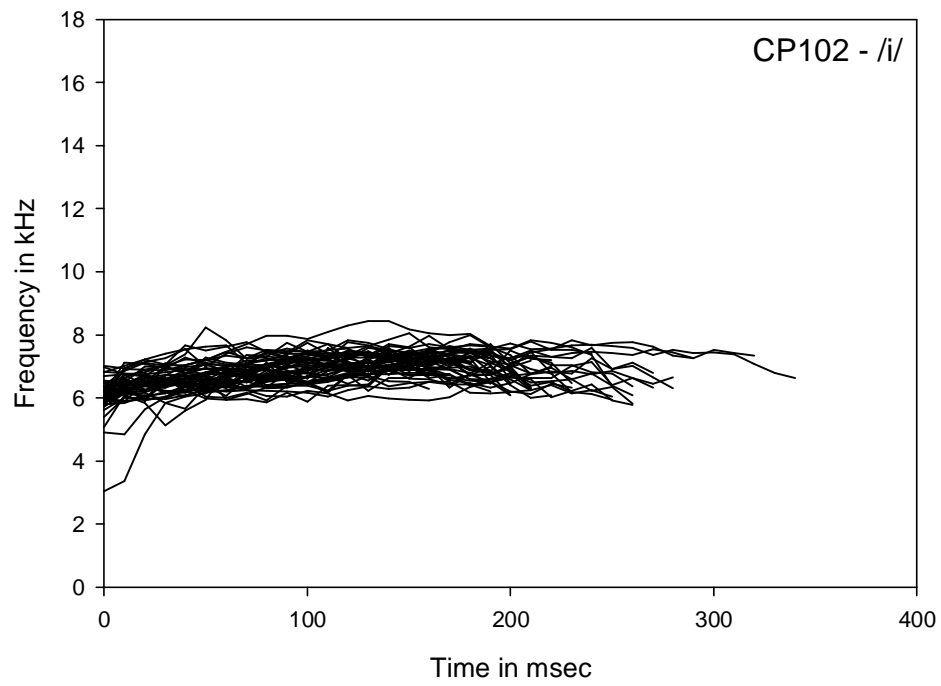
Figure 17, continued.
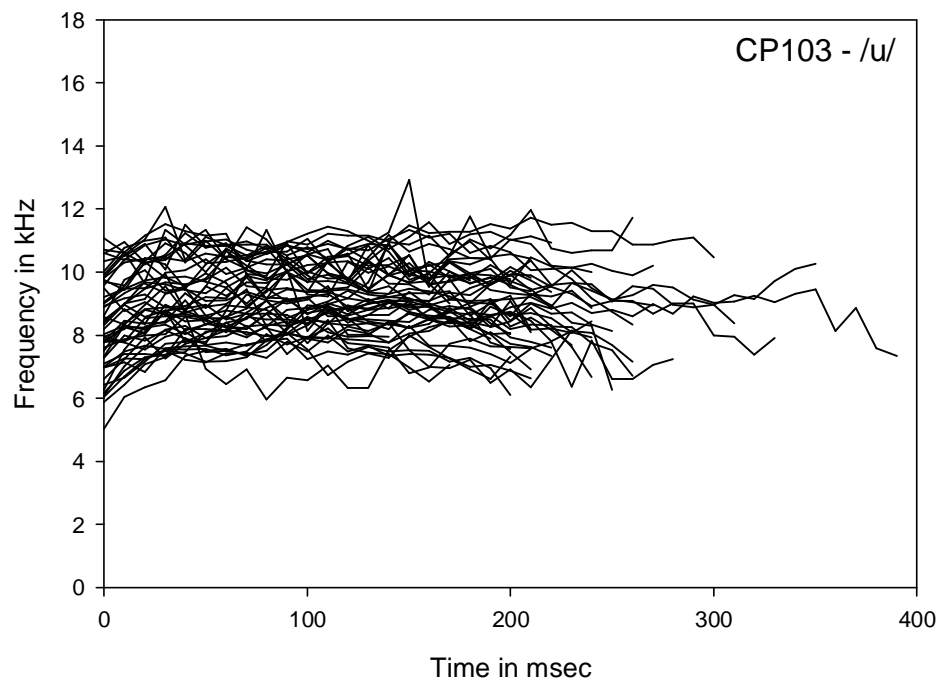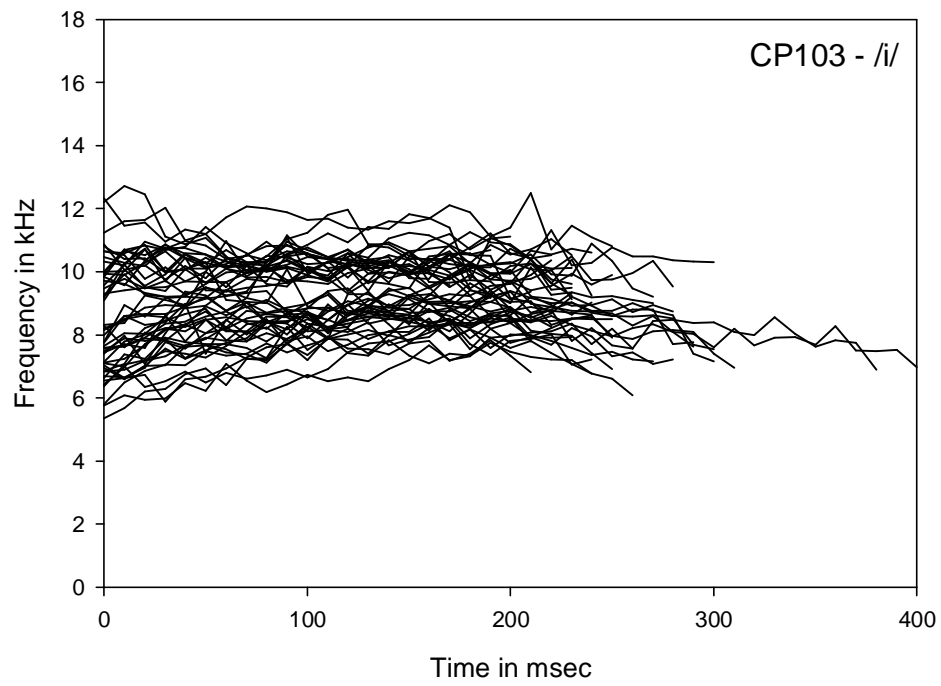
Figure 17, continued.
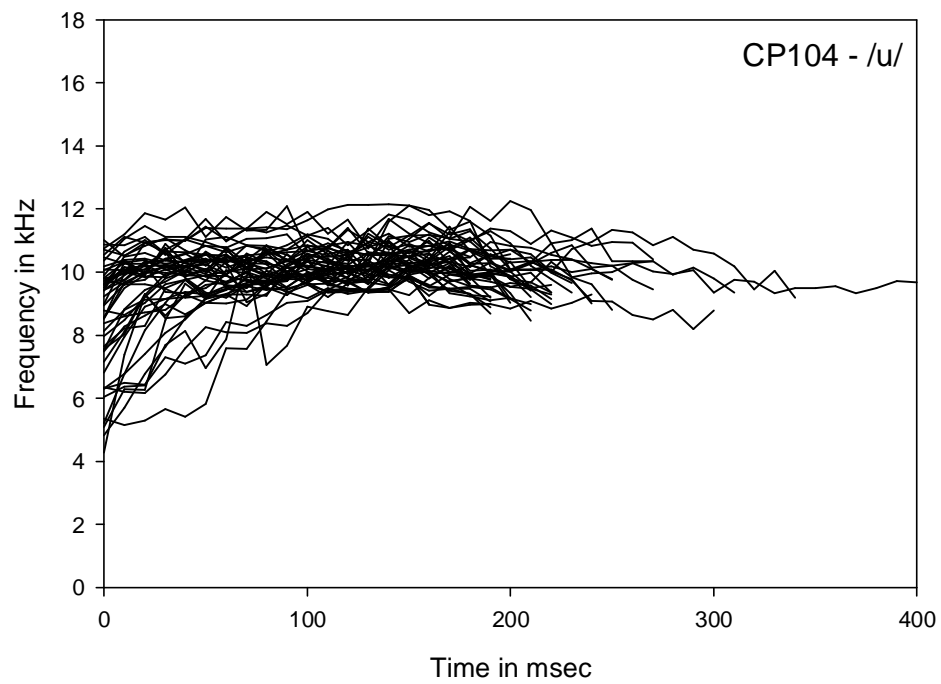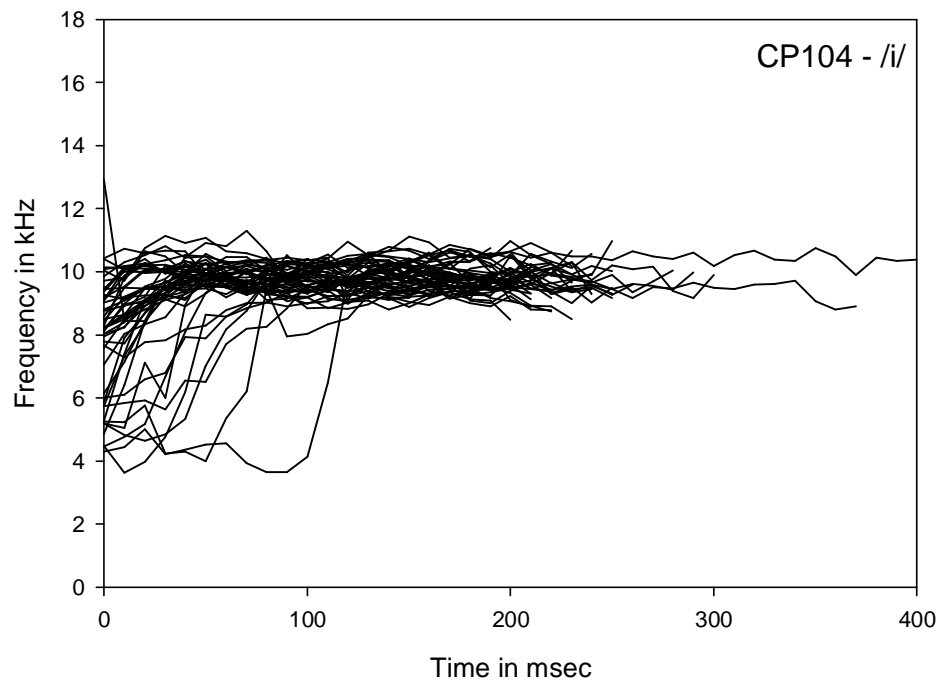
Figure 17, continued.
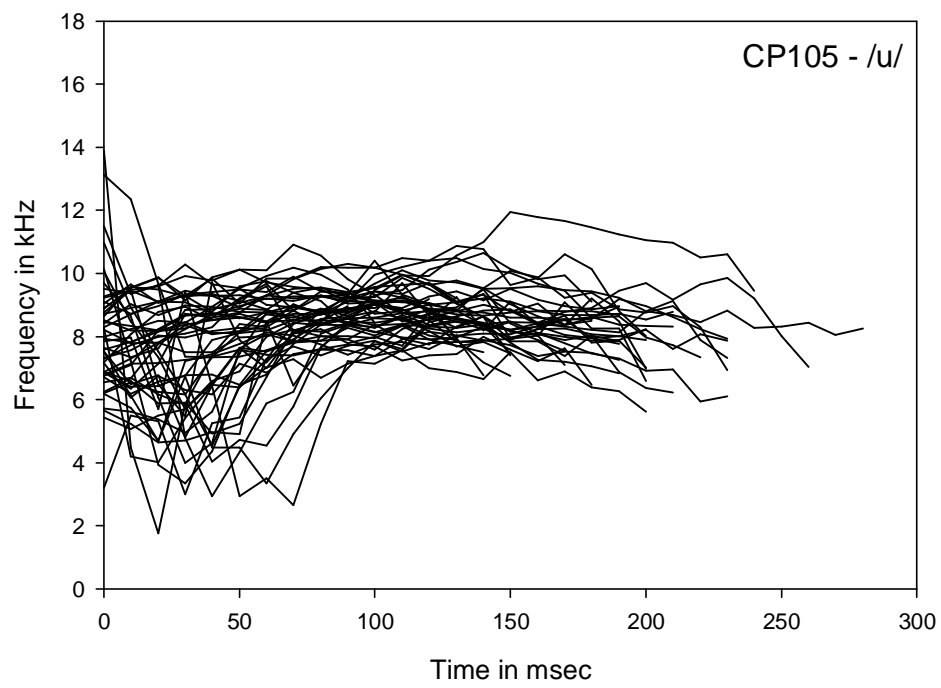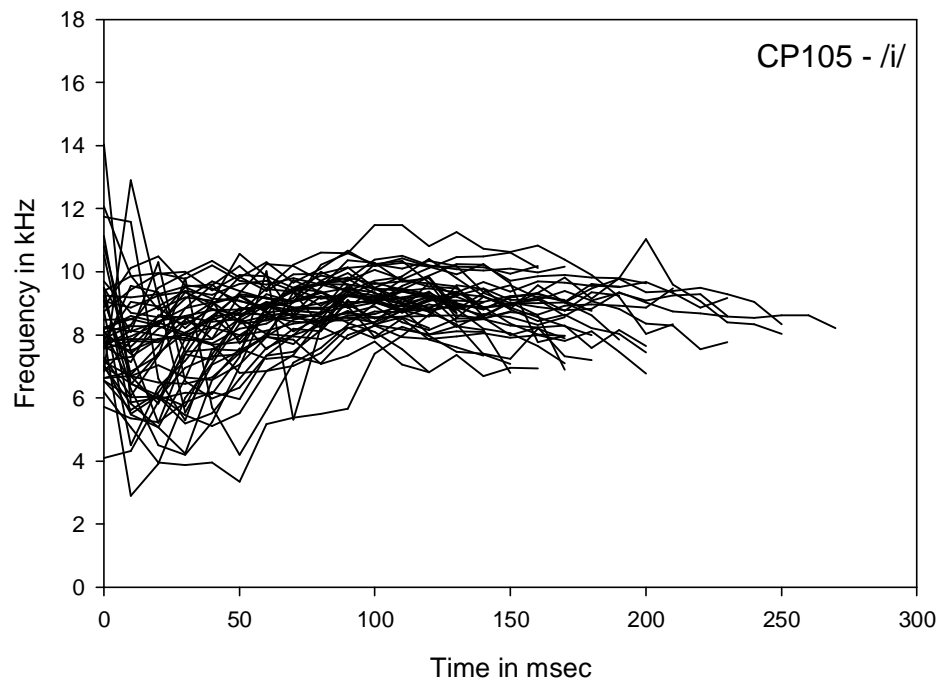
The magnitude of distinction between [s] and [ʃ] is displayed in Figure 18. As described above, only 2 subjects in the CLP group showed a productive distinction between the two fricatives (CP102 & CP103). All of the participants in the CLP group were male, so there was no opportunity for a comparison between male and female speakers.

Figure 18. Mean FSM for [s] and [ʃ] showing the magnitude of difference between the two consonants for the CLP group. Mean FSM is shown in kHz and the error bars display variability of production for both consonants.
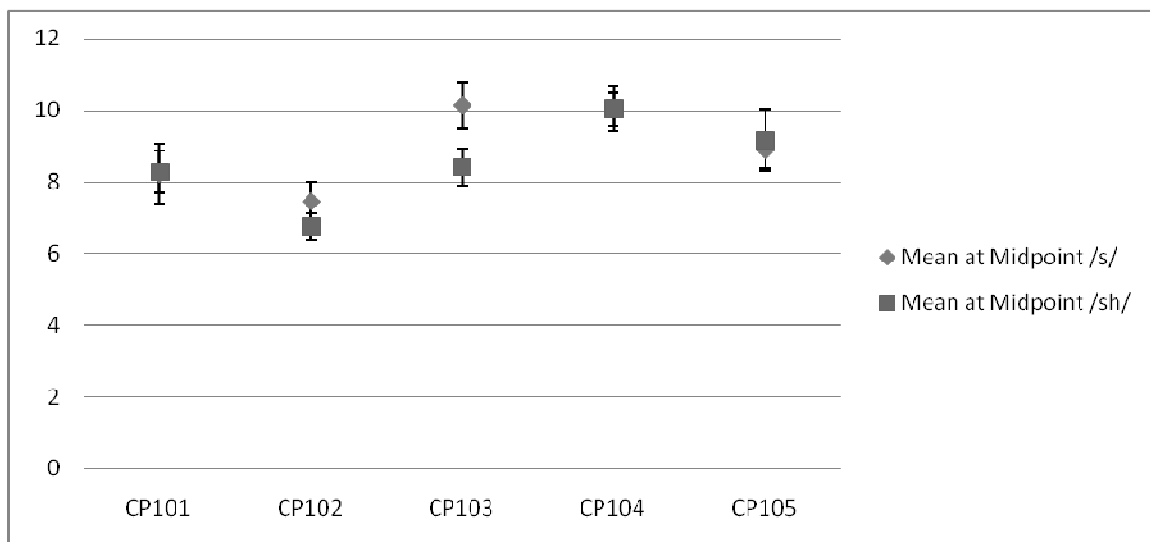
Fricative Duration

The duration for each fricative token ([si], [su], [ʃi], and [ʃu]) for each speaker was estimated to the nearest 10 ms (described in the Methods section), and then averaged across token. The results from this analysis are detailed in Table 16.

