Developmental Acoustic Analysis of the /r/ Phoneme

by

Brittany Patricia Becvar

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved April 2017 by the
Graduate Supervisory Committee:

Tamiko Azuma, Chair
Juliet Weinhold
Visar Berisha

ARIZONA STATE UNIVERSITY

May 2017

ABSTRACT

The purpose of this study was to identify acoustic markers that correlate with accurate and inaccurate /r/ production in children ages 5-8 using signal processing. In addition, the researcher aimed to identify predictive acoustic markers that relate to changes in /r/ accuracy. A total of 35 children (23 accurate, 12 inaccurate, 8 longitudinal) were recorded. Computerized stimuli were presented on a PC laptop computer and the children were asked to do five tasks to elicit spontaneous and imitated /r/ production in all positions. Files were edited and analyzed using a filter bank approach centered at 40 frequencies based on the Mel-scale. *T*-tests were used to compare spectral energy of tokens between accurate and inaccurate groups and additional *t*-tests were used to compare duration of accurate and inaccurate files. Results included significant differences between the accurate and inaccurate productions of /r/, notable differences in the 24-26 mel bin range, and longer duration of inaccurate /r/ than accurate. Signal processing successfully identified acoustic features of accurate and inaccurate production of /r/ and candidate predictive markers that may be associated with acquisition of /r/.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Introduction

Literature Review

Historically, clinicians have relied on perceptual judgment to determine eligibility for remediation of /r/ (Klein, Grigos, McAllister Byun, & Davidson, 2012; Klein, McAllister Byun, Davidson, & Grigos 2013). The phoneme /r/ in typically developing children is among the last to be acquired in their phonetic inventory (Smit, Hand, Freilinger, Bernthal, & Bird, 1990; Sander 1972; Goldman & Fristoe, 2000; Templin & Darley, 1969). If one considers the point at which 50% of children have acquired a phoneme and the point at which 90% of children have acquired it, we could refer to this as the acquisition window for a particular speech sound. Currently, there is a large acquisition window for the phoneme /r/; this window ranges from 3 years old to as late as 8 years old depending on the norms that are implemented. Smit at al. (1990) and Templin and Darley (1969) reported data that would describe the acquisition window as between ages 3 and 8 years old. Sanders' (1972) data would identify an acquisition window between ages 3-6 years old. Finally, Goldman & Fristoe (2000) indicated that /r/ should be acquired by age 6. However, they used an 85% criteria rather than the 90% criteria used by the previously cited researchers. In an informal survey on placement criteria taken by the author, 3 school districts in the Phoenix Metropolitan area indicated the eligibility criterion for /r/ remediation was between 7 ½ and 8 years old. For children who have not mastered /r/ by age 3, clinicians typically wait to see if they self-correct by around age 8 before placing them in therapy. For children who do not self-correct, this results in a 3-5 year window in which inaccurate /r/ is habituated. This window of time that children are producing /r/ inaccurately is a significant period that could cause resistance in therapy, thus lengthening the remediation period longer (Merrell & Weinhold, 2014). In addition, the habituated inaccurate productions can persist through adulthood. Some examples of celebrities that have inaccurate /r/, and/or derhotacized /r/ production include Barbara Walters, Matt LeBlanc, and Sam Worthington. Residual speech sound disorder is an issue that could be addressed in early childhood to potentially prevent future concerns in speech. A significant factor that affects the ability to establish appropriate eligibility criteria is that clinicians do not have a way of predicting which children at the beginning of the acquisition window will self-correct and which children will not. Those who do not self-correct would eventually need to be placed in therapy.

Placement for therapy pertaining to the phoneme /r/ is currently based on perceptual judgement and phonetic transcription to determine severity and need for therapy. Perception is a subjective measure that is used to place children in therapy. Although this has worked for placement in the past, with current advances in computer software technology, researchers have opportunity to utilize these technologies to help acquire objective data that can give us a more in depth analysis of children's speech.

Glapsey & Macleod (2010) took a multi-dimensional approach when documenting the phonological change in a young boy with a phonological disorder by looking at accuracy of consonant productions, dynamic assessments, and an acoustic analysis measure. They noted that speech errors and distortions in many cases are not detected by the listener, which affects the outcome of assessment. They concluded that by combining these measures together, they were able to see gradual acquisition of target phonemes that would have been missed if they had focused on a single measure. Relating back to the placement of /r/, this suggests that use of other cues beyond a single measure can be useful in determining which children will or will not acquire /r/ on their own. Klein, McAllister Byun, Davidson, & Grigos (2013), examined multidimensional approaches to investigate children's /r/ productions using perception, ultrasound, and acoustic measures. The researchers from this longitudinal study focused on children that inaccurately produced /r/, but did not compare it to accurate /r/. Ultrasound and audio data were collected for each child. They primarily used the ultrasound technology to observe articulatory changes during intervention for /r/ and associated the changes with perceptually rated accuracy and acoustics. Major findings included significant associations between qualitative tongue-shape coding, perceptual measures, and acoustic measures in children receiving therapy for inaccurate /r/. In addition, they found that perceptual, acoustic, and articulatory properties were different between consonantal and vocalic variants of /r/. With the findings of this study and the lack of research done on acoustic characteristics of /r/ there remains a need for research that focuses on the relationships among perceptual and acoustic characteristics of /r/ phoneme produced by children. Using methods that include current technology in addition to perceptual judgment, acoustic features, and phonetic transcription will give an in depth look and aid in describing the properties of children's /r/ production in greater detail.

