DISENTANGLING FREQUENCY AND TALKER VARIABILITY IN A STATISTICAL LEARNING TASK INVESTIGATING CHILD SPEECH ACQUISITION

By

AMANDA K GOOD

Bachelor of Arts in Linguistics

McGill University

Montreal, QC

2014

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
May, 2017

DISENTANGLING FREQUENCY AND TALKER VARIABILITY IN A STATISTICAL LEARNING TASK INVESTIGATING CHILD SPEECH ACQUISITION

Thesis Approved:

Dr. Peter Richtsmeier, Ph.D.

Thesis Adviser

Dr. Ramesh Kaipa, Ph.D.

Dr. Sabiha Parveen, Ph.D.

Name: AMANDA K GOOD

Date of Degree: MAY, 2017

Title of Study: Disentangling frequency and talker variability in a statistical learning task investigating child speech acquisition

Major Field: COMMUNICATION SCIENCES AND DISORDERS

Abstract:

Past research suggests that statistical patterns in a child's native language influence the child's speech production accuracy (Stoel-Gammon, 1998; Plante, Bahl, Vance, & Gerken, 2011). In addition to hearing some forms more often than others, children also hear forms from a variety of talkers in their environments (parents, caretakers, etc.). This variable, known as talker variability, can have facilitative effects on children's ability to reproduce nonwords (Plante et al., 2011; Richtsmeier, Gerken, Goffman, and Hogan, 2009).

To further investigate frequency and talker variability, the current study employed a within-subjects design to expose 3-4 year olds to four levels of experimental frequency, with and without talker variability. The results of this study suggest a benefit for perceptual frequencies greater than 1, and for production practice. Benefits from talker variability should not be ruled out, but may be less robust than basic perceptual frequency.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1

II. METHODOLOGY 6

Participants 6
Materials 7
Procedure 9
Analysis 13

III. RESULTS 16

Production Accuracy 16
Semantic Learning 19

IV. DISCUSSION 21

Production Accuracy 21
Semantic Learning 23
Clinical Significance 23
Limitations 24

V. CONCLUSION 26

REFERENCES 27

APPENDICES 29

LIST OF TABLES

Table Page

I 8
II 15

LIST OF FIGURES

Figure Page

1 10
2 11
3 13
4 17
5 18
6 18

CHAPTER I

INTRODUCTION

Past research has suggested that statistical patterns in a child's native language influence the child's speech production accuracy. In particular, the frequency of forms has been shown to affect production accuracy, with more frequently heard forms generally being produced more accurately (Messer, 1967; Munson, 2001; Plante, Bahl, Vance, & Gerken, 2011). These forms are often sound patterns within words, typically referred to as phonotactics. Messer (1967) investigated preschool-aged children's ability to pronounce English-like and non-English-like phonotactic patterns in nonwords within a repetition task. Nonwords were considered to be non-English-like due to violations of English phonotactic constraints. "Rdek" would be an example of a nonword that violates English phonotactic constraints because English does not allow /rd/ to be located at the beginning of a word. The children were accurate in discriminating which nonwords were English-like and which were non-English-like. The children were also less accurate in producing the non-English-like nonwords than the English-like nonwords.

More recently, researchers have turned their attention toward the frequency of phonotactic patterns or sequences within a language. Munson (2001) examined the effects of English phonotactic pattern frequency on production accuracy and production duration in preschool children though a word repetition study. Phonotactic pattern frequency was measured by calculating the pattern’s frequency in an electronic corpus of English words (Hoosier Mental Lexicon; Pisoni, Nusbaum, Luce, & Slowiacek, 1985) and with subjective wordlikeness ratings of the stimuli nonwords. The children repeated infrequent sequences (sequences with a low corpus frequency and low wordlikeness rating) less accurately and produced them with longer durations than frequent sequences (sequences with a high corpus frequency and high wordlikeness rating).

Coady, Evans, and Kluender (2010) investigated which factors aid children, ages 7 to 10 years, with specific language impairment (SLI) and typically developing children to produce novel phonological patterns. The children completed a nonword repetition task that manipulated phonotactic frequency. Each nonword varied by a single manipulation, either consonant frequency or phoneme co-occurrence frequency. The children reproduced the nonwords and their repetitions were scored in terms of accuracy, onset-to-onset reaction time, and duration. The authors found that typically developing children and children with SLI were significantly more accurate when producing sounds in shorter words and in phonotactically frequent sound combinations. Similar to the Munson (2001) results, Coady et al.’s results suggest that all children benefit from increased phonotactic frequency when reproducing words.

Focusing on perceptual experience, children likely hear high frequency sequences more often. Related to that, children also likely hear frequent words and phonotactic sequences from a variety of talkers in their environments, including their parents, siblings, and caretakers. This variable, known as talker variability, can have facilitative effects on adults learning a second language (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1997; Wang, Jongman, & Sereno, 1999) and children’s ability to reproduce nonwords (Plante et al., 2011; Richtsmeier, Gerken, Goffman, and Hogan, 2009).