Table 16. Mean (SD) duration measures in milliseconds for the tokens [si], [su], [ʃi], [ʃu] for speakers in the CLP group.

Speaker	[si] (SD)	[su] (SD)	[ʃi] (SD)	[ʃu] SD
CP101	167 (34)	183 (49)	169 (42)	164 (24)
CP102	224 (45)	218 (33)	233 (39)	219 (37)
CP103	236 (29)	239 (41)	279 (51)	244 (52)
CP104	228 (25)	232 (98)	244 (66)	238 (42)
CP105	178 (44)	182 (43)	176 (38)	196 (42)
CLP group	*207 (36)*	*211 (53)*	*220 (47)*	*212 (40)*

As a group, the results of fricative duration demonstrate greater mean values compared to the typically developing participants. However, because of the small sample size and age range of the participants in the CP group, direct comparisons between the TD group and CP group cannot be made.

Fricative Perception

Identification. The individual results for the identification testing in the CLP group are shown in Table 17. The identification functions are plotted in Figure 19. In general, all participants in the CLP group showed monotonic crossover (on at least one of the fricative + vowel continua) in identification of the [s] and [ʃ] continua, despite the fact that all of the speakers produced distorted fricative production, and 3 out of 5 speakers did not make an acoustic productive distinction between the sounds. The CLP group did show increased variability in identification of tokens on the [s] end of the continuum when compared to the TD-7 and TD-11 groups. Qualitatively, several of the identification curves had appreciably shallower slopes than the participants in the TD-7 and TD-11 groups.

Table 17. Identification responses of the CLP group. The percentage of [s] responses for each step of the continuum is displayed. The percentages represent the mean responses on both [i] and [u] identification tasks for each speaker.

	Step 1 - [s]	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7 - [ʃ]
CP101	95	100	100	95	5	0	0
CP102	90	85	75	80	25	0	0
CP103	100	100	100	49.5	0	0.5	0
CP104	90	95	80	85	5	0	0
CP105	70	45	55	25	5	0	0
MEAN	*89*	*85*	*82*	*66.9*	*8*	*0.1*	*0*
(SD)	*(11.40)*	*(23.18)*	*(18.90)*	*(28.93)*	*(9.75)*	*(0.22)*	*(0)*
COV	*0.128*	*0.273*	*0.231*	*0.432*	*1.218*	*2.236*	*0*

Figure 19. Identification function for CLP group. The x-axis depicts the step of the continuum, starting with Step 1 ([s]) and ending with Step 7 ([ʃ]). The y-axis shows the percentage of [s] responses corresponding to each step of the continuum.
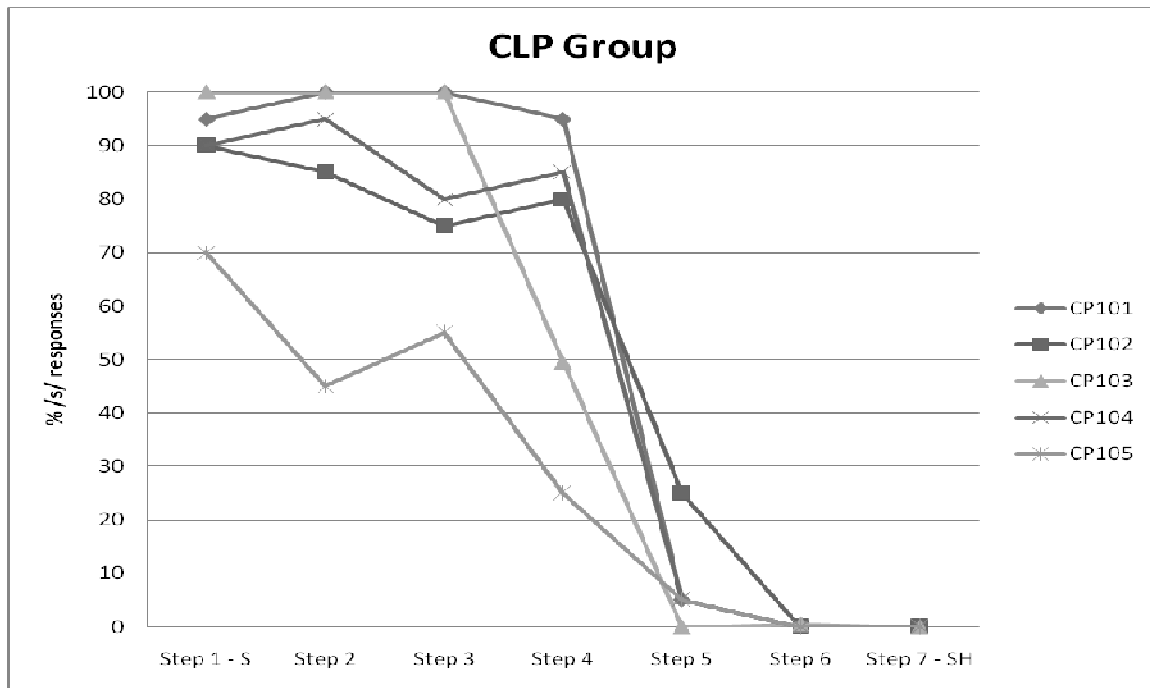
Identification function results from two participants (CP102 and CP105) warrant closer inspection. These two participants showed some difficulty with the fricative identification task, specifically on the [si] - [ʃi] continuum. Figure 20

shows the identification functions for these two participants for this specific

continuum.

Figure 20. Identification functions for participants CP102 and CP105 for the [si] - [ʃi] continuum. The x-axis depicts the step of the continuum, starting with Step 1 ([s]) and ending with Step 7 ([ʃ]). The y-axis shows the percentage of [s] responses corresponding to each step of the continuum.
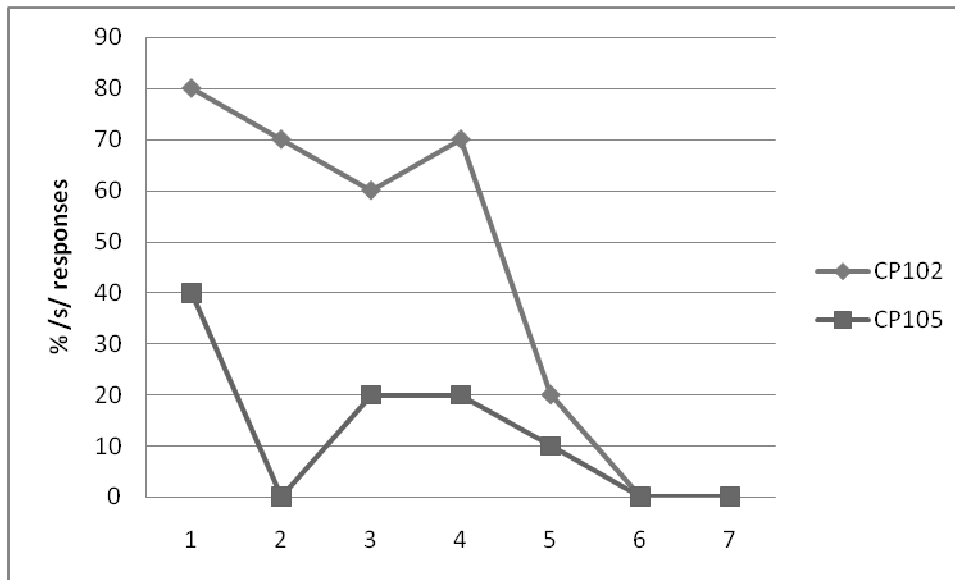
Both of these subjects were able to identify the endpoints of the continuum

with at least 95% accuracy during the perceptual identification screening (CP102

= 95% accuracy, CP105 = 100% accuracy); however, during the actual perceptual

identification task, these subjects showed performance that was particularly

different from the other participants in the CLP group and both the TD-7 and TD-

11 group. CP102 demonstrated some difficulty with the [s] end of the continuum

but not the [ʃ] end of the continuum. During the administration of the

identification task, CP105 stated "these are all 'she'" when completing the

identification task for the [si] - [ʃi] continuum. For either participant CP102 or

CP105, it was possible that the increased difficulty of the task from the screening

to the actual identification task affected their performance. Both of these

subjects had productive distortion of the fricatives [s] and [ʃ], with CP102

producing both fricatives closer to the spectral range for [ʃ] and CP105

demonstrating no productive distinction between the fricatives. For the [su] -

[ʃu] continuum, however, both participants demonstrated monotonic crossovers

similar to the rest of the study participants.

     Discrimination. Individual participant results for the CLP group are

detailed in Table 18. Similar to the TD-7 and TD-11 groups, discrimination

accuracy was higher and variability in responses was lower for older participants

than for younger participants.

Table 18. Discrimination performance for individual participants in CLP group. Mean
accuracy (SD) measures are expressed in proportion correct for 2-step and 3-step
discrimination tasks. The magnitude of productive distinction in kHz between [s] and [ʃ]
is included for comparative purposes.

	Age	Accuracy: 2-step	Accuracy: 3-step	COV: 2-step	COV: 3-step	Magnitude of Production Distinction
CP101	11 years	.78 (.15)	.80 (.17)	.19	.22	-0.04
CP102	9 years	.70 (.19)	.80 (.20)	.27	.25	0.71
CP103	8 years	.56 (.28)	.64 (.22)	.49	.34	1.73
CP104	7 years	.56 (.27)	.53 (.20)	.48	.38	0.03
CP105	9 years	.54 (.25)	.75 (.18)	.47	.24	-0.28
Average		*.62 (.23)*	*.70 (.20)*	*.38*	*.29*	*0.43*

     Evidence of Categorical Perception. While almost all of the participants

in the CLP group showed monotonic crossovers in their identification of the [s] -

[ʃ] continua for both the [i] and [u] vowels, with the exception noted above

(CP105 for the [si] - [ʃi] continuum), evidence for categorical perception was

examined in the CLP group. The mean results of the 2-step discrimination task

and the average identification function were plotted in the same figure. This

analysis was completed in order to examine if the participants demonstrated evidence of discriminating better across the phonetic boundary than within each phonetic category. Figure 21 below displays the results.

Figure 21. Identification function and discrimination accuracy results by comparison graphed simultaneously for the CLP group. For the discrimination results, the comparisons are represented by the steps of the continuum for the identification function (2 = 1 vs. 3, 3 = 2 vs. 4, 4 = 3 vs. 5, 5 = 4 vs. 6, 6 = 5 vs. 7).
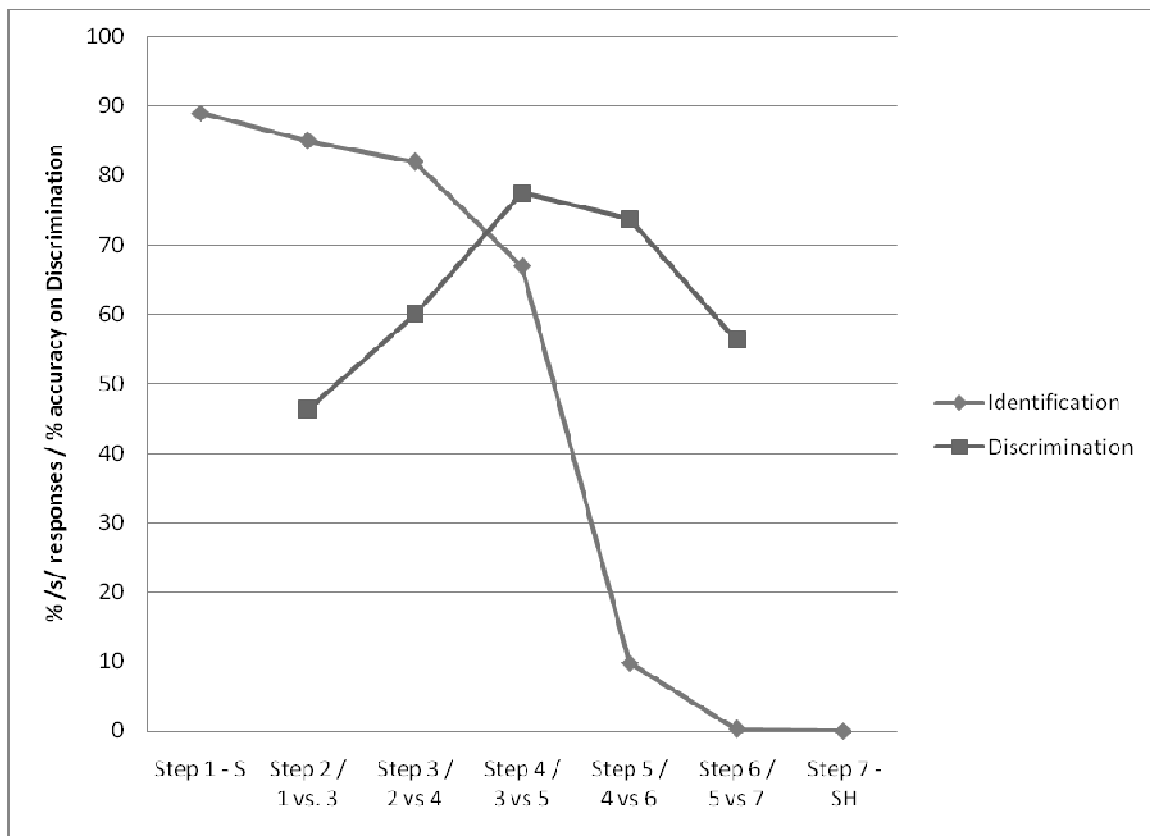
As a group, the participants with repaired cleft lip and palate did show greater difficulty consistently identifying tokens on the [s] end of the continuum as [s] productions compared to the participants in the TD-7 and TD-11 group. However, similar to the typically developing children, the participants in the CLP group also showed greater accuracy in discrimination at the phonetic boundary than within each phonetic category.

Relationship between Production and Perception

Although a formal correlational analysis was not completed due to small sample size and subsequent lack of power, inspection of the magnitude of production distinction and accuracy of 2- or 3-step discrimination tasks does not show a directional relationship. The participant in the CLP group with the largest magnitude of production distinction (CP103) did not show the highest accuracy on discrimination tasks, nor did the speaker with the smallest magnitude of production distinction (CP105) show the lowest discrimination scores.

Summary of Findings

Hypothesis One

Typically developing children age 7 and 11 years of age will demonstrate a reliable distinction between [s] and [ʃ], which would reflect the speech production pattern found in adults. The findings from the analyses of fricative production indicated that most, but not all, typically developing children in this study (95%) showed non-overlapping productive distinction between the voiceless sibilant fricatives [s] and [ʃ], with varying degrees of token-to-token variability and variability in dynamic patterns of fricative production.

Hypothesis Two

The 11-year-old children will produce the fricatives [s] and [ʃ] with greater magnitude of distinction compared to the 7-year-old children, approaching adult-like productions. No developmental differences in production of the fricatives were found between the age groups of 7-year-olds and 11-year-olds. Differences in fricative duration and variability, estimated using the COV, were found between groups. The older children produced fricatives of

shorter duration compared to the younger children, and also produced [s] with less variability at midpoint compared to the younger children.

Hypothesis Three

The older children will demonstrate steeper identification function curves and higher levels of accuracy in fricative discrimination than the younger children. Differences in fricative perception were found between the TD-7 and TD-11 groups, with the older children displaying qualitatively steeper slopes on identification functions, with higher accuracy and decreased variability on tests of fricative discrimination compared to the younger children.

Hypothesis Four

Typically developing children who demonstrate greater distinction between[s] and [ʃ] and less variability in production will have higher discrimination abilities than children with a decreased magnitude of distinction in production. In this study, no relationship was found between fricative production and perception in the typically developing children.

Hypothesis Five

Children with repaired cleft palate, who demonstrate distorted fricative production judged by the examiner, will show spectral overlap measured by spectral moment analysis. We may further predict that these same children will also show shallower identification curves and less accurate discrimination of [s] versus [ʃ] on the perception tasks. Children with repaired cleft lip and palate showed greater proportion of overlapping fricative production, but like the typically developing children, showed individual speaker variability in dynamic spectral patterns during production of [s] and [ʃ]. In general, the participants in

128

the CLP group showed monotonic crossovers in identification of the [s] - [ʃ] continuum despite most speakers showing no productive distinction. Fricative discrimination in the CLP group was similar to the performance of the TD-7 and TD-11 group, with older children in the CLP group demonstrating greater accuracy and lower variability compared to the younger children in this group. Similar to the typically developing children, there did not seem to be a relationship between production and perception of the voiceless sibilant fricatives in the participants with repaired cleft lip and palate.