In the past, a range of methodologies have been used to study the phoneme /r/. These include x-ray imaging (Westbury, Hashi, & Lindstrom, 1998), magnetic resonance imaging (Zhou, Espy-Wilson, Boyce, Tiede, Holland & Choe 2008), and ultrasound imaging (Gick, B., Bacsfalvi, P., Bernhardt, B. M., Oh, S., Stolar, S., & Wilson, I. 2008). In addition, acoustic analysis has previously been used to identify that low third formant (F3) and decreased F3-F2 differences determine accurate /r/ (Chaney, 1988; Peterson and

Barney 1952; Dalston, 1975; Flipsen, Shriberg, Weismer, Karlsson, & McSweeny, 2001). When comparing inaccurate /r/ to accurate /r/, research has focused on F3 or the F3-F2 difference, but largely has neglected formant frequencies above F3 with an exception of Zhou et al (2008), who studied F4 and F5 in /r/ production. They studied subjects who produced bunched and retroflex variants of /r/ and found that the frequency difference between F4 and F5 is different between the variants. This suggests that there are acoustic cues in the signal in addition to F3 that can be used to identify /r/; more research can provide us with additional information. Formant frequency is an objective measure but has not been widely used to describe the transition of children self-correcting the phoneme /r/, nor has it been widely adopted for utilization as a remediation tool. If clinicians could determine a set of developmental acoustic markers of /r/, it could potentially lead to identification of predictive factors that could to be utilized for placement in therapy.

Acoustic analysis using signal processing technology is automatic and is produced with a reduced time frame. It may provide information beyond what is provided by F3 acoustic analysis. A literature review shows no evidence of signal processing techniques having been used to analyze /r/.

Aims of the Study

The first goal of this study was to use signal processing to identify acoustic properties that can be associated with accurately produced /ɝ/ and /r/ in children, and compare them to corresponding properties of children's inaccurate productions. The second goal of this study was to conduct a longitudinal study, by re-recording each participant every 2-3 months until they acquired /r/ or the point at which the investigation ended. The main goal of the longitudinal aspect of this research was looking to identify any acoustic markers associated with changes in the children's /r/ accuracy. Two different outcomes were anticipated: one path was that the child would self-correct on their own without any intervention needed, and in the other path the child would not self-correct, thereby continuing to produce /r/ inaccurately. Currently, there are no data that can associate certain acoustic features with the likelihood of self-correction or, conversely, non-acquisition.

Method

Participants

Children that participated in this study were recruited from private and public schools in the Phoenix metropolitan area. The accurate group included 23 children ages 5-8 (11 boys and 12 girls) who accurately produced /r/. For each phoneme position (prevocalic, vocalic, and postvocalic) participants were able to produce the phoneme accurately in at least 80% (range 80-100%) of their tokens with only the accurate tokens being analyzed. The inaccurate group included 12 children ages 5-8 (6 male, 6 female) who did not produce /r/ accurately. Children were included in this sample if their accuracy was at or below 50%, and only inaccurate tokens were analyzed. Finally, in the longitudinal group, 8 participants who produced /r/ inaccurately were followed every 2-3 months to examine changes in their /r/ production. Table 1 outlines characteristics for the participants of the longitudinal study.

Per parental report, all children were monolingual English speakers with normal hearing and cognitive development. None of the participants had received any form of speech therapy before or throughout the study. After signing the consent forms, parents completed a survey regarding the child's language, hearing, cognitive development, and the birthdate. Study participants received no compensation for their participation.

Table 1

Longitudinal Group Characteristics

Participant	Age	Gender	Number of observations	Observation period (months)	Accuracy at beginning	Accuracy at end
51001	5	Male	4	12	0%	0%
51002	5	Male	4	12	0%	0%
51003	5	Male	3	9	0%	0%
52001	5	Female	3	20	25%	96%
52002	5	Female	3	9	31%	79%
52003	5	Female	3	9	0%	42%
62001	6	Female	2	6	60%	80%
71001	7	Male	2	6	0%	0%

Design

This research is a cross sectional study of children ages 5-8 with accurate production of /r/ and inaccurate production of /r/. In addition, a longitudinal study examined children with inaccurate /r/ ages 5-7, following them every 2-3 months and recording their progress.

Procedure

Each child was tested individually in a quiet room with the experimenter being the only other person in the room. Children wore a Sennheiser PC151 Binaural Headset with noise-cancelling microphone and Wavesurfer software was used to record and edit children's speech (Sjölander & Beskow, 2004). Computerized stimuli were presented on a PC laptop computer. Children were asked to do five tasks. The first task, spontaneous speech elicitation, asked the children to describe a picture. From this, a baseline percentage accuracy was calculated. The second task, picture naming, required children to spontaneously name four pictures which contained /r/ proceeding each of the four corners of the vowel quadrilateral: /rɑk, ræbɪt, riŋ, ruf/, and four pictures which contained /r/ in postvocalic position:/ɪr, bɛr, fɔr, kɑr/ for a total of 8 pictures. The pictures were selected from clip art images from online source. Next, a consonant vowel (CV) imitation task was implemented by instructing the children to imitate the examiner. For example, the examiner would say vocalic /r/ and then would point to the child to do the same. Imitation included vocalic /r/ in isolation and in CV syllables following voiced stops: /gɝ, bɝ, dɝ/. There were three repetitions of each of these stimuli. The fourth task was a consonant vowel consonant (CVC) imitation task, in which children imitated CVC syllables with /r/ in the prevocalic position preceding corner vowels: /rip, rit, rik; rup, rut, ruk; rɑp, rɑt, rɑg; ræp, ræt, ræk/. Each word was imitated one time. The fifth task, vowel consonant (VC) imitation, required children to imitate postvocalic VC syllables following front vowels /ɪr/ and /ɛr/ and back vowels /ɔr/ and /ɑr/; each of these was repeated three times.

Thus, there were 16 tokens of /r/ in prevocalic position (4 spontaneous, 12 imitation), 12 vocalic imitation tokens, and 16 tokens of /r/ in postvocalic position (4 spontaneous, 12 imitation). See Appendices A and B for the stimuli used in the production tasks and the imitation testing protocol. If the children did not produce the targeted stimulus correctly, the examiner would implement delayed imitation. For example, if the child said “stone” for “rock”, the examiner would say, “That is a rock, can you say that?”

The responses were determined to be accurate or inaccurate based on perceptual judgement by 5 trained listeners. A certified speech language pathologist double scored 20% of all language samples and 20% of all tokens. Inter-rater reliability for language sample scoring was 94%; for tokens, it was 93%. This research had IRB approval (STUDY00001430).