Plante, Bahl, Vance, and Gerken (2011) demonstrated that both phonotactic frequency and talker variability can affect children's production of sound sequences. The authors presented nonwords to children with SLI and typically developing children. These nonwords were systematically manipulated for English phonotactic frequency and the frequency of presentation within the experiment, or experimental frequency. For example, /bɑktem/ was highly frequent in the experiment and /kt/, the medial consonant sequence, is frequent in English. On the other hand, /fɑʃpem/ was highly frequent in the experiment but /ʃp/ is infrequent in English. Experimental frequency was counterbalanced in this study so that some nonwords, like /bɑktem/, were of high experimental frequency for some children and low experimental frequency for others. For words that the children heard in the high experimental frequency condition, words were presented ten times in a different talker's voice each time, utilizing the concept of talker variability. Words in the low experimental frequency condition were presented only once, and as a result, by only one talker. The authors found that both the manipulations of phonotactic frequency and experimental frequency affected the accuracy and response time for nonword production for both typically developing children and children with SLI. High English frequency and high experimental frequency increased production accuracy and reduced response times in both groups. These findings suggest that simple manipulations of the child's input can affect a child's representation of word forms, and this can facilitate production with only short-term exposure and little articulatory practice. One limitation of this study is that, since all high experimental frequency nonwords were presented by multiple talkers, the effects of experimental frequency and talker variability cannot be teased apart.

It is important to note that the benefits from talker variability in phonological learning have been inconsistently observed. In an adult second language learning study, Lively, Logan, and Pisoni

(1993) found that Japanese speakers failed to generalize discriminations of /l/ and /r/ when exposed to a new talker. However, in a similar experiment, Wang, Spence, Jongman, and Sereno (1999) found that English speakers were more accurate in identifying Mandarin tones when talker variability was high and showed increased generalization to new speakers. The uncertainty of the role of talker variability in speech learning and acquisition warrants further research.

To investigate the effects of experimental frequency and talker variability on speech learning, Richtsmeier, Gerken, Goffman, and Hogan (2009) exposed typically developing four-year-olds to nonwords. The children would hear some words ten times and some words only once. Of the words that were heard ten times, some words were presented by only one speaker and some were presented by ten different speakers. These researchers found no significant effect of experimental frequency when all words were presented by one speaker. However, when frequent words were presented by ten talkers, children produced them with higher accuracy and shorter production latencies. This suggests that talker variability is a significant factor in perceptual learning.

In sum, previous research on production development suggests that phonotactic frequency, perceptual experimental frequency, and talker variability all have facilitative effects on children's production speed and accuracy. However, a binary model in which some words are heard rarely (once) and some are heard ten times does not accurately represent a child's linguistic experience. To further investigate the effects of experimental frequency and talker variability, children should be exposed to more levels of experimental frequency to better reflect the role of talker variability in natural language learning. In the following experiment, typically developing children were exposed to nonwords that were presented ten times, six times, three times, or once. Thereby creating a more ecologically valid experimental design in regards to both experimental frequency and talker variability.

The following study aims to answer if talker variability and experimental frequency can affect children's ability to produce consonants in nonwords. Due to the findings of the previously discussed studies, the researchers hypothesize that multiple talkers and increased experimental frequency will facilitate the children's accuracy.

CHAPTER II

METHODS

Participants

Children between the ages of 3;0 and 4;8 (*M* = 3;9.6) were recruited for the experiment. Children were recruited through daycares and preschools in the Stillwater, Oklahoma area. Fliers were left at the preschools and parents voluntarily contacted the researchers. Children were also recruited through advertisements in local newspapers. All children were monolingual native English speakers and met criteria for typical development. The researcher administered the *Goldman-Fristoe: Test of Articulation – 2* (GFTA-2) to all participants and the children yielded an average score of 113. No children with a *GFTA-2* score of 85 or below participated in the experiment. Children who had a history of failing hearing exams, had any neurological or developmental problems, had a history of speech or language problems, or had a history of speech-language therapy, were excluded from the experiment by a pre-experimental phone screening. Two children were excluded in this way. Another six participants were unable to complete the experiment. The remaining 32 participants, 17 male and 15 female, were included in the analysis. All included participants passed a pure tone hearing screening bilaterally at 1000, 2000, and 4000 Hz at 25dB.