CHAPTER 5

Discussion

The purpose of this study was to advance the current knowledge base regarding production and perception of the voiceless sibilant fricatives [s] and [ʃ] in two groups of typically developing children, age 7 and 11. Developmental differences in production and perception were investigated, as well as the relationship between production and perception of [s] and [ʃ]. Finally, a group of five children with repaired cleft lip and palate between 7 and 11 years of age were included in the study to investigate if differences exist in perception or production in children with obligatory limitations in early development of speech production and perception skills compared to typically developing children.

Fricative Production

The first two research questions were designed to examine fricative production in typically developing children, ages 7 and 11. For these research questions, the first spectral moment (FSM) was utilized rather than other spectral moments because of its usefulness in providing both qualitative and quantitative information about fricative production, and because it reflects the different articulatory configurations for these two sounds. The FSM for each fricative token (twenty-five repetitions each of [si], [su], [ʃi], and [ʃu]) production was graphed over the course of the fricative for each participant in order to create a time history plot and to examine fricative production of participants

qualitatively. These plots were inspected visually to evaluate the overall distinction in production and gain information regarding inter-speaker and intra-speaker variability in production. Quantitatively, the FSM reflected the spectral mean of fricative production, determined by place of constriction in the oral cavity. The FSM has been shown to distinguish [s] from [ʃ] categorically. In addition, the FSM requires little subjective judgment to compute, is retrieved from the acoustic waveform and spectrogram with minimal effort, and thereby easily accessible for analysis from a quantitative standpoint.

Intra-speaker Distinction between [s] and [ʃ]

The first research question was designed to determine if typically developing children, ages 7 and 11, produced a reliable FSM distinction between the voiceless sibilant fricatives. It was hypothesized that all children would produce a reliable distinction between [s] and [ʃ] on all tokens. In order to answer this question, the time history plots of the FSM for all fricative productions were visually inspected for overlapping distributions, and the FSM at midpoint was evaluated for overlap.

No spectral overlap between [s] and [ʃ] was apparent in the majority of the typically developing participants. This finding was consistent with results from studies by Nittrouer and colleagues (Nittrouer, 1995; Nittrouer et al., 1989) who found a FSM distinction between [s] and [ʃ] in groups of children ages 3 through 7 years of age. Nissen and Fox (2005) also found a FSM distinction when examining the FSM at midpoint for [s] and [ʃ] in younger children (age 5), but contrary to findings of Nittrouer and colleagues, did not find a FSM distinction in children age 3 and 4 years of age. In the present study, individual speakers were

131

found to produce a consistent distinction between [s] and [ʃ] in both the TD-7 and TD-11 age groups. In addition, because production of individual speakers was examined, inter-speaker differences in token-to-token variability and dynamic patterns of fricative production were discovered among the speakers both within and across age groups and gender.

One speaker in the TD-7 group (B708) showed a pattern of overlapping time history plots and produced [s] and [ʃ] with the smallest quantitative distinction of approximately 700 Hz, compared to the other typically developing children in this study. It was curious that this speaker was able to produce [s] and [ʃ] that were judged to be distinct from one another, yet the FSM analyses overlapped for almost 30% of his productions. When listening to this speaker's productions, his [s] productions sounded as if lower frequency noise was present, indicating that this speaker may be making [s] productions with a slightly backed posterior tongue posture, approaching [ʃ]. When inspecting this speaker's mean FSM for [s] and [ʃ], the FSM for [s] is closer to the range of [ʃ] productions for participants in the TD-7 group (refer to Table 6 for frequency information). This participant's small magnitude of spectral distinction between the two fricatives still maintained a perceptual distinction, judged as correct production of [s] and [ʃ] by the examiner, when he produced these sounds. The finding that a typically developing speaker could produce a small FSM magnitude of distinction between [s] and [ʃ] and yet maintain a perceptual distinction between [s] and [ʃ] would have been lost if only group averages were used in analysis. By investigating individual speaker productions, the range of normal variation in magnitude of distinction could be examined. Although other researchers have found that

groups of typically developing children make a consistent spectral distinction between the fricatives [s] and [ʃ], a small distinction may not be an indicator of disordered speech production.

Effect of Vowel Context on Fricative Distinction

The effect of vowel context on fricative differentiation and fricative variability was informally examined through qualitative inspection of the time history plots of [s] and [ʃ] production, separated by vowel. Nittrouer and colleagues (1989) studied fricative-vowel syllable production in children ages 3, 4, 5, and 7 years, and adults. The purpose of this study was to examine how well the speakers produced the [s] and[ʃ] contrastively and how much the following vowel context of [i] and [u] affected this productive contrast. The researchers found that the degree of fricative contrast increased as age increased, and the coarticulatory effect of the vowel [i] and [u] decreased as age increased. In the current study, the coarticulatory effect of the vowel was specific to the fricative produced. For [s] productions, the participants in the TD-7 and TD-11 groups showed similar coarticulatory effects of the following vowel; the [si] tokens were produced with higher mean FSM values than [su] tokens. This finding was expected given the known effect of the lip rounding during articulation of [u] on lengthening of the vocal tract, effectively lowering the resonating frequencies of the fricatives. For [ʃ] productions, the effect of the vowel was only slightly more pronounced in the younger children compared to the older children.

When individual speaker data were examined, different effects of the following vowel on fricative production were observed. Some participants showed greater overlap in fricative distributions on the [u] vowel compared to [i],

some participants showed the opposite effect, and some participants showed no effect of the vowel. These findings were found in participants in both the TD-7 and TD-11 groups. Because individual speakers, in both age groups, were showing different patterns of fricative differentiation in either the [i] or [u] vowel context, making conclusions regarding developmental trends may be precluded.

Individual Speaker Patterns

One observation of individual speaker variation was the performance of speaker B711, also in the TD-7 group. This participant showed the most dynamic changes across the fricative for his tokens compared to all other typically developing speakers, and demonstrated the greatest amount of variability on [s] and [ʃ] (reflected in the coefficient of variation). These findings are also reflected in the time history plots of his fricative productions. Anecdotally, this participant had the most difficulty with the production task. He required approximately 23 re-recordings of productions due to inaccuracy. For example, he required repeated attempts to produce the carrier phrase with the correct token named by the examiner. The difficulty seemed to be with the coordination of repeating the carrier phrase while simultaneously attempting to recall the token spoken by the examiner. The task appeared to tax his system in a way that required the pace of the task to be slowed down, and he spoke in a deliberate, controlled manner in order to maintain accuracy. This was reflected in his fricative durations, which were considerably longer than any other normal participants. This participant was the only participant who was missing a tooth, incidentally. While this variation in oral structure may have had some effect of increasing variability of production of [s] and [ʃ], the separate contributions of the missing tooth and the

observed production difficulties cannot be determined by the methodology of the current study.

Fricative Duration and Variability

Increased duration of speech units and inter- or intra-speaker variability of temporal measures have been implicated as possible indices of speech timing and motor control in child speakers (Kent et al., 1980; Smith, 1978, 1992; Tingley & Allen, 1975; Weismer et al., 1982). Studies have compared the speech of children and adults and found that children demonstrate longer segment durations than adults (Kent et al., 1980; Smith, 1978) and that children demonstrate more variability in temporal measures compared to adults (Kent et al., 1980; Tingley et al., 1975). In addition, Smith and colleagues (Smith, Sugarman, & Long, 1983) found that when children spoke at faster than normal rates, comparable to those of adults, their variability was still greater than that of adults. Greater spectral variability of [s] production has been found in child speakers compared to adults as well (Munson, 2004). Participant B711's greater duration and greater spectral variability in [s] and [ʃ] production may be reflective of less stable motor control than other participants in the TD-7 group, evidenced by his difficulty in completing the demands of the production task. He may have needed to employ strategies such as reducing his rate of speech in order to gain greater motor control and perform accurately on the task.

Developmental Differences in Fricative Production

In order to examine if developmental differences in fricative production existed between the TD-7 and TD-11 groups, the FSM data including average FSM at midpoint, the magnitude of distinction at midpoint, the average overall FSM,

and the coefficient of variation were investigated. It was hypothesized that the TD-11 group would show a larger magnitude of distinction between the two fricatives, indicating continued developmental progress after age 7 years toward adult-like speech production.

In this study, independent sample student's t-tests found no statistically significant differences between groups on the magnitude of the distinction at midpoint. These results were contrary to what had been hypothesized, based on previous studies that older children produced fricatives with greater magnitude of distinction compared to younger children, and that children produced fricatives with less differentiation than adults (McGowan et al., 1988; Nittrouer, 1995; Nittrouer et al., 1989; Nittrouer et al., 1996). It could be assumed given the results of these studies that the magnitude of distinction between [s] and[ʃ] would have continued to increase as children got older, as their speech approaches more adult-like productions. The results of these previous studies, however, only included children up to age 7, compared to adults.

Several different factors could be contributed to an individual child's magnitude of distinction of the fricatives [s] and [ʃ]. For example, significant growth of the vocal tract has been found to occur at approximately the age of the children in the TD-11 group (Fitch & Giedd, 1999). Larger vocal tracts will resonate lower frequencies compared to smaller vocal tracts. Vorperian and Kent (2007), in their review of acoustic data on the development of vowel production, stated that "a basic principle in relating anatomic change to acoustic correlates is that the length of the vocal tract determines the overall pattern of formant frequencies" (p. 1512). The differences in oral and/or pharyngeal cavity size

between the children in the two age groups may have affected the absolute values of the resonating frequencies for one or both of the fricatives produced by the child participants. Fitch & Giedd (1999) used magnetic resonance imaging to study vocal tract morphology objectively of normal humans age 2 – 25 years, using a cross-sectional study design. Results of the study indicated a steady gradual lengthening of the vocal tract throughout the stages of child development, as well as a positive correlation between vocal tract length and body size. In their study, they characterized children as "pre-pubertal" prior to age 10, "peri-pubertal" from approximate ages 10 to 14 ½ years, and "postpubertal" from ages 14 ½ years to 25 years. Significant differences in vocal tract length between the female and male children at the peri-pubertal and post-pubertal stages were found, but no significant differences were found in vocal tract length between the genders at the pre-pubertal stage. Their results indicated that children at age 11 may be entering a significant time of growth and development of the vocal tract, with gender differences beginning to emerge (Fitch et al., 1999). A period of growth, beginning at approximately age 11, would be expected to account for growth in vocal tract size and subsequent lowering of resonant frequencies for fricative production in the TD-11 group.

Differences in fricative duration may or may not have contributed to individual differences in magnitude of distinction between [s] and [ʃ]. The younger participants were found to produce statistically significant longer fricative durations compared to the older participants, when analyzing group means. Greater fricative duration may be directly related to a generally slower rate of speech in general in the younger children. Therefore, it is possible that

younger children may demonstrate greater articulatory excursion and less coarticulation than the older children, contributing to a possible greater magnitude of distinction in the younger children. B711, however, the participant with the longest fricative durations, did not produce the largest magnitude of distinction between the fricatives. Duration has been implicated as an index of speech motor control (Kent et al., 1980; Smith, 1978, 1992; Tingley et al., 1975; Weismer et al., 1982), and may be related more to a child's individual motor skill development in fricative production, rather than having a direct relationship to magnitude of distinction between fricatives. It should be noted that the results of this study finding the TD-7 group producing the fricatives approximately 50 ms duration longer on average than the TD-11 group is a greater difference than what has been found in the literature. For example, in a study by Lee and colleagues (1999), the mean difference in fricative duration between the 7-year-olds and 11-year-olds in the study was 20 ms. In the current study, the small sample size and possible outlying results from participant B711 may have influenced the finding of a larger mean difference in fricative duration between groups. It should be noted that children in the TD-7 group produced fricatives with durations similar to the TD-11 group, and vice versa. Therefore, the statistically significant results of group differences in fricative duration found in the current study should be viewed cautiously.

Variability in fricative production could be another factor in a child's ability to demonstrate a smaller or greater magnitude of distinction between [s] and [ʃ]. As a group, the older children showed a smaller coefficient of variation of [s] production at midpoint compared to the younger children, although this

difference between groups was not present for COV for [ʃ] at midpoint. This

finding of less variability in production in older compared to younger children is

similar to what has been found in other studies (Kent et al., 1980; Lee et al., 1999;

Smith, 1978). Children who produce less token-to-token variability in [s] and [ʃ]

production may or may not produce the fricatives with greater distinction

compared to other children with greater variability.

    Differences in speaking style could also contribute to a child's magnitude

of distinction in fricative production. Each child had varying degrees of attention

and motivation for the tasks in this study protocol. Children in the younger age

group appeared more motivated to please the examiner and "perform," while the

older children tended to be more extrinsically motivated, by impressing the

examiner with the speed at which they could finish the task, or appearing excited

for the possibility of being paid for participation. Another stylistic difference may

have been that the older children may have not had to try as hard to be accurate

at the production task, so with a lack of difficulty came a lack of effort, and

therefore possibly a decrease in magnitude of distinction between [s] and [ʃ].

Again, these stylistic differences are purely suggestive of possible individual

differences in child participation that were casually observed during the course of

data collection.

Summary of Findings – Fricative Production

    Evidence in the current literature base indicates that refinement of speech

production skills may be occurring into adolescent years (Lee et al., 1999).   In the

present study, all of the typically developing child speakers, in both age groups,

produced a measurable difference between [s] and [ʃ] using the FSM, although

the range of the magnitude of distinction varied widely. The majority of these children (95%) produced the sounds with very little spectral overlap in production, similar to what has been found for normal adult speakers (Hagle, 2002; Haley, 2002; Haley et al., 2000). Some evidence for effects of vowel context on fricative production was found, although the effect varied across individual speakers and within age groups. No statistically significant differences in magnitude of fricative distinction were found between the younger and older children. This is contrary to findings from other studies, which have investigated fricative production in children age 7 and younger, and found a trend of increasing distinction between [s] and [ʃ] with age. The lack of a difference in fricative distinction between the 7 and 11 year olds in this study may indicate that development of fricative production has possibly reached a plateau somewhere between these ages. However, similar to the current literature, results of this study found developmental differences in fricative duration and [s] variability between the two age groups, with the younger children demonstrating longer fricative duration and greater spectral variability compared to the older children. The evidence that significant vocal tract growth may be occurring at approximately 10 to 11 years of age could be contributing to changes in resonating frequencies of fricatives in the older children while motor skill for fricative production continues to improve. Finally, gender and/or stylistic differences in speaking style may have contributed to some of the variability that were observed among individual speakers.

Fricative Perception

The third research question was designed to investigate if developmental differences existed in fricative perception between two groups of typically developing children age 7 and 11. To this end, both perceptual identification and discrimination tasks were used for data analysis and comparison, in order to examine both phonetic category boundaries and discrimination abilities within and between phonetic categories in typically developing children.

Fricative Identification and Discrimination

All of the typically developing children in this study showed monotonic crossovers in fricative identification, with the TD-7 group showing slightly more variability in [s] responses than then TD-11 group. The TD-11 group showed greater accuracy and less variability in discrimination responses compared to the TD-7 group; these group differences were found to be statistically significant. Previous studies have documented developmental shifts in fricative perception, related to perception becoming more segmental in nature as children grow older, and the developmental shifts occur at approximately 8 years of age (Nittrouer, 1992, 2002; Nittrouer et al., 1993; Nittrouer & Miller, 1997; Nittrouer et al., 1987). The children in the TD-11 group may have been able to access additional cues within the fricative + vowel syllables to make decisions on the discrimination task, above and beyond what the younger children may have used, in order to be more accurate in their responses. These results are in line with the hypothesis of the Developmental Weighting Shift (DWS), which suggests that children attend to different aspects of the speech signal as experience with their native language increases (Nittrouer, 2002). Therefore, the group differences in

discrimination are not surprising, due to the fact that the older children have approximately 4 more years of language exposure and experience than the younger children. Evidence for diminished linguistic experience having a negative effect on the development of adult-like perceptual skill has been found in children with a history of low socioeconomic status with or without a history of chronic otitis media (Nittrouer, 1996). In addition, evidence for ongoing maturation of phonemic categorization in children ages 6 through 12 years was found by Hazan & Barrett (2000), in their study investigating identification of various phonemic contrasts (including the [s] - [ʃ] contrast) in children and adult listeners. The authors discovered that children in the age range of 6 to 12 years continued to develop phonemic categorization skills, measured by steepness of identification functions and calculation of the phoneme boundaries for the contrasts studied. The oldest children, age 12, however, were still not as consistent in responses as adult listeners.

In order to determine whether the participants in the TD-11 group were performing at adult-like levels in perceptual skill, the results of the Perkell et al. (2004) study, which included adult listeners, could provide some basis for comparison. The adults in the Perkell et al. (2004b) study were reported to demonstrate monotonic crossovers in their identification of a [s] - [ʃ] continuum, although specific data for the identification functions was not provided. In the present study, the participants in the TD-7 and TD-11 group performed similarly on identification testing. In addition, the range of proportion correct on the discrimination task in the current study was similar to the results of the TD-11 group. In the Perkell et al. (2004b) study, the adults achieved proportion correct

of 0.6 – 1.0 on a 2-step discrimination task, and in the present study, the participants in the TD-11 group achieved a range of 0.54 - 0.83 on a relatively similar task.