Scoring

Sound files were recorded and trimmed using Wavesurfer (Sjölander & Beskow, 2004) and scored for perceptual accuracy by 5 trained listeners. Each listener removed contiguous phonemes so that only the perceived /r/ component in the utterance remained for each individual file. Because formants do not always follow stereotypic trajectory, perception was used as the final criterion for trimming. Guidelines (see Table 2) were provided as additional cues to aid the listeners in determining the start and end points of each trimmed /r/ file. While trimming each audio file, the listener marked four sets of time points: The beginning and end time of the word and the beginning and end time of the /r/. Because formants do not always follow stereotypic trajectory, perception was used as the final criterion for trimming. Guidelines (see Table 2) were provided as additional cues to aid the listeners in determining the start and end points of each trimmed /r/ file.

Table 2

Trimming Guidelines

	Vocalic /r/	Prevocalic /r/	Post-vocalic /r/
Starting point of /r/ phoneme	Onset of voicing Lowest F3 point	Onset of voicing Lowest F3 point	F3 drop F2 rise Preceding vowel ends
Ending point of /r/ phoneme	Termination of voicing F3 rise	F3 rise Where transition to vowel begins	When voicing is terminated F3 rise

Results

Analysis

2,195 trimmed files were analyzed. All files were analyzed using a filter bank approach. Filters were centered at 40 frequencies based on the Mel-scale. Mel-scale frequency is perceptually based, focuses on spectral shape, and describes the spectral energy of a given speech segment. This scale was utilized because it better captures the logarithmic range of human perception at higher frequencies. Log energy was calculated at each center frequency and Mel-frequency log spectral plots were derived.

Welch *t*-tests were used to compare spectral energy of tokens between the accurate and inaccurate groups across Mel-centered frequencies. Longitudinal files were analyzed implementing the same process described above.

Parallel analysis was conducted for duration of the edited sound files between accurate and inaccurate groups and within the individuals in the longitudinal group.

Descriptive Analysis

For both accurate and inaccurate productions of /r/, Mel-frequency spectral plots were created for prevocalic, post-vocalic, and vocalic /r/ positions. Most notable differences occurred in the 24-26 bin range in all three positions. Vocalic and postvocalic plots displayed similar curves, showing a peak of acoustic energy at bin 6, steep drop from bins 6-9, and consistent level until a slight rise at bins 21 and 22. At bin 23 there is a dip and then a second rise at bins 24-26. After bin 26, the curve for both accurate and inaccurate goes down for the continuation of the curve. For the lower frequencies, the accurate group had higher spectral energy with notable differences from the inaccurate group at bins 1-4 and a peak at bin 6. After bin 6, the curvature between the groups stays similar until reaching the higher frequency bins with the first significant difference at bin 24. The accurate group shows a peak in spectral energy at bins 24-26 and from there a steep drop between bins 26-28 putting them lower in spectral energy in comparison to the inaccurate group. The prevocalic curve looks different than the other two allophones discussed above. The curve displays more significant differences between mel bins 9-20 compared to the other two allophones. In addition, it does not have a significant difference in spectral energy between accurate and inaccurate groups at bin 27. Finally, the prevocalic curvature has a slight rise at 23 and from bin 24 on the spectral energy gradually decreases. In parallel to the vocalic and postvocalic positions, the prevocalic curve is similar between bins 26-28 with a steep drop in spectral energy. It should be noted that for all three positions, in the lower frequencies, the accurate group has higher spectral energy compared to the inaccurate group. In the higher frequencies, the inaccurate group has higher spectral energy. See Figures 1-3.

Statistical Analysis: Welch *t*-tests

Welch *t*-tests were done comparing spectral energy between the means of the accurate and inaccurate group tokens. Results are displayed below in Figure 1 (Vocalic /ɝ/), Figure 2 (Prevocalic /r/), and Figure 3 (Postvocalic /r/). Analyzing the spectral energy based on /r/ different allophones, we found that each allophone had notable differences from the other. Using Bonferroni correction for multiple comparisons, a *p* value of .00125 was necessary to identify significance. Those mel bins with *p* values below .00125 for vocalic were 1-4, 6, 24-39, for prevocalic were bins 1-5, 9-11, 15, 20, 22-26, 28-40, postvocalic were bins 2-6, 24-26, 28-39. Asterisks were used in each table to note mel bins that had significant differences in spectral energy between the accurate and inaccurate groups. T values ranged from t (299-384) = -19.9-11.48 for vocalic, t (461-468) = 18.13-7.61 for prevocalic, and t (213-266) = -11.7-8.18 for postvocalic /r/. Significant *p* values ranged from <.0002 to <4.47 x 10^{-61}.