Materials

Materials included eight CVCCVC nonwords (Munson, 2001; Richtsmeier et al., 2009). The eight words were: /pɛmtəs/, /niʃkət/, /mæfpəg/, /fugdən/, /sabləf/, /tʌvʧəp/, /bozjəm/, and /gɪsnək/. Thirteen adult female speakers of English were recorded producing the nonwords. Sixteen sounds were included in the word medial consonant sequences, and no sounds were repeated in more than one word-medial sequence. For example, the word medial sequence in /pɛmtəs/ is /mt/, so /m/ and /t/ do not appear in the medial sequence of the other seven nonwords. All words had word-medial CC Biphone Probability of .0015 or lower and had no phonological neighbors according to an online corpus search. This criteria has been used by past studies (Richtsmeier, et al., 2009; Plate et al. 2011). As noted above, highly frequent English sound sequences are more accurately produced by children (Munson, 2001; Cody, Evans, & Kluender, 2010; Plante et al., 2010). Therefore if the English frequency of the sound sequences composing the nonwords are low, more errors will be made to be analyzed. The eight words were therefore relatively similar in terms of their phonotactic properties. Table I below provides a summary of the phonotactic probabilities and neighborhood densities of the experimental words.

Table I: Phonological properties of the experimental words and the word-medial consonant sequences. All words had word-medial CC Biphone Probability of .0015 or lower. Phonological neighborhood information is also given. Syllable neighbors and neighbor syllabic ratio were not controlled, but number of neighbors was consistent. All nonwords had zero neighbors.

Consonant Sequence	Shriberg Categories	Word	CC Biphone Probability	Neighbors (IPhOD)	1st Syllable Neighbors	2nd Syllable Neighbors	Neighbor Syllable Ratio
sn	late-early	gɪsnək	.0003	0	9	2	4.5
zj	late-early	bozjəm	.0000	0	12	0	-
vʧ	mid-mid	tʌvʧəp	.0000	0	9	3	3
bl	early-late	sabləf	.0015	0	8	4	2
gd	mid-early	fugdən	.0001	0	3	6	0.5
fp	mid-early	mæfpəg	.0000	0	16	3	5.33
ʃk	late-mid	niʃkət	.0000	0	7	8	0.875
mt	early-mid	pɛmtəs	.0002	0	10	1	10

To manipulate the variable of experiment frequency, the words were divided into four experimental frequency conditions that differed in the number of exposures of the nonword. Two of the nonwords were heard ten times, two were heard six times, two were heard three times, and two just once. Another variable the experiment manipulated was talker variability, divided into single talker and multiple talker conditions. Either a single talker or multiple talkers presented each nonword. Experimental frequency and talker variability were crossed within the experiment. It should be noted that since experimental frequency varied from one to ten, talker variability also varied as a consequence of experimental frequency. For the two words in the experimental frequency levels of three, six, and ten, one was in the single talker condition and the other was in the multiple talker condition. Therefore, a nonword in the multiple talkers condition and the experimental frequency level of ten, would be presented by ten talkers, and a nonword in the single talker condition and the experimental frequency level of three would be presented by one talker three times. This can be seen in Figure 2. By default, the two words in the experimental frequency level of one were heard by a single talker.

Procedure

Participants were brought in for a single experimental session by their parents. The experiment took place in a quiet room. Children sat at a child-sized table with the computer screen with speakers on either side. Speaker volume was set to a comfortable level and was consistent across all participants. Presentation of the experiment was controlled by Paradigm computer software (http://www.paradigmexperiments.com/). The experimenter sat to the left of the child, and provided directions and reinforcement. The child controlled the pace of the experiment by touching the computer screen or clicking the mouse, which would bring about a new repetition of the target word or an opportunity for the child to reproduce the target word, as shown in Figure 1. The experimenter could also move the experiment forward if the child preferred to passively listen. However, only the experimenter controlled when one word set ended and a new set began. A set, for the purpose of this experiment, was one, three, six, or ten repetitions of one nonword and one production of the nonword from the child, as seen in Figure 2.

The experimenter explained to the child before starting the experiment that he or she would play a game concerning a set of make-believe animals. A colorful hand-drawn picture of a make-believe animal accompanied each exposure of the nonwords (Ohala, 1999).

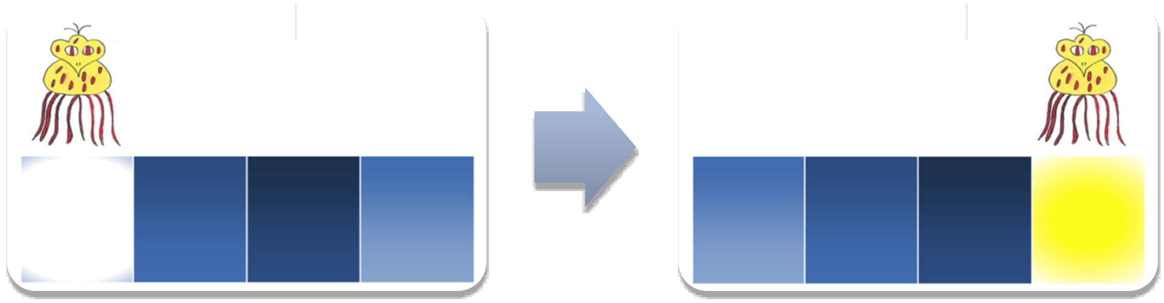

Figure 1: Above depicts a set from the first block of the experiment. There were three repetitions of the target word, as depicted by the first three blue boxes, and one opportunity for the child to produce the target word, as depicted by the last blue box. As each repetition was produced the corresponding blue box turned white and the picture lined up with the box below it. When it was time for the child to produce the target word the last blue box turned yellow.