Another methodological difference between the Perkell et al. (2004b) study and the present study was the type of stimuli used for tests of fricative discrimination skill. The stimuli used for the Perkell et al. (2004b) study were described as hybrid stimuli, with synthetic fricative noise combined with natural vowel portions recorded from male and female speakers. The authors did not specify characteristics of the formant transition between the end of the frication and the start of the vowel portion, only stating that "each of the resulting sibilant portions was concatenated with the same naturally produced vowel and final [d] portion" (p. 1263). Although the methodology in the Perkell et al. (2004b) study and the present study were similar in the type of discrimination task used, the stimuli utilized for the tasks may have contained different information useful for listeners to make decisions in the discrimination task. In addition, Perkell and colleagues (2004b) indicated that some of the adult listeners who achieved 100% accuracy on the discrimination task may have had perceptual skills above and beyond what was able to be tested using this specific task. It is unclear whether the participants in the TD-11 group were not performing at adult levels on the discrimination task, or if the dissimilarity in stimuli between the Perkell et al. (2004b) study and this study could be attributed to the difference in discrimination task scores. With the lack of an adult control group in this study, direct comparison between child and adult listeners cannot be completed.

One aspect of categorical perception that both the TD-7 and TD-11 group demonstrated was that for fricative discrimination, participants in both groups displayed higher discrimination scores at the phonemic boundary between [s] and [ʃ] than at the ends of the continuum (shown in Figures 14 and 15). This psychophysical phenomena has been well documented for categorical perception of various phonetic contrasts (see Guenther, Husain, Cohen, & Shinn-Cunningham, 1999, for a review of the literature with regard to this effect). The developmental differences in discrimination accuracy are apparent when inspecting Figures 14 and 15, with the TD-11 group achieving greater accuracy on all comparisons tested compared to the participants in the TD-7 group.

Summary of Findings - Fricative Perception

The results from this study found evidence of developmental differences in fricative perception between the two groups of typically developing children, ages 7 and 11. The TD-7 and TD-11 group showed similar performance on the identification task; however, the TD-11 group demonstrated higher levels of accuracy on the discrimination task, with less variability, as compared to the TD-7 group. Because of the lack of an adult control group in this study, it is unknown whether or not the participants in the TD-11 group may have been demonstrating adult-like perceptual skill on the discrimination task.

Relationship between Production and Perception

The fourth research question was designed to investigate if a linear relationship existed between typically developing children's production and perception of [s] and [ʃ]. One reason this relationship is important to explore is because what can be learned about typically developing children can be applied to

144

the assessment and treatment of children with speech sound disorders. Studies investigating this relationship in children have mainly focused on children with speech disorders (Estrem & Broen, 1989; Hoffman et al., 1985; Rvachew et al., 1989), Perkell and colleagues (Perkell, Guenther et al., 2004; Perkell, Matthies et al., 2004) found a straightforward relationship between adults' ability to discriminate phonetic contrasts for both vowels and the fricative contrast [s] and [ʃ] and the magnitude of productive distinction for these contrasts. The methodology of the study by Perkell et al. (2004) served as a framework for this investigation of the relationship between production and perception of [s] and [ʃ] in typically developing children. Similar to the Perkell et al. (2004b) study, both a productive measure of distinction (FSM at midpoint) and perceptual discrimination tasks were used in order to investigate the relationship between fricative production and perception in two groups of typically developing children, age 7 and 11.

Contrary to what was hypothesized, the typically developing participants' production and perception measures showed no significant linear relationship. There are several possible explanations for why this relationship was not present in this sample of children. First, methodological differences existed between this study and the Perkell et al. (2004b) study. For example, in the Perkell et al. (2004b) study, the researchers separated the adult participants into high and low achievement groups, based on their discrimination performance. Then, the group means for the production measure were compared between the two groups. Therefore, it is unclear whether all of the speakers in the high achievement group all showed the largest magnitude of distinction in the

production measure, on an individual basis, compared to the speakers in the low achievement group. In addition, the stimuli used in the Perkell et al. (2004b) study were not fully described, as mentioned previously. The information contained in these stimuli, in terms of cues available for fricative discrimination, compared to the stimuli used in this study, are not transparent. It could be that this information in the hybrid fricative + vowel stimuli could be used specifically by the listeners in the Perkell et al. (2004b) that were able to achieve 100% accuracy on the discrimination task, and may be directly related to their productive distinction.

In the current study, individual participant data were used to examine the relationship between production and perception. When examining both the discrimination performance and magnitude of production distinction values of the participants, no evidence for true "high" or "low" achievers emerged. Only one participant, in the TD-7 group (G705), produced both the largest magnitude of distinction and achieved the highest discrimination scores on both the 2- and 3-step trials (within the TD-7 group). No subject in the TD-11 group displayed this type of pattern; in fact, the participant with the highest 2-step discrimination results in the TD-11 group (B1107) demonstrated the smallest magnitude of production distinction. In order to rule out whether stylistic production differences were responsible for the small magnitude of production distinction, this participant agreed to re-record a portion of the production task, and was given specific instructions to speak "as clear as possible" and to emphasize the fricative production in the token words.

This participant was brought back approximately 2 months after completing the initial protocol and instructed to produce the tokens "as clear as possible" and to "emphasize" the fricative in the token word. He was only required to produce approximately half of the productions (48) compared to the original production task. A comparison of time history plots for this participant is provided in Appendix 9. Quantitatively, the magnitude of distinction at midpoint increased approximately 1000 Hz, and this difference in distinction can be attributed to mean FSM values for [ʃ] decreasing by approximately 800 Hz and mean FSM values for [s] increasing by approximately 200 Hz. The range of FSM values at midpoint for both [s] and [ʃ] were reduced as well. Although this speaker's magnitude of distinction increased from his initial completion of the study protocol to his revised production recordings, his quantitative distinction was still smaller than 60% of the other participants in the TD-11 group. Finally, the specific instructions given to B1107 to be clear and emphasize the fricative production was reflected in productions approximately 50 ms longer than his initial recordings. Even with the re-recording of tokens with specific instructions for production, B1107, the highest achiever on the discrimination task in the TD-11 group, did not produce the largest magnitude of distinction on the production task.

Another potential explanation for finding an absence of a linear relationship between production and perception may be because there was a limited range of performance on the discrimination measures for both groups of children. The range of proportion correct on the 2-step discrimination task for the TD-7 group was 0.55 - 0.69 (range = 0.14). The range of proportion correct

147

on the 2-step discrimination task for the TD-11 group was wider, at 0.54 – 0.86, but 7 out of 10 of the children scored in the range of 0.65 – 0.79 (range = 0.14), which was the same as the TD-7 group. Interestingly, the range of proportion correct for the 3-step discrimination task was wider for the TD-7 group, ranging from 0.52 – 0.83, whereas the range for the TD-11 group was smaller, from 0.63 – 0.95, with 8 out of 10 of the participants scoring between 0.80 – 0.95. In the Perkell et al. (2004b) study, the performance of the normal adults on 2-step discrimination ranged from 0.6 – 1.0, with the authors speculating that a ceiling effect was reached, with some adults likely having even better perceptual abilities than the 1.0 proportion correct was representing. In the current study, none of our participants reached 1.0 proportion correct. In addition, the limited range of discrimination performance may be masking the presence of a possible linear relationship between production and perception skills.

Third, we attempted to examine the relationship between production and perception in child speakers, rather than adults. It could be assumed that both perceptual and production skills are changing and likely improving, even at these ages. Evidence for continued improvement with age in both perceptual skill (e.g., Nittrouer, 1992) and speech motor control, reflected in smaller segment durations and less variability (e.g., Lee et al., 1999), has been found. Studying this relationship between perception and production in an immature, probably variable system may also have prevented a relationship between production and perception from being found. The typically developing children in these two age groups may have widely varying language experience, which can also affect perceptual skill (Nittrouer, 1996), although all of the children in this study passed

148

a language screening, indicating language abilities grossly within normal limits. The children in this study could have different levels of skill with motor planning for speech production, and in fact, the participants showed diverse patterns of variability in fricative production on the individual time history plots. Because of the ages chosen for this study, the children in the TD-11 group may have been in some stages of experiencing periods of significant growth, particularly in the male speakers (Fitch et al., 1999), causing changes in the vocal tract that may not yet be stable with regard to production patterns. Overall, the typically developing children in this study may have perceptual and productive systems that are not as mature and stable as adults, introducing significant variability, and possibly going through stages of rapid vocal tract growth, therefore potentially masking a straightforward relationship between perception and production.

Last, the relationship between perception and production may certainly be present, even if it was not found in this particular study. The Perkell et al. (2004b) study found a positive relationship between perception and production using FSM and discrimination measures; evidence for this relationship, however, was found in *groups* of adult speakers. Therefore, it is not certain that the positive relationship between perception and production of fricatives found in this study would be present in individual participants.

The relationship between these skills in children should be investigated further, to determine possible relationships between different measures of perception or production, or to investigate how this relationship changes over time in individual children. Individual children at both 7 and 11 years of age may have many factors interacting that could affect both their speech perception and

production development, including continued experience with their native language, speech motor control changes for [s] and [ʃ] productions, and changes in "fricative space" due to vocal tract growth.   All of these factors may be interacting differently in individual children, making the relationship between perception and production possibly fluid and changing at any point in a child's development.

Summary of Findings - Relationship between Production and Perception

Although a positive linear relationship was not found between production and perception of fricatives in the two groups of typically developing children in this study, it does not mean it does not exist. Methodological differences between studies that have found this relationship in adults (Perkell et al., 2004b) and this study were apparent and could be attributed to the lack of the relationship in this study. Variability in both perceptual skill and production measures may have precluded the ability to find a straightforward relationship between these two areas of development in typically developing children in this study.

Production and Perception in Children with Repaired Cleft Palate

The final research question was designed to investigate production and perception of [s] and [ʃ] in a group of children with repaired cleft lip and palate. It was hypothesized that children with repaired cleft palate with distorted fricative production would show spectral overlap of [s] and [ʃ] estimated by the FSM. If production and perception were assumed to co-evolve for these targets, we predicted that these same children would also show shallower identification curves and less accurate discrimination of [s] versus [ʃ] on the perception tasks.

Fricative Production

Compared to the typically developing children, a greater proportion of the children with repaired cleft palate demonstrated overlapping distributions and a greater degree of spectral overlap evidenced by the time history plots of fricative production. Three of the five children with repaired cleft palate demonstrated overlapping fricative distributions. These three participants, CP101, CP104, and CP105 demonstrated a magnitude of distinction of approximately 0 kHz, measured by the FSM at the midpoint of the segments. Although the time history plots for participant CP102 showed overlapping distributions, this speaker maintained some degree of distinction (700 Hz) between [s] and [ʃ] when the FSM at midpoint was averaged across the fricative, although this distinction is small compared to the majority of the typically developing children.

Despite the prevalence of overlapping distributions, the fricative productions of all the participants in the CLP had different degrees or patterns of distortion. There are several possible reasons for the reduction (CP102) or lack of a distinction (CP101, CP104, and CP105) between [s] and [ʃ] in these speakers. Occlusal status may be the primary factor affecting fricative production, for a number of different reasons. If a child presents with an underbite, it is possible that the space available for fricative production is reduced within the oral cavity. This could be described as a decrease in "fricative space," causing a shift in articulation place for the fricatives to a more posterior position, due to the upper jaw being positioned posterior to the lower jaw. The participants in this study with an underbite (CP101 and CP102) produced [s] and [ʃ] with reduced to no distinction. Occlusal status could also distort fricative noise being produced by

151

the children with repaired cleft palate. Several of the children in the CLP group

had missing teeth, malrotated teeth, or unusual anterior bite patterns

(protruding premaxilla, underbite and crossbite simultaneously). The

characteristics of the turbulence passing through the anterior bite patterns of

these children could distort the resultant noise created during fricative

production, affecting FSM values in different ways for individual children. CP104

was the only participant who was believed to produce a true articulation

substitution of [s] for [ʃ]. Therefore, this participant had not yet established the

motor skills required for [ʃ] production. This participant's mother also indicated

that he had just "begun to hear the difference" between [s] and [ʃ] in the speech

of others. The identification functions for [s] and [ʃ] for CP104 were similar to

those of the 7-year-old typically developing children, but this participant's

discrimination scores were among the lowest scores achieved when compared to

the scores of children in the TD-7 group. The possibility of a relationship

between this participant's production skill and discrimination performance may

exist.

It was interesting to find that the values for the standard deviation of

fricative production in the CLP group were within the range of standard deviation

demonstrated by the participants in the TD-7 and TD-11 group. This may

indicate that the children in the CLP group produced the fricatives with the same

degree of variability as the typically developing children, but the place of

articulation was shifted, causing a shift in mean FSM values. Whether this shift

in articulation place could be specifically attributed to dental arch anomalies,

such as the highly variable anterior bite patterns displayed in the CLP group,

"fricative space" limitations, or altered sensory feedback caused by scar tissue resulting from initial cleft palate repair, are variables to investigate in future research. In a study of adult speakers with aphasia with and without co-existing apraxia of speech, in addition to normal controls, Haley, Ohde and Wertz (2000) examined fricative production using analysis of the FSM. The authors found that almost all speakers with aphasia and co-existing aphasia and apraxia of speech demonstrated overlapping fricative distributions across repeated [s]and [ʃ] productions, yet the speakers presented with variable degrees of speech involvement overall, indicating likely different etiologies of speech impairment. Within the present study, although the majority of the participants with repaired cleft lip and palate produced fricatives with an almost non-existent magnitude of distinction, the etiology of the differences in fricative production among the participants in this group may have included effects of obligatory structural limitations causing distortion in fricative production and observed developmental articulation errors.

Fricative Perception

The participants in the CLP group showed the same pattern of monotonic crossover in fricative identification as the typically developing children, with the exception of the two participants (CP102 and CP105) on the [si] - [ʃi] continuum, mentioned in the results section. Participant CP102 had a significant history of otitis media and subsequent pressure-equalizing (PE) tube placement. CP102 currently had PE tubes placed at the time of his participation in the study. This positive history of middle ear disease, still requiring intervention at 9 years of age, may be one possible cause of his mildly different results on the identification

153

task, compared to the typically developing children and some of the children in the CLP group. Participant CP105, however, had an early history of otitis media, with the same number of ear infections reported as for CP102; however, no current issues with hearing or middle ear disease were reported by the parent for this participant. Participant CP105 was the only child in the study who did not demonstrate a monotonic crossover for a continuum, and his results for the [su] - [ʃu] continuum demonstrated a monotonic crossover, similar to the rest of the participants in the CLP group and typically developing children. His isolated difficulty with the [si] - [ʃi] continuum is not easy to explain.

Effects of severe, recurrent otitis media during early childhood on development of speech perception skills have been found in children with a history of otitis media (OM) with and without receptive / expressive language delays (Clarkson et al., 1989). These findings are contrary, however, to what was found in the present study. Although CP102 and CP105 demonstrated some difficulty with identification of the [si] - [ʃi] continuum, their discrimination scores and evidence of categorical perception overall was not markedly different from the other participants in the CLP group or the typically developing children. Both participants showed better discrimination scores at the boundaries of the [si] - [ʃi] and [su] - [ʃu] continua than at the ends of the continua, and their discrimination scores were within the range of performance compared to the other participants in the study.

Contrary to the findings in the study by Clarkson et al. (1989), the participants in the CLP group achieved results on the discrimination task within the range of performance of the TD-7 and TD-11 groups. A similar trend of the

154

older children in the CLP group (CP101 and CP102) achieving higher proportions correct on both the 2-step and 3-step discrimination task compared to the younger children (CP103 and CP104) was also observed.

In the current study, differences in performance on the identification task by the children in the CLP group compared to controls were more apparent than on the discrimination task. The results from the fricative identification task do provide some evidence that phonetic categorization skills may not be stable and consistent in all children with repaired cleft lip and palate. Factors that may contribute to difficulties with categorical perception of these sounds, in addition to other phonetic contrasts, should continue to be investigated in this population of children.

Relationship between Production and Perception

Within the CLP group, results demonstrating a relationship between perception and production were limited. Although the results of the discrimination task were similar to the TD-7 and TD-11 groups, the production results were different, with the children in the CLP group demonstrating a lower and sometimes non-existent magnitude of distinction between the two fricatives. However, some variability in fricative identification, albeit in only one or two participants, was observed. It could be possible that the variability or difficulty with fricative identification for participants CP102 and CP105 could be due to the lack of productive distinction. Participant CP102 showed a greater magnitude of distinction in production between the fricatives, and also showed less difficulty identifying the [si] - [ʃi] continuum in the perceptual identification task. Participant CP105 showed (and verbally expressed) significant difficulty

identifying [si] and [si]-like tokens within the perceptual identification task, and also demonstrated higher FSM values at midpoint for the [ʃ] than the [s] fricative.