Figure 1

Vocalic /r/ Log Mel Spectral Plot

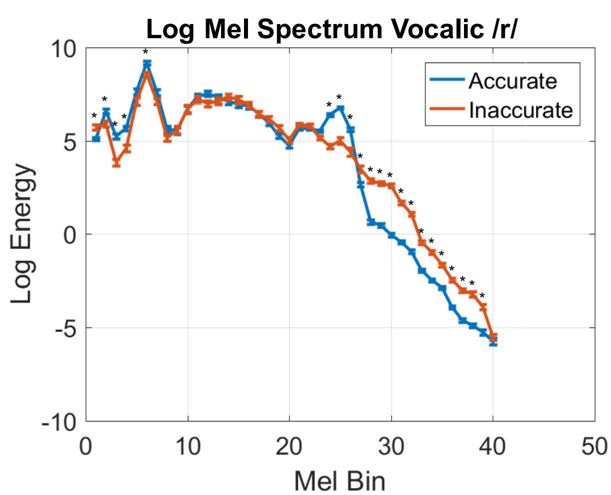


Figure 2

Prevocalic /r/ Log Mel Spectral Plot

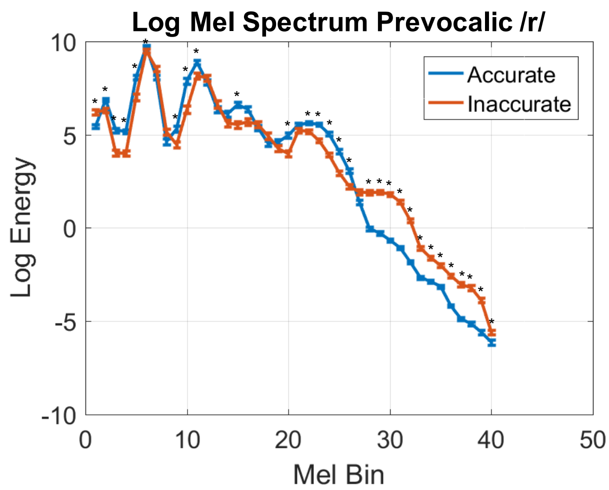
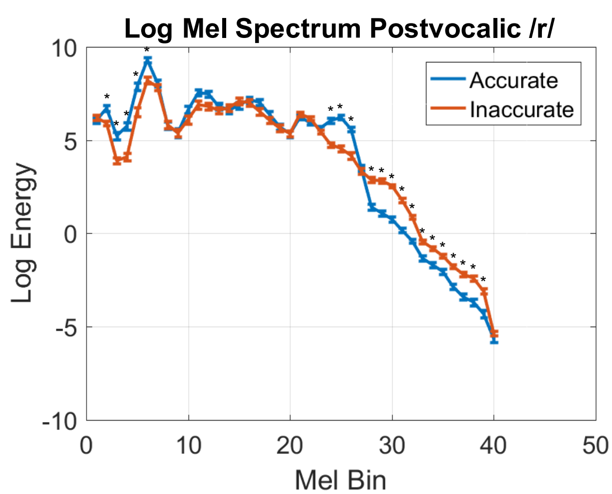


Figure 3

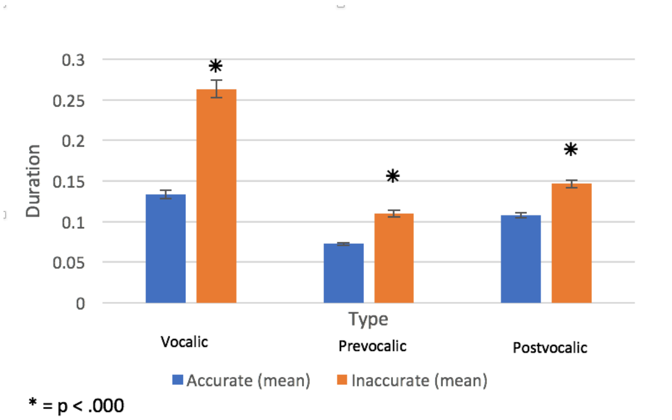
Postvocalic /r/ Log Mel Spectral Plot



Welch *t*-tests that were conducted comparing mean duration of edited sound files between accurate and inaccurate groups revealed the inaccurate group used longer duration for all three /r/ allophones compared to the accurate group. See Figure 4.

Figure 4

Mean Duration of Three Allophones of /r/



Longitudinal Results

In the longitudinal study 3 out of the 8 children displayed much greater progress than the others. The individuals (52001, 52002, and 62001) were considered children that "acquired" /r/. Although the criteria dictate 80-100% accuracy to define acquisition, subject 52002 reached 79% accuracy and was considered as part of the acquired group. The baseline and final results for the two groups (acquiring and non-acquiring) in the longitudinal study are displayed below with the baseline and final curves for each allophone. Comparison of the non-acquired group baseline to the acquired group baseline revealed significant differences in spectral energy in all allophonic variations. Range across significant mel bins are reported here. For vocalic, t (89-90) = -8.73-9.09; for prevocalic, t (110-121) = -9.68-6.26; for postvocalic, t (82-108) = -3.92-9.59. In all variations, in the lower frequencies the non-acquired groups had higher spectral energy and in the higher frequency the acquired group has higher spectral energy. Significant differences in bin numbers include: mel bins 1-4, 6, 24-39 for vocalic, 1-6, 9-11, 15, 20, 22-26, 28-40 for prevocalic, and 2-6, 24-26, 28-39 for postvocalic. Notably, at mel bin 19 there is the change in spectral energy for both groups. Overall, most of the mel bins displayed significant differences between the acquired and non-acquired groups showed comparable patterns as described above in the accuracy versus inaccurate comparisons. It should be noted that all tokens included in the baseline data set were inaccurate; any accurate tokens produced by these individuals were removed from the analysis. After baselines were compared, log mel spectrum plots were created for the individuals who acquired /r/ during the longitudinal study. For each individual, plots were made for the three allophonic variations comparing the time point at which they were inaccurate until the time they acquired /r/. See Figures 5-7 below. Results from each individual that acquired /r/ displayed similar curves. At the baseline recording, which is the point the child was inaccurate, all allophonic variations displayed lower spectral energy at the lower frequencies than the final time (accurate) the child was recorded, which is the point they were accurate. In addition, in all variations at the higher frequencies, the baseline was higher in spectral energy compared to the final time (accurate). Both of these characteristics are parallel with the comparison of low and high frequencies of the accurate versus inaccurate groups. Depending on the allophone, around mel bin 24 spectral energy shifted lower each successive time the child was recorded. See Figures 8-10. Note that asterisks are not placed on the final spectral plots, because they conform to the patterns presented earlier in the accurate vs. inaccurate results.