The experiment was comprised of four blocks. The first block will be referred to as the test because it contained the primary manipulations of experimental frequency and talker variability. As shown in Figure 2 below, the test manipulated experimental frequency and talker variability, with experimental frequency represented by number of ears and talker variability represented by the color of the circles. The test consisted of sets in which the child heard each nonword a variable amount of times while looking at the make-believe animal paired with that nonword. The child produced each nonword after hearing adult productions matching the word's experimental frequency level. For example, /gɪsnək/ was presented six times by six different talkers while /pɛmtəs/ was presented ten times by a single talker. Nonwords retained their experimental frequency and talker variability conditions across a set within the same experiment. Each word was produced by the child a total of three times during the test (separate from the number of times a child heard it or the number of talkers who produced it).

Figure 2 – The above is an illustration of a portion of the test condition for List 1. Each circle with an ear represents an exposure, while each speaking child represents an opportunity for the participant to produce the nonword. The color of the circles represents different talkers. For example, /gɪsnək/ is heard six times (because six ears appear below /gɪsnək/) and is presented by multiple talkers (each circle under /gɪsnək/ is a different color).

To avoid confounding sequence and experimental frequency, the items were distributed across eight lists, with each word appearing as the experimental frequency level of ten in two lists, six in two lists, three in two lists, and one in the last two lists.

Following the test, the second experimental block was an ABX discrimination task in which the child heard a minimal pair that differed by one feature on one consonant phoneme. An example of this would be /bæd/ and /dæd/, where the first consonants, /b/ and /d/, differ only in the feature of place of articulation. Each word was paired with an appropriate picture. The pictures and word sets were presented one at a time and then the target word was verbally presented once more. The child was told to point to the picture he or she had heard. The purpose of this task was to determine if the child could distinguish the 16 sounds used in the nonword consonant sequences from other easily confusable sounds.

The third block aimed to investigate semantic learning through a set of word form-to-animal semantic probes. The experimenter presented the child with pictures of three of the make-believe animals and asked which animal was associated with a nonword that the experimenter produced. For example, the experimenter might ask, “Which animal is pemmtes?” The child would then point to the animal that he or she believed was associated with the presented nonword. No emphasis was placed on learning the association between words and animals during or before the test condition. Thus, to the extent that children learned the associations, they did so implicitly.

The fourth and final block consisted of a production posttest in which each nonword was presented once and the child reproduced the word. Each nonword set was produced twice. This is illustrated in Figure 3 below. Therefore, by the end of the posttest, the child had the opportunity to say each nonword five times.

Figure 3 – The above is an illustration of a portion of the posttest condition for List 1. Each circle with an ear represents a perceptual exposure, while each speaking child represents an opportunity for the participant to produce the nonword. The color of the circles represents different talkers. Unlike in the test condition (Figure 2), each word is heard only once before being produced.

The experimenter continued the experiment only when the child indicated that she or he was ready. The child's parents observed the experiment in the experiment room. They did not help the child with the experiment but occasionally encouraged the child if she or he became unhappy or distracted. The experiment stopped if at any time the child communicated an unwillingness to continue. The child was given a small prize when the experiment ended.

Analysis

Each nonword was produced three times in the test and twice in the posttest, yielding five total productions of each nonword. Each production of each nonword was transcribed phonetically and then scored on a 3-point scale. A score of 3 indicated an optimal production in which all features of voicing, manner of articulation, and place of articulation were accurately produced. A score of 2 indicated that one feature was not produced accurately but the two other features were correctly produced. A score of 1 indicated that two or more features were incorrectly produced but an attempt was made at producing the sound. A score of 0 indicated that the sound was deleted. The scores of all four consonant sounds in the nonword were summed and each word production was given a score out of 12 possible points. For example, if the target word was /gɪsnək/ and the child produced [gizkə] then the child would receive 3 points for a correct production of /g/, 2 points for an incorrect production of /s/ that differed only in voicing, 1 point for an incorrect production of /n/ that differed in place, manner, and voicing, and 0 points for no attempt at the word-final /k/. These scores would be totaled and the child would receive 6 points for his or her production. This scoring system has been used in similar experiments such at the Richtsmeier et al. (2009) study.