Whitehill and colleagues (2003) found poor phonetic identification of a [t] – [k] continuum in children with cleft palate who lacked productive distinction between these two consonants. When comparing the results of this study with the results of Whitehill et al. (2003), the difficulty with the [s] - [ʃ] identification task in children who did not demonstrate accurate production of these sounds was not as apparent. However, results from the Whitehill et al. (2003) study and this study provide evidence that perceptual deficits may exist in children with repaired cleft palate who demonstrate merged productions for specific phonetic contrasts. The results of both of these studies have potentially important clinical implications. Often, production accuracy is the main outcome measure of speech assessment that receives attention in the clinical evaluation of children with cleft lip and palate. The results of this study indicate that implementing perceptual assessment of phonetic contrasts, in order to determine the possible contribution of these skills to production accuracy, could be an additional component of comprehensive evaluation of speech and language skills in this population of children. If perception of certain phonetic contrasts is indeed impaired to some degree in children with cleft lip and palate, particularly incorrectly produced phonetic contrasts, recommendations for perceptual training of these contrasts could be recommended as another aspect of speech intervention. Evidence for benefit of speech perception training on facilitation of improved sound production skills has been found for children who misarticulated [ʃ] (Rvachew, 1994). The effectiveness and usefulness of this type of perceptual training would

certainly require further research before implementation could take place in clinical settings.

Summary of Findings - Production and Perception in Children with Repaired Cleft Palate

The participants in the CLP group showed a higher proportion of overlapping fricative distributions and a decreased magnitude of a productive distinction between [s] and [ʃ]. These two findings showed some relationship to structural limitations caused by dental arch anomalies in these children, with the possibility of reduced "fricative space" decreasing the degree to which the fricatives can be produced in a differentiated manner.

Most of the participants in the CLP group showed similar phonetic categorization results compared to the typically developing children; however, some evidence of difficulty with fricative identification was found in two participants. Performance of the participants in the CLP group on tests of fricative discrimination was similar to that of the typically developing participants.

The relationship between production and perception of [s] and [ʃ] warrants further investigation, particularly to examine the relevance of factors in this population that could be interacting to affect this relationship, such as structural limitations in production and possible deficits in perceptual categorization of specific phonetic contrasts.

Limitations of the Current Study

Several limitations of the current study warrant discussion. First, some methodological changes could have controlled some of the variability present in

the current study, which may have affected the relationship between perception and production, if one does exist in typically developing children. For example, a specific model for the production task was not provided for the participants. Although specific instructions were given to the participants for the task, these instructions could have been interpreted differently by any of the participants. Smith and Kenney (1994), however, investigated the effects of experimentally controlling temporal variability on production of [s] in children (age 7 to 11 years of age) and adults. In the control condition, the participants were given general instructions about saying the stimulus words and phrases clearly and naturally after the examiner; in the experimental condition, the participants were instructed to be "as consistent as possible in terms of loudness, duration, intonation" (p. 700). The investigators found small differences between the control and experimental conditions: an approximate 5 ms difference between the adult control and experimental conditions, and approximately 10 – 15 ms difference between control and experimental conditions for the children. In addition, variability in duration between the control and experimental conditions for the child participants was virtually identical. While these results do not directly speak to spectral variability, which was specifically studied in the present investigation, the results from the Smith and Kenney (1994) study provided evidence that participants are relatively consistent when performing repetitive speech tasks, such as the task in this investigation, whether given specific instructions or not.

Another example of how inter- or intra-speaker variability may have been increased in this study was related to the number of production tokens required

of the speakers. Although the study was piloted on 3 children prior to initiating

the investigation, only one of the children was in the 7-year-old range. Having

children in the TD-7 group produce 100 tokens for the production task may have

been taxing and performance may have decreased over the course of the task,

particularly for participants that may have found the task to be difficult in terms

of motor planning and execution. Only one subject (B711), however, showed

specific difficulty completing the task accurately. Regardless of age, 100 tokens

may have become tedious or uninteresting for any of the child participants in this

study. Participants may have lost interest in producing the tokens consistently or

with their best effort as the task continued, and the loss of interest may have

occurred at different times during the task for any participant. Therefore, the

choice of number of tokens to be representative of a child's speech production

skills, while maintaining interest and motivation for the task, should be an

important decision when conducting further studies such as the present one.

Additional factors such as dialect and vowel context may have introduced

variability into the production data. Not controlling for dialect could introduce

vowel differences that may have affected production of the preceding fricative

consonant in some speakers. Furthermore, the choice of the vowel [u] may have

increased variability in the preceding fricative consonant. For some speakers, the

vowel [u] introduced additional overlap of fricative targets in both the TD-7 and

TD-11 groups. A study of vowel acoustic space development in children

(Vorperian et al., 2007) found increased variability of F2 for the vowel [u] across

the developmental age range studied (ages 0 – 20 years). This variability was

attributed to dialectical influences, variability in articulation, and growth of the

vocal tract. Having the participants in this study produce the fricatives [s] and [ʃ] with the vowel [u] may have introduced additional variability above and beyond what is due to articulation development and the development of speech motor control in the children studied.

The use of different methods of measuring fricative production needs to be considered. While the FSM has been identified as a useful measure that corresponds well with fricative place of articulation, the use of other spectral moments may be appropriate when studying the children's speech production. Both the first and third (skewness) spectral moment were described as "essential to characterize [s] independent of phonetic context" (Flipsen et al., 1999, p. 675), and the second moment (variance) may be important if the onset or offset of [s] frication is being analyzed. Flipsen and colleagues (1999) did not find that the analysis of the fourth spectral moment (kurtosis) added any information regarding [s] characteristics in typically developing adolescent speakers (age 9 – 14 years). Nittrouer (1995) came to a similar conclusion as Flipsen et al. (1999) regarding the fourth spectral moment in a study investigating developmental trends in [s] and [ʃ] production, but a different conclusion regarding the third spectral moment. She concluded that the third spectral moment provided information that was similar, and therefore redundant, to information provided by the first spectral moment, when investigating age-related patterns in fricative development (Nittrouer, 1995).

A more comprehensive examination of fricative production, however, including additional moments beyond the FSM, may provide additional information regarding *disordered* fricative production in children. As

160

mentioned, the children with repaired cleft lip and palate produced [s] and [ʃ] with different degrees and patterns of distortion that may or may not be accurately described with auditory-perceptual analysis alone. Acoustic data regarding fricative production in disordered speech may be a more sensitive and reliable method to classify characteristics of speech production. A comprehensive spectral moment analysis of fricatives, for example, in children with repaired cleft palate, could provide additional information regarding production characteristics that could then be analyzed in combination with other child factors (such as objective measures of occlusal status) to obtain a more in-depth description of a child's speech production development.

Another limitation of this study was the lack of an adult control group. The addition of an adult control group in this study would have been useful to compare the performance of the children in this study to adult performance. Because there were no group differences in magnitude of distinction in [s] and [ʃ] between the TD-7 and TD-11 group, conclusions regarding continued development in fricative production toward adult-like speech production could not be made. In addition, it cannot be determined if the participants in the TD-11 group, for example, performed similarly to adult levels of perceptual performance on the discrimination task, without the inclusion of an adult control group that had completed the same study protocol.

The effect of frequency of otitis media (OME) events on production or perception results in the control group was not able to be determined by the methodology used in this study. Although parents were asked to report on frequency of otitis media, the responses from parents appeared to lack confidence

and precision, because many parents were unable to provide a number that they felt was accurate. Therefore, the variable of OME frequency was not used in quantitative analysis. Qualitatively, the participants with reportedly greater frequency of OME (10 – 12 events) did not demonstrate productive or perceptual results that were outside the range of performance from the other participants in their respective age groups. Future studies of production or perception should include a more objective variable, such as history of otitis media garnered from medical records, or a more specific measure of hearing status beyond a screening alone.

The analysis of the production data with regard to the CLP group was limited. Although anterior bite patterns were described, objective measures to classify and categorize dental arch relationships were not utilized, and these measures may have provided important information with regard to fricative production for the participants in this group. In addition, the sample size was very small, which also limited the analysis in this group. Future studies would benefit from an objective measure of occlusal status, in addition to documentation of type of cleft, type of surgical repair, and age of surgical repair, for example, in order to determine the relative contribution of any of these variables to the development of perception and production in these groups.

Future Directions

The necessity for longitudinal data collection, in addition to cross-sectional studies, seems required for a more complete understanding of the development of speech production and perception in children. In the present study, a wide range of individual speaker patterns in fricative production were

162

observed, including a considerable range of variability in production. Although *group* differences in *average* discrimination performance were found between the older and younger children, overlap in discrimination performance was observed between individual children in each age group. Observing and documenting individual child speaker patterns over time may be useful to better understand development of characteristics of speech production. With respect to children with disordered speech, studying performance over time may provide insights in terms of efficiency or effectiveness over the course of treatment, for example.

The utility of spectral moment analysis as a clinical measure warrants further attention as well. If spectral moment analysis can be used to objectively measure and document consonant production characteristics, then this analysis method could be useful for the practicing clinician. Acoustic analysis has become more readily available and accessible to the clinician, with free analysis programs available through the internet, for example. However, training and exposure with acoustic analysis methods will likely be necessary in order to integrate these methods into those traditionally used for assessment of speech production, namely phonetic transcription. Acoustic analysis may be a useful augmentation to these methods and provide objective and quantifiable information regarding speech production that is not currently being utilized.

One possible method of affirming the use of acoustic analysis would be studying the relationship between the auditory perceptual analysis of the speech signal and its corresponding acoustic analysis. For example, what is the relationship between the perception of fricative quality for the children with

smaller or greater magnitude of production distinction? This would be particularly interesting to study in children with repaired cleft palate who did not make an acoustic distinction between the sounds. Would listeners be able to distinguish between the two fricatives? If the acoustic analyses showed strong correspondence to the auditory perceptual targets, then the usefulness of employing acoustic analysis to supplement auditory perceptual analysis could be supported.

Continuing to study the utility of spectral moment analysis for children with repaired cleft palate could lead to employing the measure for clinical applications related to speech outcome. For example, spectral moment analysis could be used to document status of fricative production pre-and post-surgically during the course of orthognathic (jaw) surgery common to children with repaired cleft palate. These measures could provide objective information for the cleft palate team in terms of surgical outcome related to speech characteristics. Many children with repaired cleft palate produce sounds posterior to their ideal place of articulation (i.e. pharyngeal fricative substitutions for lingua-alveolar fricative targets). Using spectral moment analysis to provide objective information regarding place of articulation for fricative targets could also be useful, in order to document how production could change during the course of speech treatment.

Conclusions

The purpose of this study was to investigate if a relationship between production and perception of the voiceless sibilant fricatives [s] and [ʃ] exists in a sample of typically developing children, age 7 and 11, and if developmental

differences in these skills existed between these two groups of children. In addition, production and perception of [s] and [ʃ] in a small group of children with repaired cleft lip and palate were compared to the performance of the typically developing children.

While almost all of the typically developing children produced [s] and [ʃ] with a consistent distinction, inspection of individual speaker data showed varying degrees of token-to-token variability and variability in dynamic patterns of fricative production. The range of magnitude of distinction in fricative production varied both within and across groups as well. For example, children who produced both large (4.4 kHz) and small (0.7 kHz) distinctions between [s] and [ʃ] were both perceived by the examiner to produce these sounds correctly. Individual differences in fricative production between speakers would have been lost if group analyses alone had been utilized.

Fricative perceptual skill appears to continue to improve with age, similar to results of previous studies. However, individual differences in performance were apparent, with some younger listeners performing within the range of the older children and vice versa.

The relationship between perception and production of [s] and [ʃ] in typically developing children may exist, although it was not found in this study. We expected the children in this study to perform similar to previous findings regarding these skills in adults. Methodological differences between studies of adults and the present study may have precluded the finding of the same type of relationship between perception and production in typically developing children. Multiple factors, including continued improvement with age in both perceptual

skills and speech motor control development and the possibility of significant vocal tract growth during the ages of the children in this study, may have prevented a straightforward relationship between perception and production to emerge.

The children with repaired cleft lip and palate showed a reduction in magnitude of distinction between [s] and [ʃ] compared to the typically developing children, with some participants showing complete spectral overlap. Fricative production in the children with repaired cleft lip and palate was considered to be distorted in all speakers, and this auditory-perceptual observation was reflected in the analysis of the FSM for these speakers. In addition, evidence of diminished perceptual performance in both identification and discrimination tasks in some participants was found. While the specific relationship between perception and production of [s] and [ʃ] could not be definitively determined, the results of this study indicate that future research is warranted to investigate these skills in a larger population of children with repaired cleft lip and palate.

Appendix 1

Order of Stimuli for Identification Screening Test

1. Sh
2. Sh
3. S
4. S
5. S
6. Sh
7. S
8. S
9. S
10. Sh
11. Sh
12. S
13. S
14. Sh
15. S
16. Sh
17. Sh
18. S
19. Sh
20. Sh

Pictures and Orthographic Representations for Perceptual Identification
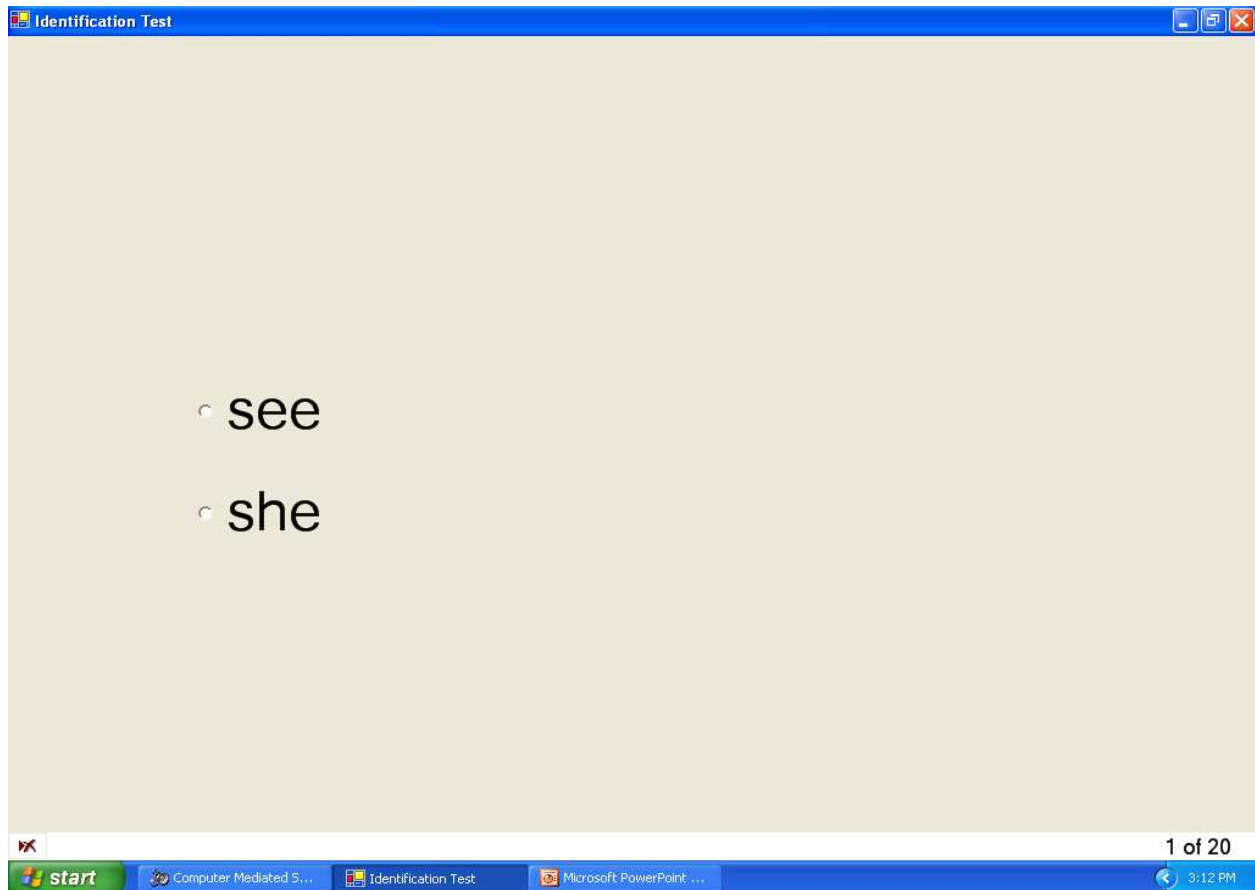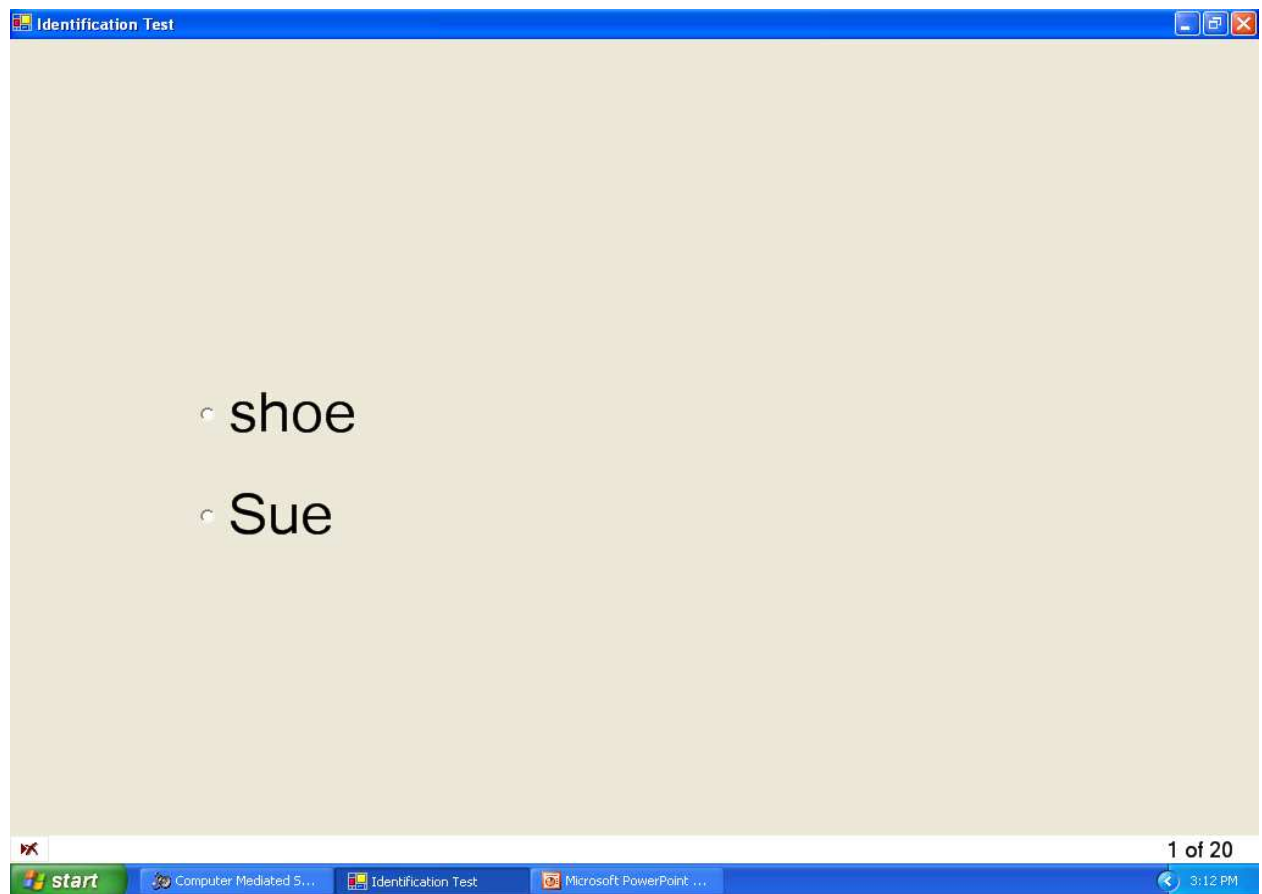Screening/Task