Figure 5

Vocalic Baseline Spectral Plot *Vocalic Final Spectral Plot*

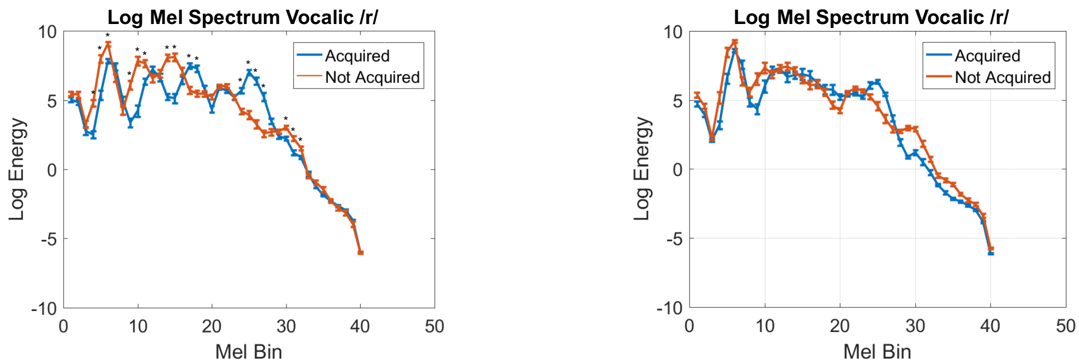


Figure 6

Prevocalic Baseline Spectral Plot *Prevocalic Final Spectral Plot*

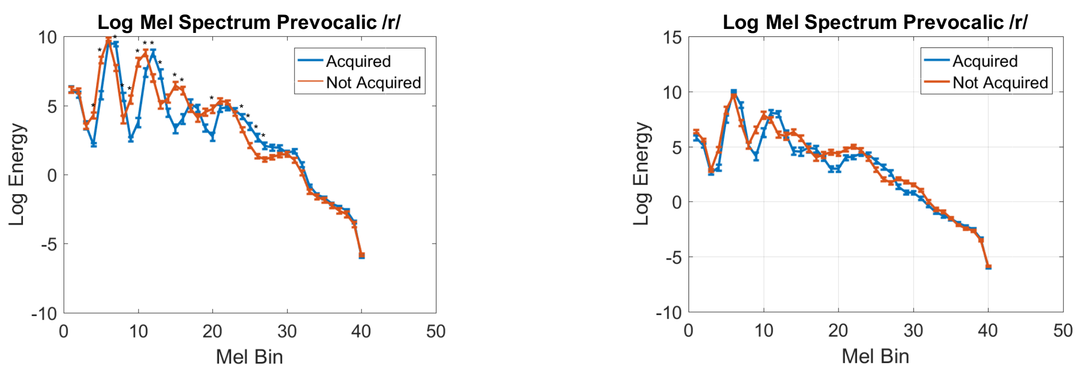


Figure 7

Postvocalic Baseline Spectral Plot *Postvocalic Final Spectral Plot*

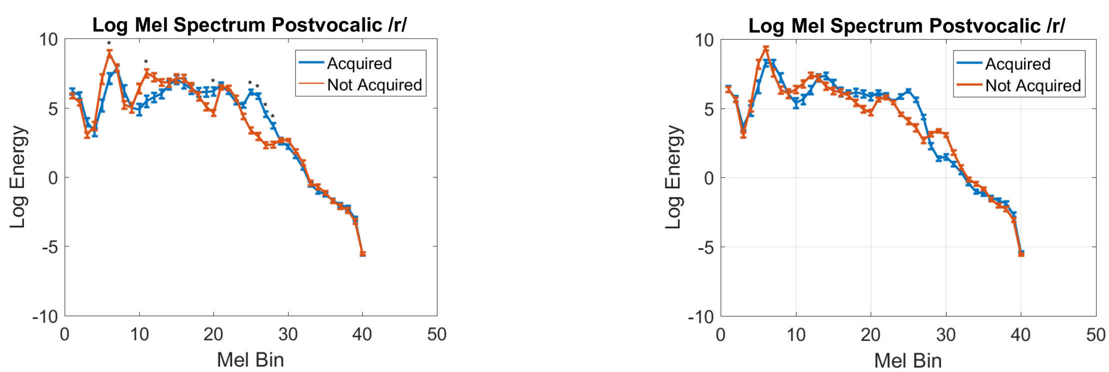
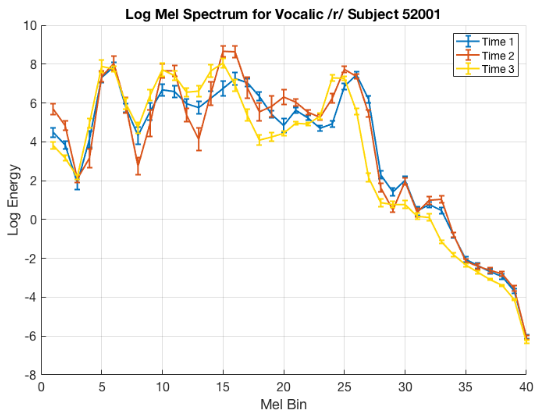
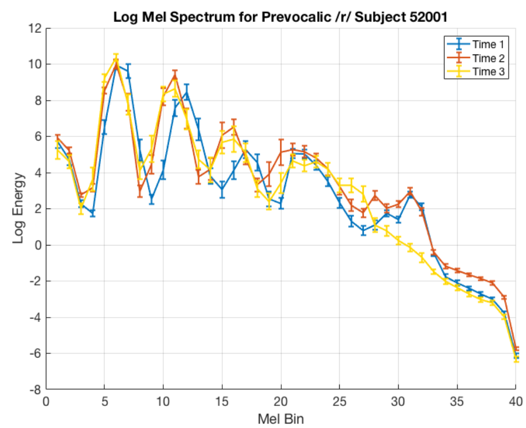