For the semantic learning probes, a score of 1 or 0 was given for each correctly matched animal and nonwords. However, the semantic probes were designed to allow for conclusions based on the experimental frequency and talker variability variables. Target animals (the animals representing the correct answer for a given probe) always had an experimental frequency level of three, six, or ten. For one probe, children heard one target animal produced by a single talker, and the competitors were other single talker animals with different frequencies. For another probe, the target animal was produced by multiple talkers, as were the competitors, and each option had a different experimental frequency. The third target animal varied between the single talker and multiple talker conditions across lists. For half of the lists, the target animal was a multiple talker word, and it was paired with a competitor animal with the same experimental frequency from the single talker condition. The second competitor had a frequency of one. For the other half of the lists, the target animal was a single talker word, but the competitors were equivalent. Using this setup, the results of the semantic probes were analyzed in terms of the questions set out below in Table II.

Table II – The above table breaks down the possible questions of significance in the block 3-semantic learning probe task. Semantic Probe T-Test Values and Interpretation.

Question	*t*	Mean Accuracy	SD	*p*
Did children respond significantly above chance overall to all three questions?	3.92	.52	.28	.000*
Did children respond significantly above chance when the target animal was experimental frequency 3?	1.89	.50	.51	.068
Did children respond significantly above chance when the target animal was experimental frequency 6?	2.10	.50	.46	.044
Did children respond significantly above chance when the target animal was experimental frequency 10?	3.32	.58	.42	.002*
Did children respond significantly above chance when the target animal was produced by multiple talkers?	2.36	.48	.37	.025
Did children respond significantly above chance when the target animal was produced by a single talker?	3.50	.56	.38	.001*
Did children respond significantly above chance when the target animal was produced by a single talker, and one of the competitors was an item with the same experimental frequency produced by multiple talkers?	1.89	.50	.51	.068
Did children respond significantly above chance when the target animal was produced by multiple talkers, and one of the competitors was an item with the same experimental frequency produced by a single talker?	2.61	.56	.50	.014

CHAPTER III

RESULTS

Production Accuracy

Effects of all variables on mean production accuracy can be seen in Figure 4 below. In the examination of production accuracy, the factors of test condition, experimental frequency, and talker variability were analyzed. Two analyses of variance were performed to analyze the data. The first ANOVA analyzed the effects of test condition and experimental frequency [1, 3, 6, 10] on production accuracy. The distribution of talker variability was only relevant to the three, six, and ten levels of the experimental frequency factor, so to analyze experimental frequency in its entirety, it made sense to leave out talker variability in this first analysis. The second ANOVA analyzed the effects of test condition, experimental frequency [3, 6, 10], and talker variability on production accuracy. Because both single talker and multiple talker experimental frequency levels of one would both ultimately be presented by one talker, it made sense to leave the experimental frequency level of one out of this ANOVA. Sphericity violations were found in both ANOVAs and Huynd-Feldt Corrections were conducted to address the violations.

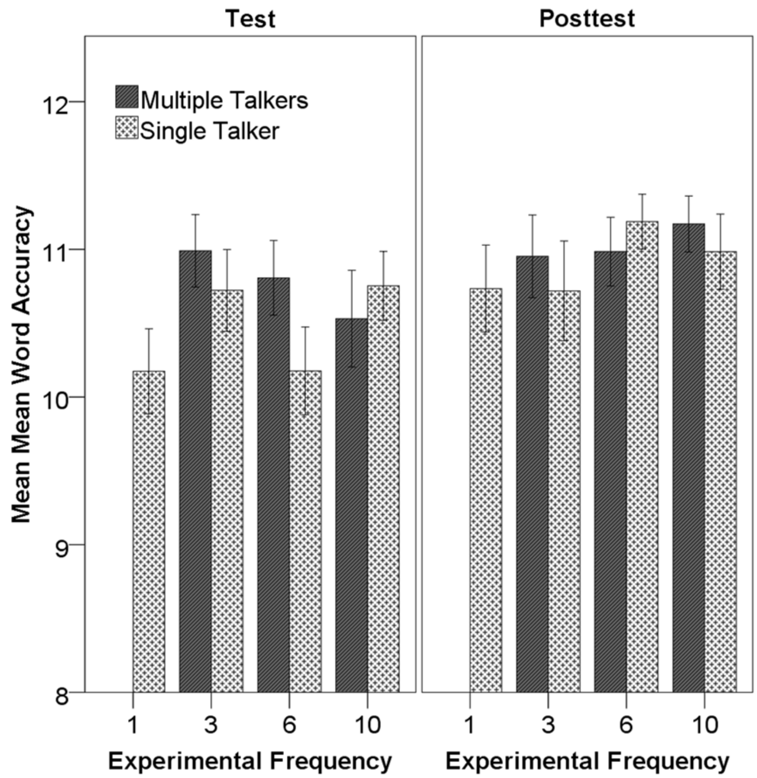

Figure 4 – Mean word accuracy (y-axis) is broken down by factors of experimental frequency (x-axis) and talker variability (multiple talkers represented by dark grey bars and single talk represented by light grey bars). The test condition is shown in the left panel and the posttest condition is shown in the right panel.