SHE:

SEE:
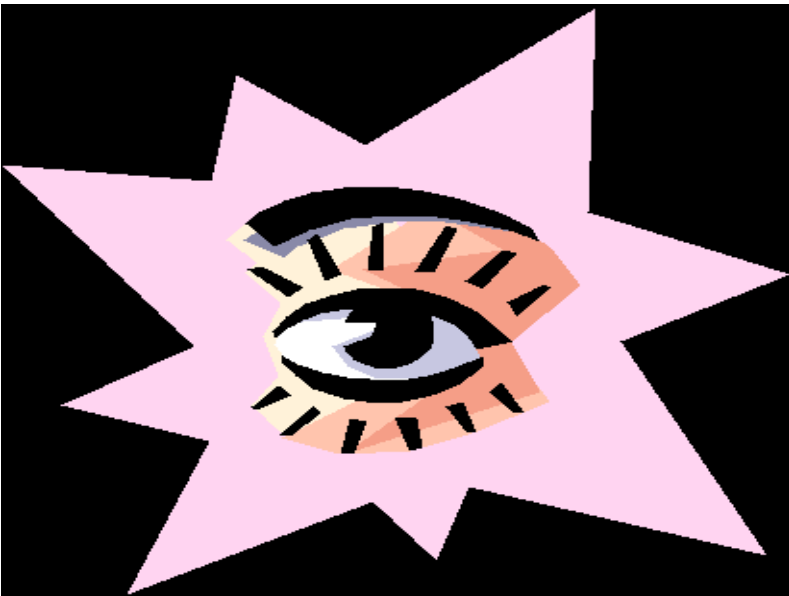
SHOE:

SUE:

Appendix 3

Visual Representation of Identification Task from the SpeechMeasures Software
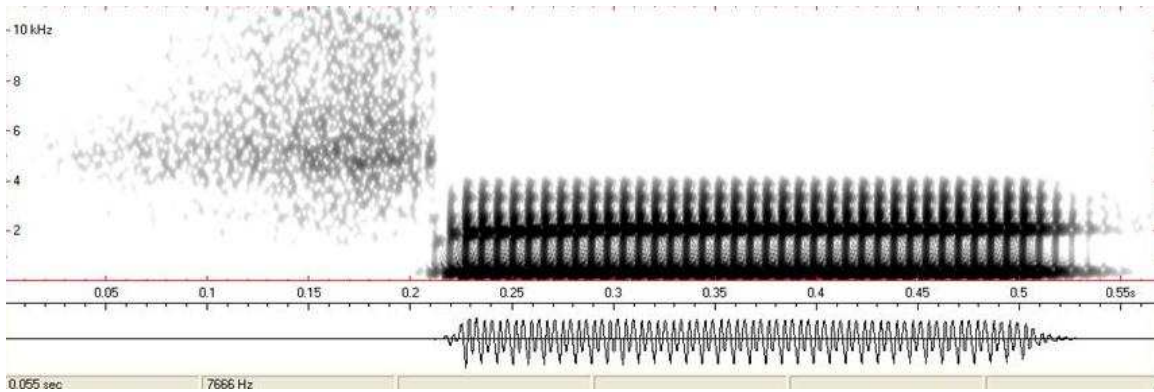Program

Appendix 4

Order of Words for Production Test

|		
1. She	38. She	75. Sue
2. See	39. See	76. Shoe
3. Sue	40. Sue	77. See
4. Shoe	41. See	78. Sue
5. See	42. Shoe	79. Shoe
6. Sue	43. She	80. She
7. Shoe	44. Sue	81. Sue
8. She	45. She	82. Shoe
9. Sue	46. Sue	83. She
10. Shoe	47. See	84. See
11. She	48. Shoe	85. Shoe
12. See	49. She	86. She
13. Shoe	50. See	87. See
14. She	51. Sue	88. Sue
15. See	52. Shoe	89. See
16. Sue	53. See	90. Shoe
17. See	54. Sue	91. She
18. Shoe	55. Shoe	92. Sue
19. She	56. She	93. She
20. Sue	57. Sue	94. Sue
21. She	58. Shoe	95. See
22. Sue	59. She	96. Shoe
23. See	60. See	97. She
24. Shoe	61. Shoe	98. See
25. She	62. She	99. Sue
26. See	63. See	100. Shoe
27. Sue	64. Sue	
28. Shoe	65. See	
29. See	66. Shoe	
30. Sue	67. She	
31. Shoe	68. Sue	
32. She	69. She	
33. Sue	70. Sue	
34. Shoe	71. See	
35. She	72. Shoe	
36. See	73. She	
37. Shoe	74. See	

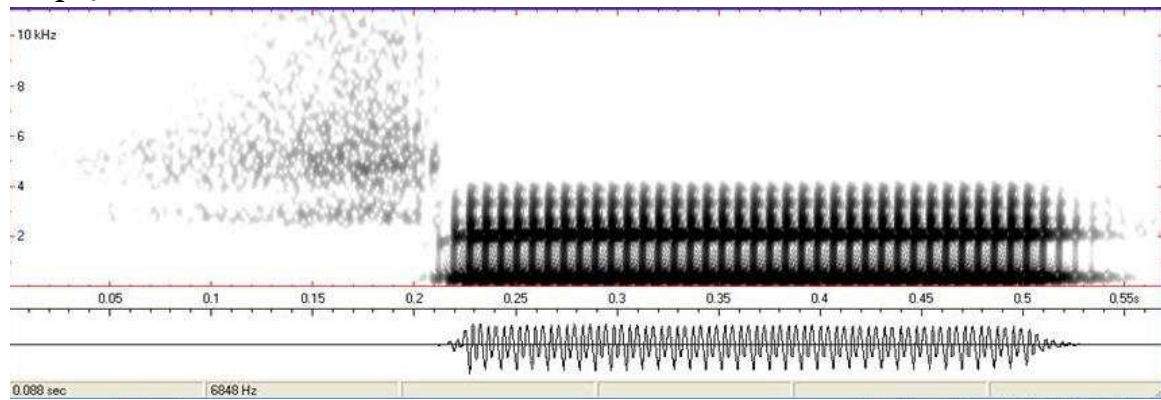# Appendix 5

Spectrograms of [si] – [ʃi] continuum
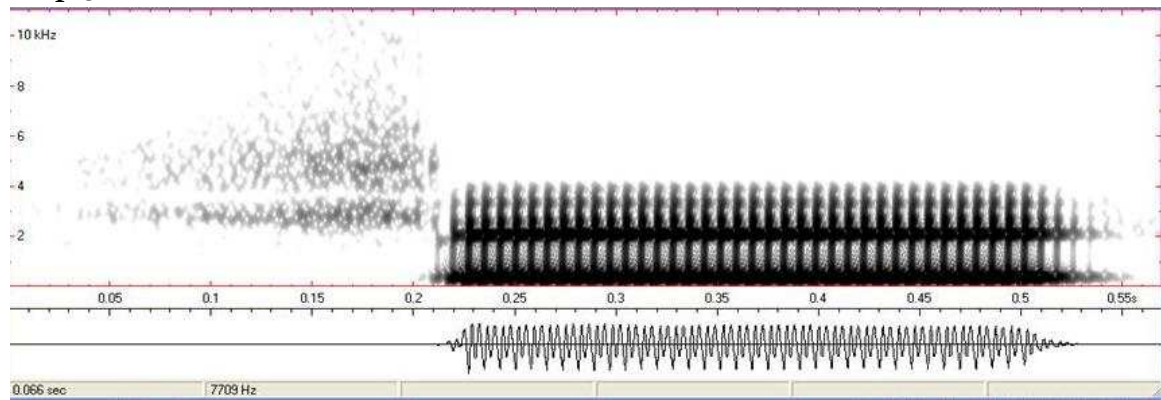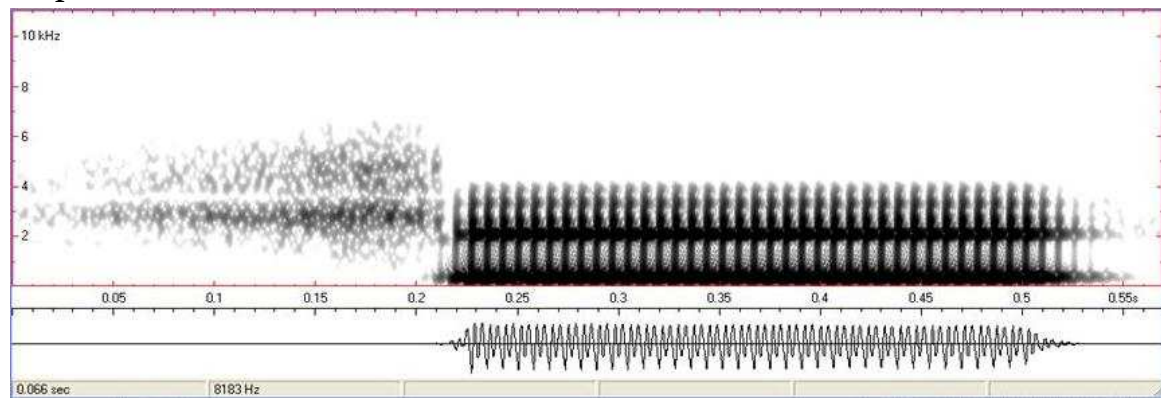
Step 1 ([si])
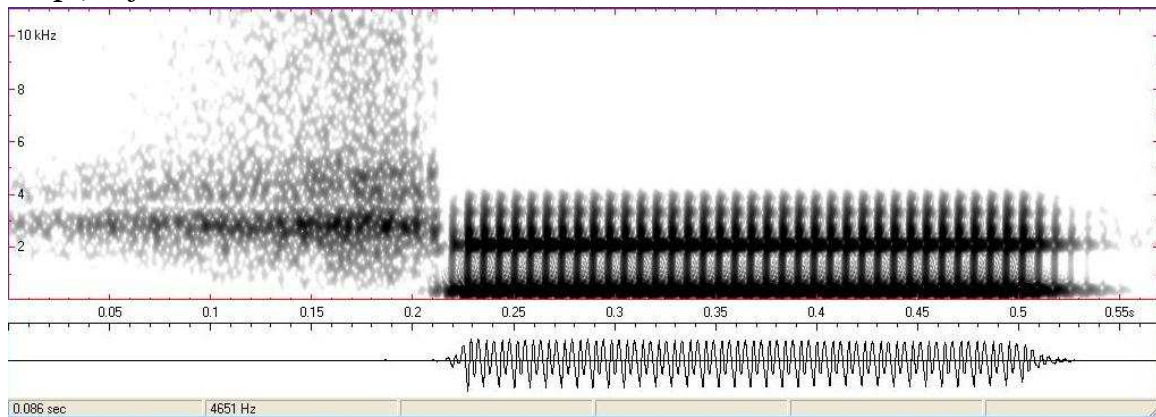

Step 2

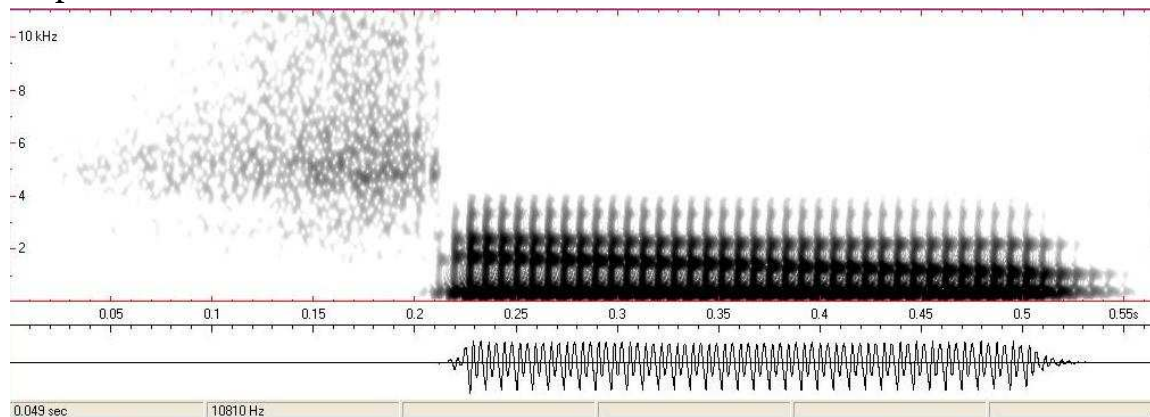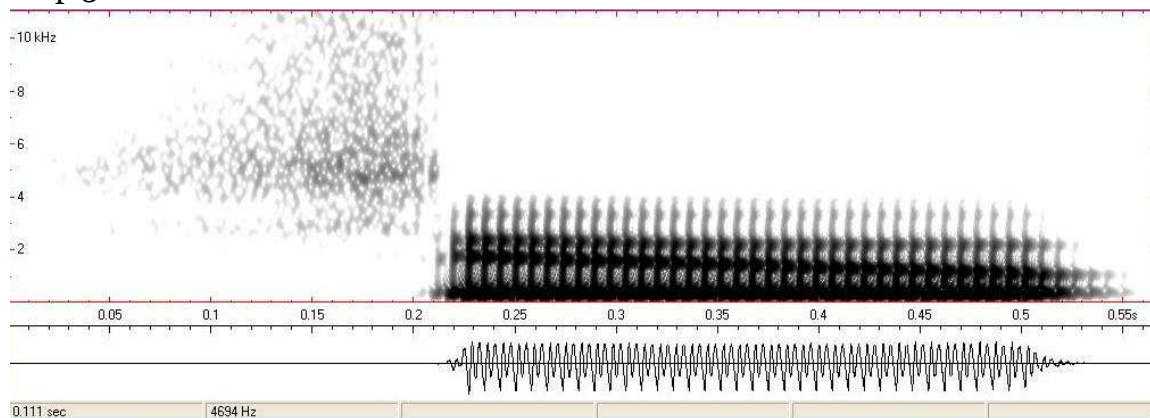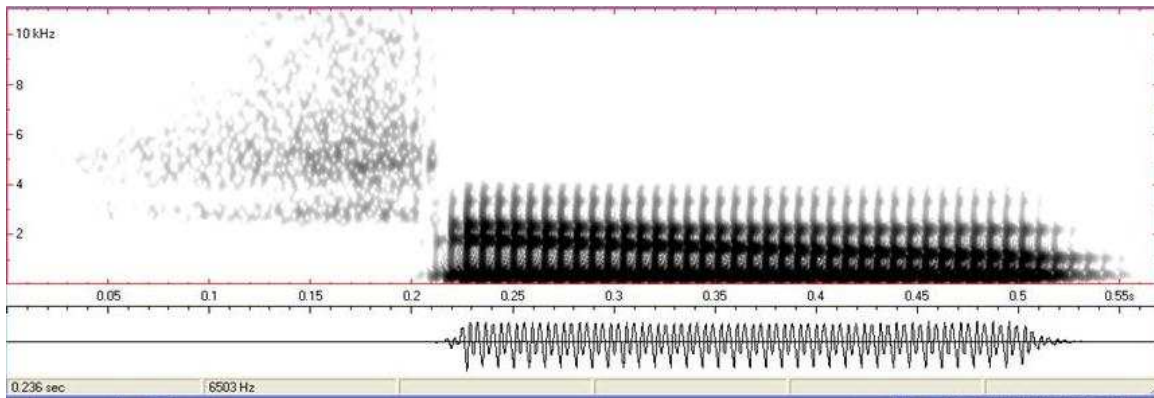