Figure 8

52001 Acquisition of /r/ Over Time





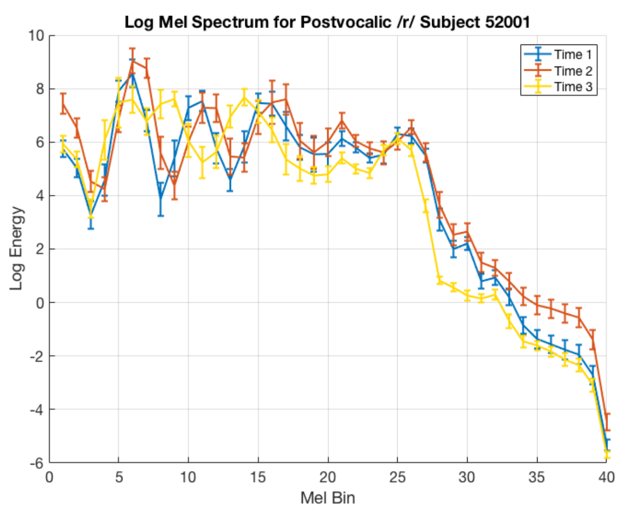


Figure 9

52002 Acquisition of /r/ Over Time

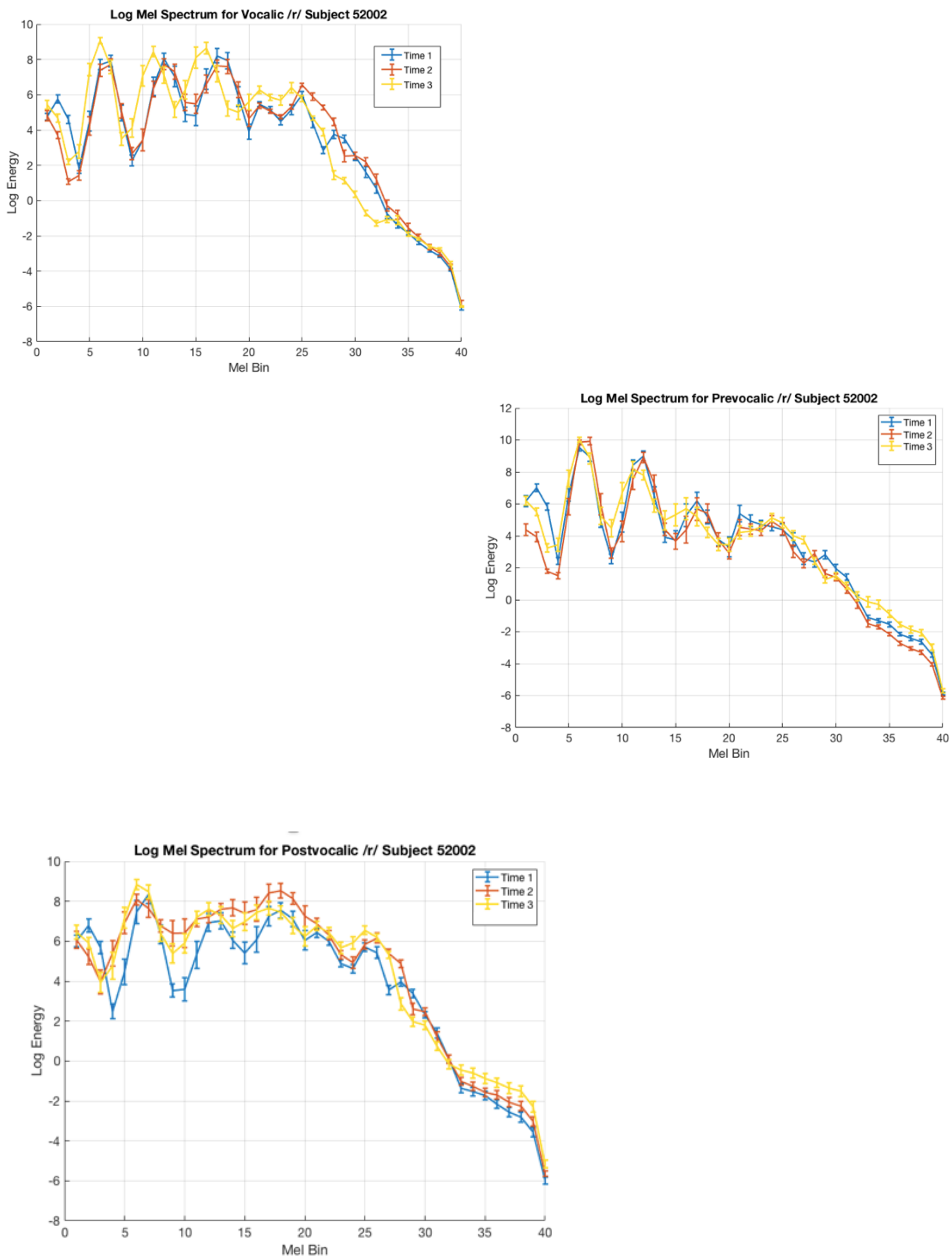
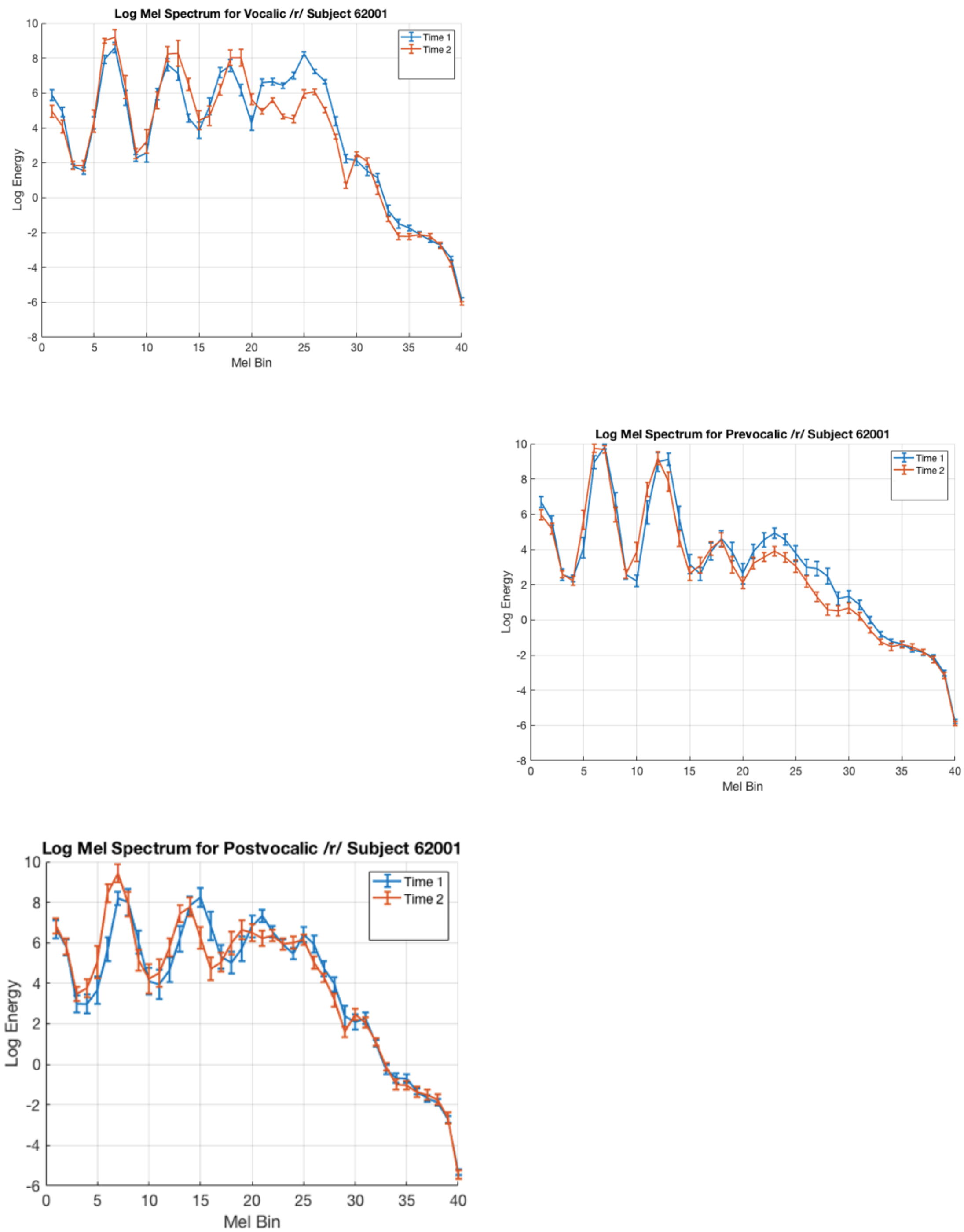