The production accuracy scores were first examined in a 2×4 (Test Condition [test, posttest] × Experimental Frequency [1, 3, 6, 10]) ANOVA. A significant main effect of test was found, with children performing significantly more accurately in the posttest than test condition, $F(1, 32) = 12.30$, $p = .001$. The mean word accuracies for the test conditions can be compared in Figure 5 below. A main effect of experimental frequency was not found to be significant, $F(2.50, 79.80) = 1.61$, $p = .201$. A near-significant trend was found in the interaction between test and frequency $F(2.03, 65.07) = 3.06$, $p = .053$. To explore this interaction, simple effects ANOVAs were conducted. A significant effect of frequency was found in the test condition, $F(7.76, 83.62) = 3.343$, $p = .028$, but not in the posttest condition, $F(2.40, 76.44) = 1.000$, $p = .384$.

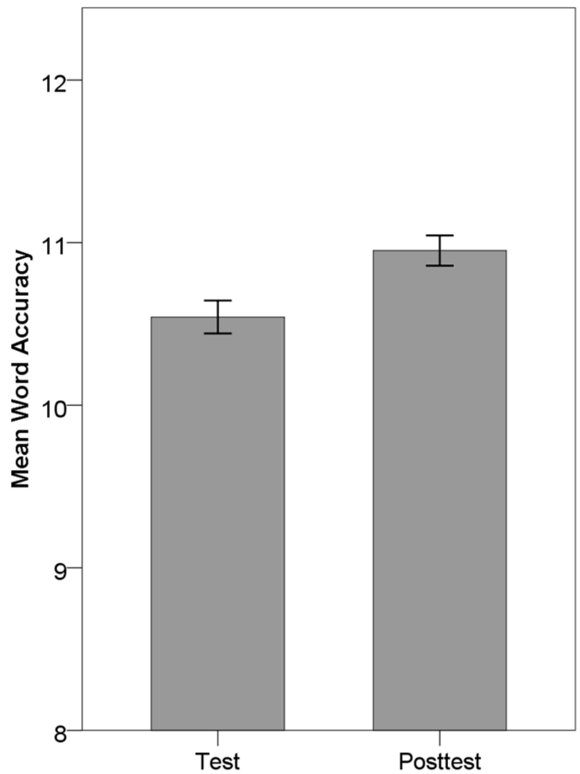

Figure 5 – Mean word accuracy (y-axis) is broken down by test condition (x-axis). Mean word accuracy is significantly higher in the posttest than the test.

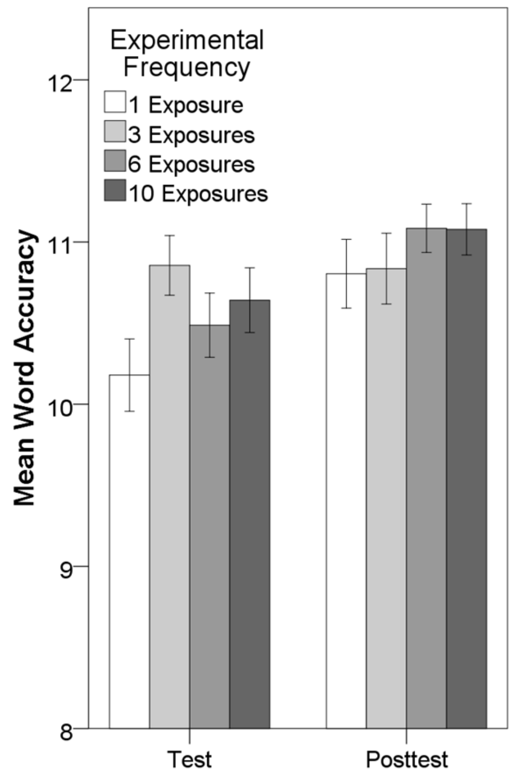

Figure 6 – Mean word accuracy (y-axis) broken down by test condition (x-axis) and experimental frequency (shading of bars). In the test condition, it appears that there exists a significant difference between the frequency level of 1 and all other levels (3,6, &10). In the posttest, it does not appear that experimental frequency levels differ significantly.

As seen in Figure 6 above, accuracy for words with the experimental frequency level of one appears to be lower than all other experimental frequency levels in the test condition. To assess if any experimental frequency level greater than one was beneficial to production accuracy in the test condition, a post hoc comparison of the experimental frequency of one versus the combination of three, six, and ten was conducted. The comparison yielded near significance, *F*(1, 32), $p = .055$, suggesting that all three frequencies greater than one conferred a production advantage.