Step 3

Step 4

Step 5

Step 6

174
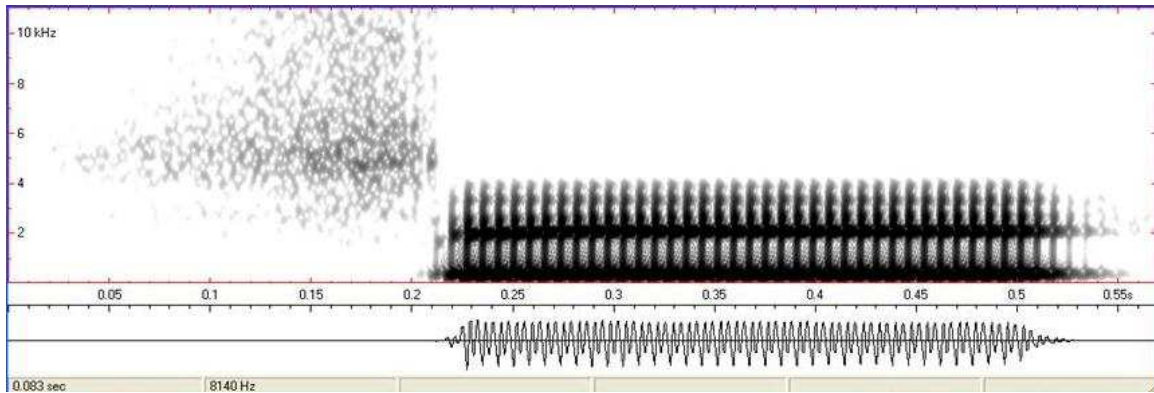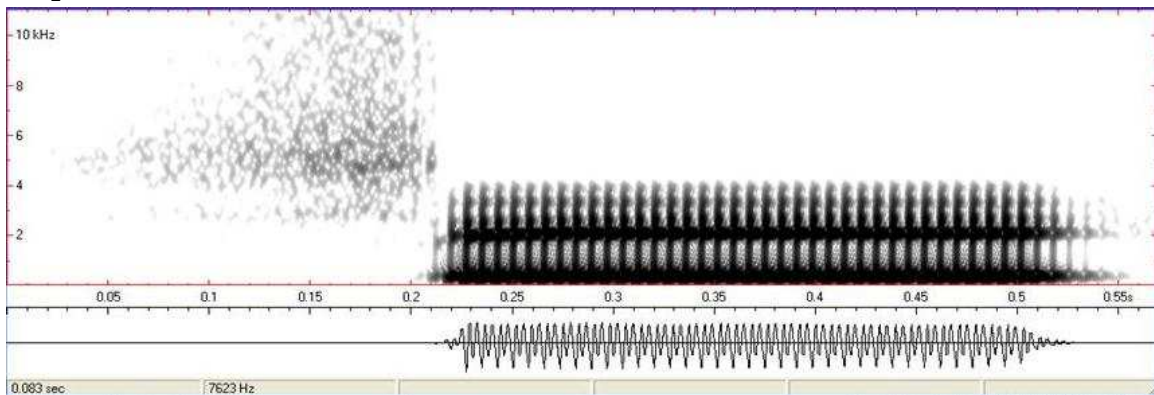
Step 7 ([ʃi])

Appendix 6

Spectrograms of [su] – [ʃu] continuum
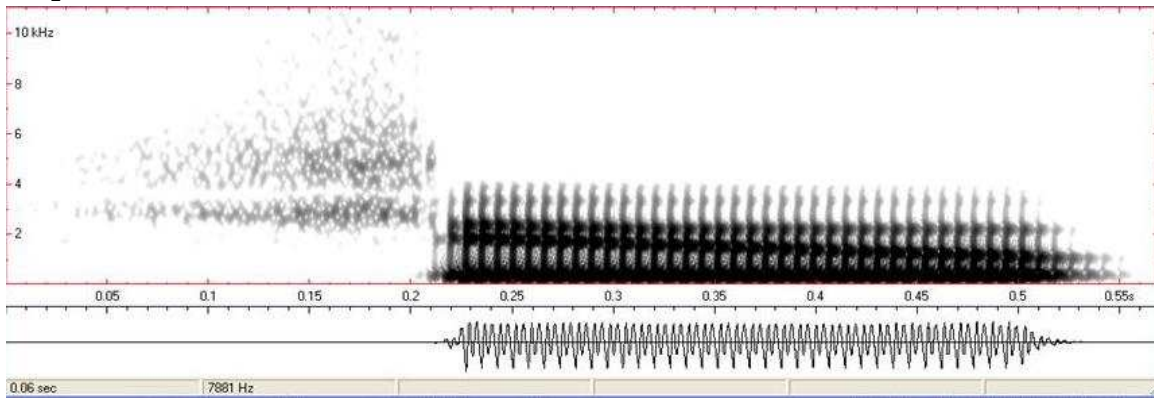
Step 1 ([su])

Step 2

Step 3
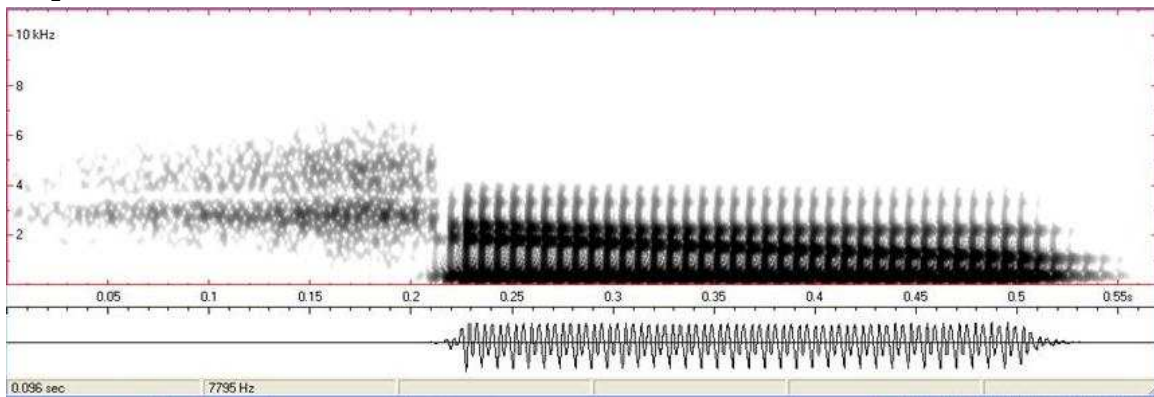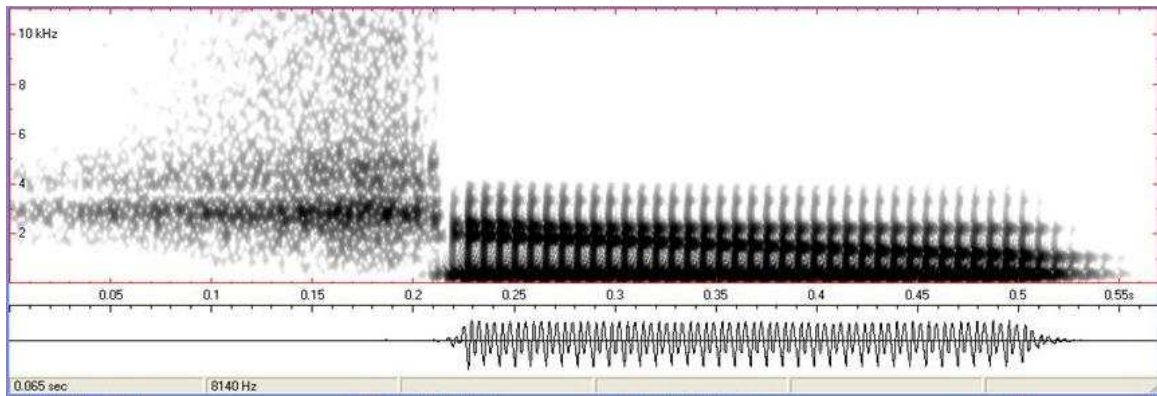
Step 4

Step 5

Step 6

177
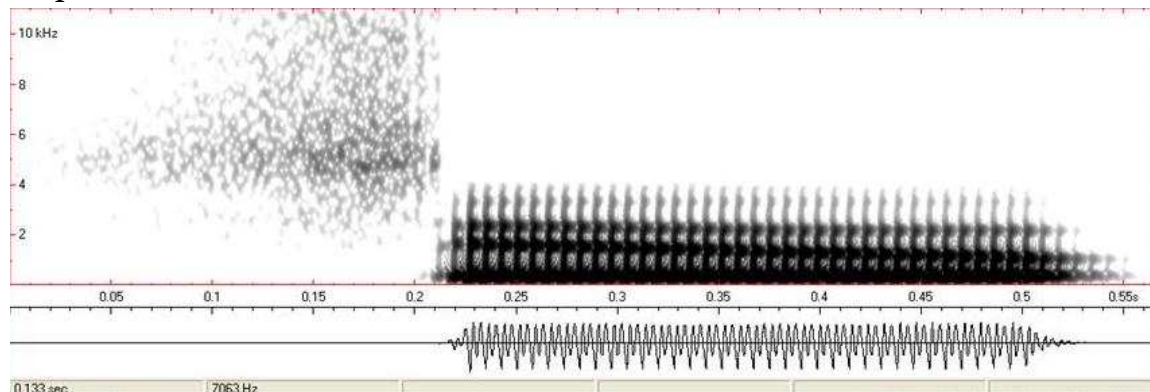
Step 7 ([ʃu])

Appendix 7

Order of Continuum Steps for the Identification Test

1. Step 5
2. Step 4
3. Step 2
4. Step 6
5. Step 1
6. Step 7
7. Step 3
8. Step 4
9. Step 7
10. Step 6
11. Step 1
12. Step 2
13. Step 5
14. Step 3
15. Step 6
16. Step 4
17. Step 5
18. Step 3
19. Step 2
20. Step 7
21. Step 1
22. Step 5
23. Step 7
24. Step 4
25. Step 2
26. Step 1
27. Step 6
28. Step 3
29. Step 5
30. Step 6
31. Step 7
32. Step 2
33. Step 4
34. Step 3
35. Step 1
36. Step 5
37. Step 1

38. Step 3
39. Step 4
40. Step 2
41. Step 7
42. Step 6
43. Step 1
44. Step 3
45. Step 6
46. Step 4
47. Step 2
48. Step 7
49. Step 5
50. Step 1
51. Step 2
52. Step 7
53. Step 6
54. Step 5
55. Step 4
56. Step 3
57. Step 4
58. Step 3
59. Step 1
60. Step 5
61. Step 6
62. Step 7
63. Step 2
64. Step 3
65. Step 4
66. Step 1
67. Step 7
68. Step 2
69. Step 5
70. Step 6

Appendix 8

Visual Representation of the Discrimination Task from the SpeechMeasures Software Program
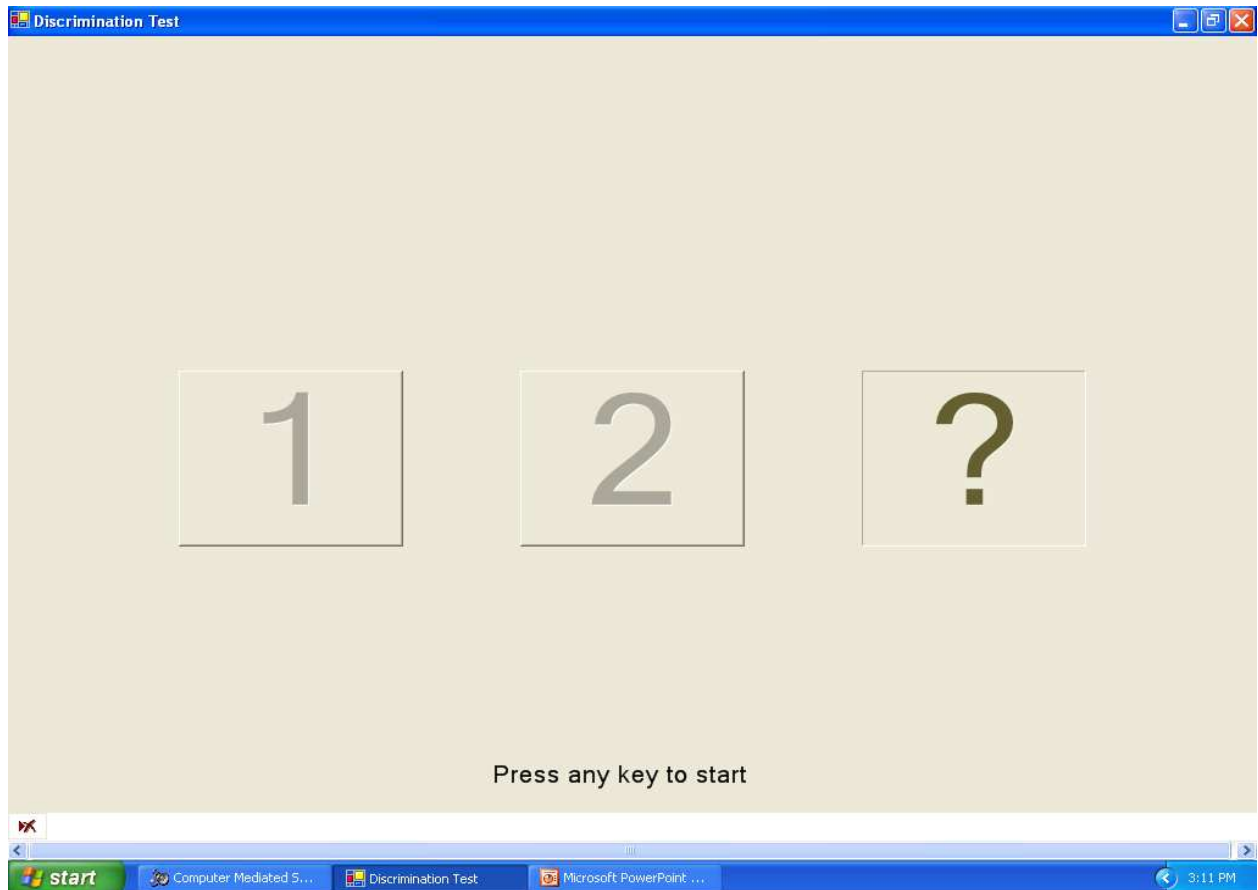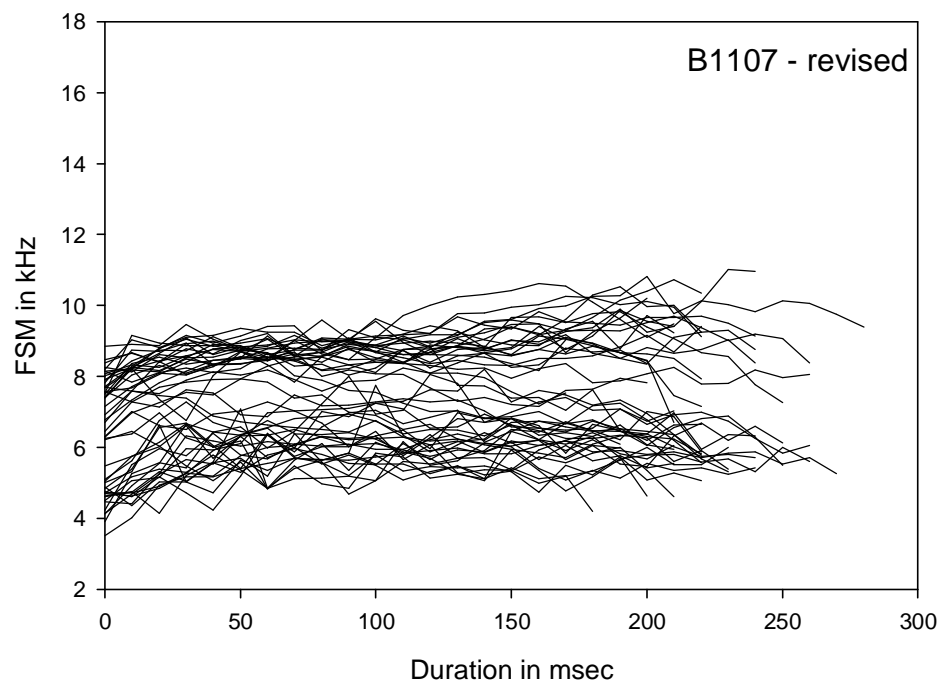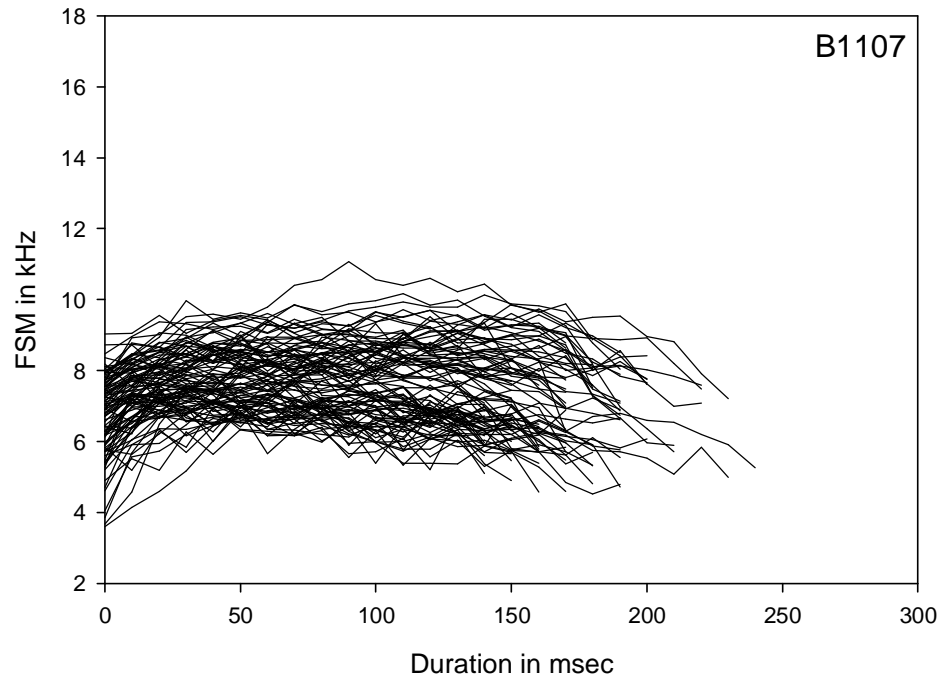
Time history plots of fricative production for participant B1107. The first plot was the original recording, the second plot was the second recording.