Figure 10

62001 Acquisition of /r/ Over Time



Discussion

In this study, the author first endeavored to determine if signal processing could be utilized to identify acoustic properties that can be associated with accurately and inaccurately produced /ɝ/ and /r/. In addition, a longitudinal study was conducted to identify acoustic markers that could be associated with changes in the children's /r/ accuracy. With perceptual judgment being the measure used to place children in therapy currently, use of objective measures like signal processing in addition to low third formant, can help identify children's productions of /r/ in greater depth. Just as Klein, McAllister Byun, Davidson, & Grigos (2013) used ultrasound and acoustic measures in conjunction with perception to get a more detailed look at the properties of children's /r/ production, the author was able to identify specific acoustic features using signal processing that appear to be candidate markers for differentiation of accurate and inaccurate production of /r/. This suggests that continued research using signal processing software is warranted to further describe differences and investigate acquisition of this complex phoneme.

Accurate Versus Inaccurate /r/ Production

One major finding that emerged from this research is that there is a significant difference in acoustic spectral energy between accurate and inaccurate production of /r/. Notable differences were found in bins 24-26 for all allophones of /r/. A major similarity throughout all allophones of /r/ was that comparison of the accurate to inaccurate group, revealed the accurate group had higher spectral energy at lower frequencies and the inaccurate group had higher spectral energy at the higher frequencies. Significant differences were noted in mel bins 9-20 for prevocalic /r/ compared to the other two allophones. This suggests that prevocalic /r/ is qualitatively different than the other allophones. Although vocalic /r/ and postvocalic /r/ display similar mel log spectral plots, acoustic energy for mel bins 24-26 is higher in vocalic than postvocalic allophones. Thus, each allophonic variation of /r/ can be considered to be different from one another and rather than being considered as one phoneme should instead be examined individually.

Changes in Duration of /r/ Between Allophone Groups

Comparison of the duration of time for the different allophonic variations revealed that, all three positions for inaccurate /r/ group had longer duration than the accurate group. Thus, signal processing provided another marker that differentiates between an accurate and inaccurate /r/.

Longitudinal Findings

Comparison of the acquired group's baseline recordings revealed significant differences from that of the group that did not eventually acquire /r/. This suggests that frequencies in which there are significant spectral energy differences could be considered as potential predictive markers for children who will self-correct. These differences in spectral energy are notably similar to the differences in the accurate versus inaccurate groups in terms of spectral energy at different frequencies. This is an interesting finding because both the acquired and not acquired groups were completely inaccurate at baseline yet they still had differences in spectral energy, suggesting that those differences could possibly be used as candidate predictive markers. When studying the three children who acquired /r/ over time, differences were found from the baseline recording time to the last recording when they acquired /r/. In fact, their baseline recording displayed similar spectral energy to the inaccurate group and their final time displayed similar energy to the accurate group.

Conclusion

There is currently no known method of predicting which children within the acquisition window for the phoneme /r/ will eventually self-correct. There is a lack of methods to determine eligibility for therapy beyond clinical perception. Because children who do not self-correct will typically not receive therapy during the 3-5 year acquisition window, they are at high risk for habituation of their inaccurate production, and carry increased risk of resistance to therapy. In this study, the author sought to address this issue by utilizing current technology to see if there are acoustic markers for accurate /r/, inaccurate /r/, and to determine predictive markers of acquisition of /r/. This was done by recording children whose ages were within the acquisition window for /r/ (3-8 years old) who accurately produced /r/, and children who inaccurately produced /r/. In addition, by identifying children who were in the middle of this acquisition window and following them over time, an attempt was made to identify markers that may be predictive of non-acquisition of /r/ using acoustic signal processing. First, the inaccurate and accurate groups were compared, and as hypothesized, a set of mel bins were identified as acoustic features that differentiate accurate and inaccurate production of /r/. The most salient differences were noted for mel bins 24-26. In addition, inaccurate /r/ was associated with longer durations compared to the accurate group for all allophones examined. When looking at the longitudinal data, the baseline profile for the acquiring group was significantly different than the baseline profile of the non-acquiring group. Differences in acoustic energy between the groups at baseline for particular mel bins suggests that these values may be considered candidate predictive markers for children who will acquire /r/ on their own. Specifically, mel bins 21-30 appeared to be predictive for vocalic /r/, mel bins 24-26, 31 for prevocalic, and bins 27-30 for postvocalic. With the exception of these candidate predictive markers, the three children that acquired /r/ over time displayed curves similar to the accurate group at their final recording. Acoustic signal processing technology provides a deeper look into the acoustic features of /r/ and enables us to identify specific acoustic features associated with accurate /r/, inaccurate /r/, and changes in /r/ acquisition. The identification of predictive markers, namely specific mel bins and possibly duration, is very promising in terms of facilitating more informed determination of eligibility for placement.