Secondly, production accuracy scores were examined in a 2×2×3 ANOVA (Test Condition [test, posttest] × Talker Condition [single talker, multiple talkers] × Experimental Frequency [3, 6, 10]). No significant effect of talker variability or any significant interaction with talker variability was found. All *p*-values were greater than 0.10.

Semantic Learning

Results of the semantic learning analysis may be viewed above in Table II. In addition to the production accuracy data, we also examined the effects of the experimental frequency and talker variability variables on semantic learning. Eight one-sample *t*-tests—one for each of the questions in Table II—were conducted to analyze the effects of talker variability and experimental frequency on semantic learning. These *t*-tests measured whether the children preformed significantly above chance (33%) for all questions. Because the children had three options in all probes and if they were guessing at chance they would yield the accuracy of 33%, accuracy for all probes was compared to 33%. Because of the large number of tests conducted, a Šidák correction was applied to the alpha level. Null hypotheses were only rejected if $p < .0045$. Participants were significantly above chance when matching nonwords to their corresponding animals overall $t(31)=3.92$, $p = .000$; at matching the words and animals if the right answer had the experimental frequency level of ten, $t(31)=3.32$, $p = .002$; and when the target animal was produced by a single talker in the test condition, $t(31)=3.50$, $p = .001$.

Direct comparison ANOVAs were conducted across all different semantic conditions and no comparisons were found to be significant, ($p > .050$). These ANOVAs were preformed to investigate main effects of talker variability and experimental frequency on task performance. In this analysis the effects of talker variability and experimental frequency on matching accuracy were not significant.

CHAPTER IV

DISCUSSION

Production Accuracy

The results of this experiment suggest that experimental frequency and articulatory practice have beneficial roles in child speech learning. This is consistent with previous literature (Plante, Bahl, Vance, & Gerken, 2010; Goffman & Richtsmeier, 2015). The fact that the children were more accurate in the posttest condition than in the test, suggests that articulatory practice improves the children's accuracies, (Maas et al, 2008). The role of experimental frequency is less cut and dry. The children only showed benefit of more than one exposure in the test condition and not in the posttest. This suggests that multiple exposures to a novel word are only beneficial during the first few articulatory productions. This does not demonstrate a learning effect, as the benefit of multiple exposures did not yield significance in the posttest. A similar finding was reported by Richtsmeier and Goffman (under review), in which typically developing children only benefitted from a high frequency exposure to nonwords for the first one to three productions. Another interesting result of this experiment is that within the test condition, words with an experimental frequency of three yielded the highest production accuracy. Therefore it is not clear that more is necessarily better. This finding also contributes to the literature that three exposures to a form is enough to show generalization effects (Gerken & Bollt, 2008). The findings of this experiment suggest that the benefit of multiple exposures to a novel sound sequence may be achieved after only a few exposures and additional exposures may be unnecessary. This has implications for the use of some treatment techniques such as Auditory Bombardment, in which children are exposed to many productions of a word (Hodson & Paden, 1991). It may be the case that just three exposures are enough to improve a child's production accuracy, at least under conditions in which production immediately follows the perceptual exposure.

Benefits from exposure to multiple talkers were not found in this experiment. Words presented by a single talker and words presented by multiple talkers did not significantly differ in terms of production accuracy. This suggests that hearing a word from one or several talkers is equally as beneficial for production accuracy. Nevertheless, participants were numerically more accurate when working with words produced by multiple talkers, both in terms of their production accuracy and in terms of distinguishing between words presented with the same experimental frequency. Given previous research showing the benefits of talker variability (Plante et al., 2011; Richtsmeier et al., 2009; Wang et al., 1999), future research is needed to further investigate its effects. One such approach would be to investigate the timing of benefits from experimental frequency and talker variability. For example, it may be beneficial to conduct an experiment similar to the one described above that reverses the order of the blocks so that the posttest is completed before the test. This would cause the children to gain equal exposure for all nonwords early in the experiment and then gain differing exposure for the nonwords later in the experiment. The results of this hypothetical study could be compared to those of this study to investigate if exposure to multiple talkers is beneficial when presented sooner or later in a child's experience with that word.

Semantic learning

Because the children performed significantly above chance overall when matching the nonwords with the pictures of the paired animal, the above results suggest that the children were able to learn semantic information associated with the nonwords. This was found even though the researchers did not actively encourage the children to learn an association between the animals and the nonwords. The finding that children were significantly accurate when matching the words and animals if the right answer had the experimental frequency level of ten, suggests that increased frequency has a beneficial effect on semantic learning. In an interesting finding, the children were significantly accurate in matching nonwords with animals if the correct answer was presented by a single talker. This suggests that exposure to a single talker may be beneficial in semantic learning. However, talker variability was not a significant main effect in an ANOVA directly comparing accuracy for questions when the target word was presented by multiple talker animals to accuracy of questions when the target word was presented by a single talker. Further research is needed to assess the role of talker variability in semantic learning.