References

Baum, S. R., & McNutt, J. C. (1990). An Acoustic Analysis of Frontal Misarticulation of [S] in Children. *Journal of Phonetics, 18*(1), 51-63.

Broen, P. A., Strange, W., Doyle, S. S., & Heller, J. H. (1983). Perception and production of approximant consonants by normal and articulation-delayed preschool children. *Journal of Speech & Hearing Research, 26*, 601-608.

Ceponiene, R., Hukki, J., Cheour, M., Haapanen, M. L., Koskinen, M., Alho, K., et al. (2000). Dysfunction of the auditory cortex persists in infants with certain cleft types. *Developmental Medicine & Child Neurology, 42*(4), 258-265.

Ceponiene, R., Hukki, J., Cheour, M., Haapanen, M. L., Ranta, R., & Naatanen, R. (1999). Cortical auditory dysfunction in children with oral clefts: relation with cleft type. *Clinical Neurophysiology, 110*(11), 1921-1926.

Chapman, K. L. (1991). Vocalizations of toddlers with cleft lip and palate. *Cleft Palate Craniofac J, 28*(2), 172-178.

Chapman, K. L., Hardin-Jones, M., Schulte, J., & Halter, K. A. (2001). Vocal development of 9-month-old babies with cleft palate. *J Speech Lang Hear Res, 44*(6), 1268-1283.

Cheour, M., Ceponiene, R., Hukki, J., Haapanen, M. L., Naatanen, R., & Alho, K. (1999). Brain dysfunction in neonates with cleft palate revealed by the mismatch negativity. *Clinical Neurophysiology, 110*(2), 324-328.

Cheour, M., Haapanen, M. L., Ceponiene, R., Hukki, J., Ranta, R., & Naatanen, R. (1998). MMN as an index of auditory sensory memory deficit in cleft palate and CATCH children. *NeuroReport, 9*, 2709-2712.

Clarkson, R. L., Eimas, P. D., & Marean, G. C. (1989). Speech perception in children with histories of recurrent otitis media. *Journal of the Acoustical Society of America, 85*(2), 926-933.

Eguchi, S., & Hirsch, I. (1969). Development of speech sounds in children. *Acta Otolaryngologica Supplementum, 257*, 1-51.

Estrem, T., & Broen, P. A. (1989). Early speech production of children with cleft palate. *J Speech Hear Res, 32*(1), 12-23.

Finnegan, D. E. (1974). Speech sound discrimination skills of seven- and eight-year-old cleft palate males. *Cleft Palate Journal, 11*, 111-122.

Fitch, W. T., & 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*(3), 1511 - 1522.

Flipsen, P., Shriberg, L., Weismer, G., Karlsson, H., & McSweeny, J. (1999). Acoustic characteristics of /s/ in adolescents. *Journal of Speech Language and Hearing Research, 42*(3), 663-677.

Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: preliminary data. *J Acoust Soc Am, 84*(1), 115-123.

Fox, R. A., & Nissen, S. L. (2005). Sex-related acoustic changes in voiceless English fricatives. *Journal of Speech Language and Hearing Research, 48*(4), 753-765.

Goldman, R., & Fristoe, M. (2000). Goldman-Fristoe Test of Articulation (2nd ed.). Bloomington, MN: Pearson Assessments.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102*(3), 594-621.

Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review, 105*(4), 611-633.

Guenther, F. H., Husain, F. T., Cohen, M. A., & Shinn-Cunningham, B. G. (1999). Effects of categorization and discrimination training on auditory perceptual space. *Journal of the Acoustical Society of America, 106*(5), 2900 - 2912.

Hagle, E. (2002). *Spectral moment analysis in voiceless sibilant fricatives.* Unpublished Master's Thesis, University of North Carolina at Chapel Hill, Chapel Hill

Haley, K. L. (2002). Temporal and Spectral Properties of Voiceless Fricatives in Aphasia and Apraxia of Speech. *Aphasiology, 16*(4-6), 595-607.

Haley, K. L. (2008). *Computer-Mediated Assessment of Speech Intelligibility in Aphasia & AOS.* Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

Haley, K. L., Ohde, R. N., & Wertz, R. T. (2000). Precision of Fricative Production in Aphasia and Apraxia of Speech: A Perceptual and Acoustic Study. *Aphasiology, 14*(5-6), 619-634.

Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6-12. *Journal of Phonetics, 28*, 377-396.

Hedrick, M. S., & Ohde, R. N. (1993). Effect of relative amplitude of frication on perception of place of articulation. *J Acoust Soc Am, 94*(4), 2005 - 2026.

Hochberg, I., & Kabcenell, J. (1967). Oral Stereognosis in Normal and Cleft Palate Individuals. *Cleft Palate Journal, 4*, 47-57.

Hoffman, P. R., Daniloff, R. G., Bengoa, D., & Schuckers, G. H. (1985). Misarticulating and normally articulating children's identification and discrimination of synthetic /r/ and /w/. *Journal of Speech and Hearing Disorders, 50*, 46-53.

Hughes, G. W., & Halle, M. (1956). Spectral properties of fricative consonants. *Journal of the Acoustical Society of America, 28*(2), 303-310.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *J Acoust Soc Am, 108*(3 Pt 1), 1252-1263.

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

Kent, R. D., & Forner, L. L. (1980). Speech segment durations in sentence recitations by children and adults. *Journal of Phonetics, 8*, 157-168.

Kent, R. D., & Read, C. (2002). *The Acoustic Analysis of Speech* (2nd ed.). Albany, NY: Delmar.

Koenig, L. L., Lucero, J. C., & Perlman, E. (2008). Speech production variability in fricatives of children and adults: Results of functional data analysis. *The Journal of the Acoustical Society of America, 124*(5), 3158-3170.

Kuehn, D. P., & Moller, K. T. (2000). Speech and language issues in the cleft palate population: The state of the art. *Cleft Palate - Craniofacial Journal, 12*, 85-94.

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America, 105*(3), 1455 - 1468.

Liberman, A. M., Cooper, F., Shankweiler, D., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*, 431-461.

Liberman, A. M., Harris, K. S., Kinney, J. A., & Lane, H. (1951). The discrimination of relative onset time of the components of certain speech

and nonspeech patterns. *Journal of Experimental Psychology, 61*, 379-388.

Locke, J. L. (1980). The inference of speech perception in the phonologically disordered child. Part I: A rationale, some criteria, the conventional tests. *Journal of Speech and Hearing Disorders, 45*, 431-444.

McGowan, R. S., & Nittrouer, S. (1988). Differences in fricative production between children and adults: evidence from an acoustic analysis of /sh/ and /s/. *J Acoust Soc Am, 83*(1), 229-236.

Milenkovic, P. M., & Read, C. (1997). CSpeech. Madison, WI: University of Wisconsin-Madison, Department of Electrical Engineering.

Monnin, L. M., & Huntington, D. A. (1974). Relationship of articulatory defects to speech-sound identification. *Journal of Speech and Hearing Research, 17*, 352-366.

Munson, B. (2004). Variability in /s/ production in children and adults: Evidence from dynamic measures of spectral mean. *Journal of Speech Language and Hearing Research, 47*(1), 58-69.

Newman, R. S. (2003). Using links between speech perception and speech production to evaluate different acoustic metrics: a preliminary report. *J Acoust Soc Am, 113*(5), 2850-2860.

Newman, R. S., Clouse, S. A., & Burnham, J. L. (2001). The perceptual consequences of within-talker variability in fricative production. *J Acoust Soc Am, 109*(3), 1181-1196.

Nissen, S. L., & Fox, R. A. (2005). Acoustic and spectral characteristics of young children's fricative productions: a developmental perspective. *J Acoust Soc Am, 118*(4), 2570-2578.

Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics, 20*(3), 351-382.

Nittrouer, S. (1995). Children Learn Separate Aspects of Speech Production at Different Rates - Evidence from Spectral Moments. *Journal of the Acoustical Society of America, 97*(1), 520-530.

Nittrouer, S. (1996). The Relation Between Speech Perception and Phonemic Awareness: Evidence From Low-SES Children and Children With Chronic OM. *J Speech Hear Res, 39*(5), 1059-1070.

Nittrouer, S. (2002). Learning to perceive speech: how fricative perception changes, and how it stays the same. *J Acoust Soc Am, 112*(2), 711-719.

Nittrouer, S., Crowther, C. S., & Miller, M. E. (1998). The relative weighting of acoustic properties in the perception of [s] + stop clusters by children and adults. *Percept Psychophys, 60*(1), 51-64.

Nittrouer, S., Manning, C., & Meyer, G. (1993). The perceptual weighting of acoustic cues changes with linguistic experience. *The Journal of the Acoustical Society of America, 94*(3), 1865.

Nittrouer, S., & Miller, M. E. (1997). Developmental weighting shifts for noise components of fricative-vowel syllables. *J Acoust Soc Am, 102*(1), 572-580.

Nittrouer, S., & Studdert-Kennedy, M. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *J Speech Hear Res, 30*(3), 319-329.

Nittrouer, S., Studdert-Kennedy, M., & McGowan, R. S. (1989). The emergence of phonetic segments: evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research, 32*(1), 120-132.

Nittrouer, S., Studdert-Kennedy, M., & Neely, S. T. (1996). How children learn to organize their speech gestures: further evidence from fricative-vowel syllables. *J Speech Hear Res, 39*(2), 379-389.

Ohde, R. N., & Haley, K. L. (1997). Stop-consonant and vowel perception in 3- and 4-year-old children. *Journal of the Acoustical Society of America, 102*(6), 3711-3722.

Ohde, R. N., Haley, K. L., Vorperian, H. K., & McMahon, C. W. (1995). A developmental study of the perception of onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America, 97*(6), 3800-3812.

Ohde, R. N., & Sharf, D. J. (1988). Perceptual categorization and consistency of synthesized /r-w/ continua by adults, normal children, and /r/-misarticulating children. *Journal of Speech and Hearing Research, 31*, 556-568.

Paradise, J. L., Campbell, T. F., Dollaghan, C. A., Feldman, H. M., Bernard, B. S., Colborn, D. K., et al. (2005). Developmental Outcomes after Early or Delayed Insertion of Tympanostomy Tubes. *New England Journal of Medicine, 353*(6), 576 - 586.

Pentz, A., Gilbert, H. R., & Zawadzki, P. (1979). Spectral properties of fricative consonants in children. *The Journal of the Acoustical Society of America, 66*(6), 1891-1893.

Perkell, J. S., Boyce, S. E., & Stevens, K. N. (1979). Articulatory and acoustic correlates of the [s-[s-caron]] distinction. *The Journal of the Acoustical Society of America, 65*(S1), S24.

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., et al. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *The Journal of the Acoustical Society of America, 116*(4), 2338-2344.

Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., et al. (2004). The distinctness of speakers' (s) - (sh) contrast is related to their auditory discrimination and use of an articulatory saturation effect. *Journal of Speech Language and Hearing Research, 47*(6), 1259-1269.

Polka, L., Jusczyk, P. W., & Rvachew, S. (1995). Methods for Studying Speech Perception in Infants and Children. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 49 - 89). Baltimore: York Press.

Rvachew, S. (1994). Speech perception training can facilitate sound production learning. *Journal of Speech and Hearing Research, 37*, 347-357.

Rvachew, S., & Jamieson, D. G. (1989). Perception of voiceless fricatives by children with a functional articulation disorder. *Journal of Speech and Hearing Disorders, 54*, 193-208.

Semel, E., Wiig, E. H., & Secord, W. A. (2004). Clinical Evaluation of Language Fundamentals - 4 Screening Instrument (4th ed.). San Antonio, TX: Harcourt Assessment, Inc.

Sensimetrics. (2004). HLsyn (Version 2.2). Somerville, MA: Sensimetrics Corporation.

Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009). Perceptual recalibration of speech sounds following speech motor learning. *Journal of the Acoustical Society of America, 125*(2), 1103 - 1113.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa Articulation Norms Project and Its Nebraska Replication. *Journal of Speech and Hearing Disorders, 55*, 779 - 798.

Smith, B. L. (1978). Temporal aspects of English speech production: a developmental perspective. *Journal of Phonetics, 6*, 37 - 68.

Smith, B. L. (1992). Relationships between duration and temporal variability in children's speech. *Journal of the Acoustical Society of America, 91*(4), 2165 - 2174.

Smith, B. L., & Kenney, M. K. (1994). Variability control in speech production tasks performed by adults and children. *Journal of the Acoustical Society of America, 96*(2), 699 - 705.

Smith, B. L., Kenney, M. K., & Hussain, S. (1996). A longitudinal investigation of duration and temporal variability in children's speech production. *Journal of the Acoustical Society of America, 99*(4), 2344-2349.

Smith, B. L., Sugarman, M. D., & Long, S. H. (1983). Experimental manipulation of speaking rate for studying temporal variability in children's speech. *Journal of the Acoustical Society of America, 74*, 744-749.

SPSS 14 for Windows. (2005). (2005). Chicago: SPSS, Inc., 2005.

Strange, W. (Ed.). (1995). *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press.

Tingley, B. M., & Allen, G. D. (1975). Development of speech timing control in children. *Child Development, 46*(1), 186-194.

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

Weismer, G., & Elbert, M. (1982). Temporal characteristics of "functionally" misarticulated /s/ in 4- to 6-year-old children. *J Speech Hear Res, 25*(2), 275-287.

Whalen, D. H. (1981). Effects of vocalic formant transitions and vowel quality on the English [s]--[[s-caron]] boundary. *The Journal of the Acoustical Society of America, 69*(1), 275-282.

Whitehill, T. L., Francis, A. L., & Ching, C. K.-Y. (2003). Perception of place of articulation by children with cleft palate and posterior placement. *Journal of Speech, Language, and Hearing Research, 46*, 451-461.