Limitations of Current Study and Further Research

A limitation of the current research is the low number of subjects in the longitudinal study. There were only 3 out of 8 subjects who were considered to have acquired /r/ over the time they were recorded. It is also noted that only females self-corrected /r/ in this study and it would be beneficial to have subjects of both gender who self-correct in the future studies. A longitudinal study with a larger sample of participants with a range of ages and gender to follow longitudinally would provide a deeper and more detailed look into possible predictive markers of changes in /r/ accuracy. Not withstanding the low number of subjects in this study, the data suggest that there are differences at baseline in children who self-correct and those who do not.

A second limitation of the longitudinal section is that the 80% accuracy cut off for determining acquisition did not allow for analysis of the subjects who, although they did not reach criterion, did indeed show some improvement in their articulation of /r/. For example, subject 52003 was a member of the non-acquiring group. However, by the end of the data collection period, her accuracy had improved from 0% to 42%. This child could be considered transitional, and it would be interesting to look at her data over time and analyze what specific changes were underway. Future longitudinal study could identify a third group, transitional, whose data could be compared to both the acquiring and non-acquiring profiles. The longitudinal part of this study serves as a compelling pilot that justifies further acoustic analysis of children within the acquisition window, using a longitudinal perspective.

References

Chaney, C. (1988). Acoustic analysis of correct and misarticulated semivowels. *Journal of Speech and Hearing Research,* 31, 275-287.

Dalston, R. M. (1975). Acoustic characteristics of English /w, r, l/ spoken correctly by young children and adults. *The Journal of the Acoustical Society of America, 57*(2), 462-469.

Flipsen, P., Shriberg, L.D., Weismer, G., Karlsson, H.B., & McSweeny, J.L. (2001). Acoustic phenotypes of speech-genetics studies: Reference data for residual /ɝ/ distortions. *Clinical Linguistics & Phonetics, 15*(8), 603-630.

Gick, B., Bacsfalvi, P., Bernhardt, B. M., Oh, S., Stolar, S., & Wilson, I. (2008). A motor differentiation model for liquid substitutions. English /r/ variants in normal and disordered acquisitions. *Proceedings of Meetings on Acoustics, 1,* 1-9.

Glapsey, A. M., & Macleod, A. A. (2010). A multi-dimensional approach to gradient change in phonological acquisition: A case study of disordered speech development. *Clinical Linguistics & Phonetics, 24,* 283-299.

Goldman, R., Fristoe, M. (2010). *Goldman-Fristoe Tests of Articulation- Second Edition.* Circle Pines, MN: Pearson Assessments. (GFTA-2).

Klein, H. B., Grigos, M. I., McAllister Byun, T., & Davidson, L. (2012). The relationship between inexperienced listeners' perceptions and acoustic correlates of children's /r/ productions. *Clinical Linguistics & Phonetics, 26*(7), 628-645.

Klein, H. B., Byun, T. M., Davidson, L., & Grigos, M. I. (2013). A multi-dimensional investigation of children's /r/ productions: perceptual, ultrasound, and acoustic measures. *American Journal of Speech-Language Pathology, 22*, 540-553.

Merrell, B. and Weinhold, J. (November, 2014). *Evaluation of Placement Criteria: How Long Does /r/ Remediation Take?* Poster presented at American Speech, Language, and Hearing Association Convention, Orlando, FL.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The jJournal of Acoustical Society of America, 24*(2), 175-184.

Sander, E. K. (1972). When are speech sounds learned? *Journal of Speech and Hearing Disorders, 37*(1), 55-63.

Sjölander, K. & Beskow, J. (2004). Wavesurfer. (Version 1.8.8p4-win-i386.zip) [Computersoftware]. Retrieved from: sourceforge.net/projects/wavesurfer/.

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*(4), 779-798.

Templin, M.C., & Darley, F. L. (1969). *The Templin-Darley Tests of Articulation* (2nd ed.). Bureau of Educational Research and Service, Division of Continuing Education. Iowa City: The University of Iowa.

Westbury, J., Hashi, M., Lindstorm, M. (1998). Differences among speakers in lingual articulation of American English. *Speech Communication, 26,* 203-226.

Zhou, X., Espy-Wilson, C. Y., Boyce, S., Tiede, M., Hollard, C., & Choe, A. (2008). A magnetic resonance imaging-based articulatory and acoustic study of “retroflex” and “bunched” America English /r/. *The Journal of Acoustical Society of American, 123*(6), 4466-4481.

APPENDIX A

IMITATION PROTOCOL

Pre/post test screening protocol: **/r/** Date _____________

Participant number: _________________

Imitation of vocalic /r/ in isolation ____ ____ ____

Imitation of: /brrr/ ___ ___ ___ /drrr/ ___ ___ ___ /grrr/ ___ ___ ___

Imitation of:
rip ___ ___ ___ rit ___ ___ ___ rik ___ ___ ___
roop ___ ___ ___ root ___ ___ ___ rouk ___ ___ ___ [use /u/ vowel]
ropp ___ ___ ___ rott ___ ___ ___ rogg ___ ___ ___ [use /a/ vowel]
rap ___ ___ ___ rat ___ ___ ___ rak ___ ___ ___ [use /ae/ vowel]
/ir/ ___ ___ ___ /ɛr/ ___ ___ ___ /ɔr/ ___ ___ ___ /ɑr/ ___ ___ ___

PICTURE NAMING
rock ____ rabbit ____ ring ____ roof ____
[Cue if needed: "What part of the house is on fire?"]
ear ____ bear ____ four ____ car ____

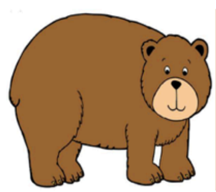
APPENDIX B

SPONTANEOUS PICTURE STIMULI







4



APPENDIX C

LANGUAGE SAMPLE STIMULUS PICTURE