Clinical Significance

For both production accuracy and semantic learning, the findings of this study are consistent with previous literature that found that effects of experimental frequency may be derived in a relatively short span of time such as one study session (Plante, Bahl, Vance, & Gerken, 2010). This is clinically relevant because many speech and language therapy sessions are approximately an hour long, similar to this study's sessions with each participant. These findings suggest that a typically developing child's production accuracy has the potential to improve in one therapy session with relatively few exposures to the target words. This may hold true for some children with speech or language disorders however further research would be needed in this area.

Furthermore, this study suggests optimal numbers of exposures necessary to see improvement in accuracy for typically developing children. For production accuracy, it appears that three exposures are enough to see improvement. This finding suggests that clinical techniques such as Auditory Bombardment in the Cycles Approach (Hodson & Paden, 1990), may be optimized with a smaller number of perceptual exposures, since in the Cycles approach target words are heard more than three times. However, for semantic learning, it appears that ten exposures can cause significant increases in accuracy. It may be the case that a larger number of exposures is optimal for semantic learning.

Limitations

As noted above, this experiment investigated the speech of typically developing children. Therefore, the above conclusions may not hold true for children with speech sound disorders or language disorders. This limitation calls for future directions in which production accuracy and semantic learning in children with atypical speech or language development are studied.

Furthermore, this experiment's results may be influenced by the homogeneity of the adult speakers who lent their voices for this experiment, as well as the homogeneity of the child participants. All of the adult speakers were female, middle class, working in a University environment, and mostly all white. It may be the case that the lack of significance found in regards to talker variability may be due to the lack in variance among the stimulus talkers.

Additionally, many of the children who participated in the experiment were from middle class white families and many of the children scored well above average on the *GFTA-2.* As noted above, the children yielded an average score of 113 when a score of 100 is considered to be average. It may be the case that the above findings may have been influenced by this homogeneity in participants.

A final limitation of this study is that ten was the highest experimental frequency level studied. Especially in regard to semantic learning, it may be the case that more than ten exposures are facilitative. To better understand the number of exposures necessary to see optimal improvement in semantic learning, future directions may be to study children's accuracy in forming associations between nonwords and pictures with more levels of experimental frequency.

CHAPTER V

CONCLUSION

In regards to production accuracy, articulatory practice and frequency of exposures appear to have beneficial effects. The effect of frequency of exposures can be seen in the short span of a one-hour experiment and few exposures (three to ten) are enough to see these effects. Exposure to multiple talkers producing the target word may not be beneficial to production accuracy.

In regards to semantic learning, both experimental frequency and multiple talkers may be beneficial. These effects can also be seen only after a few exposures during a short experimental session.

The results of this experiment imply that further research should be done in regard to the effects of talker variability on production accuracy and semantic learning.

REFERENCES

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. I. (1999). Training Japanese listeners to identify English/r/and/l: Long-term retention of learning in perception and production. *Perception & psychophysics, 61*(5), 977-985.

Coady, J., Evans, J., & Kluender, K. (2010). Role of phonotactic frequency in nonword repetition by children with specific language impairments. *International Journal of Language & Communication Disorders, 45*(4), 494-509.

Gerken, L., & Bollt, A. (2008). Three exemplars allow at least some linguistic generalizations: Implications for generalization mechanisms and constraints. *Language Learning and Development*, *4*(3), 228-248.

Hoson, B. W., & Paden, E. P. (1991). *Targeting Intelligible Speech: A phonological approach to remediation* (2nd Ed.). Austin, TX: Pro-Ed.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English/r/and/l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America, 94*(3), 1242-1255.

Maas, E., Robin, D. A., Hula, S. N. A., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American journal of speech-language pathology*, *17*(3), 277-298.

Messer, S. (1967). Implicit phonology in children. *Journal of Verbal Learning and Verbal Behavior, 6*(4), 609-613.

Munson, B. (2001). Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research*, 44, 778-792.

Plante, E., Bahl, M., Vance, R., & Gerken, L. (2011). Beyond phonotactic frequency: Presentation frequency effects word producations in speech language impairment. *Journal of Communication Disorders,* 44, 91–102.

Richtsmeier, P., Gerken, L., Goffman, L., & Hogan, T. (2009). Statistical frequency in perception affects children's lexical production. *Cognition, 111*(3), 372-377.

Richtsmeier, P. T., & Goffman, L. (2015). Learning trajectories for speech motor performance in children with specific language impairment. *Journal of communication disorders*, *55*, 31-43.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *The Journal of the Acoustical Society of America, 106*(6), 3649-3658.

APPENDICES

MINIMAL PAIRS USED IN THE THIRD BLOCK ABX AUDITORY DISCRIMINATION TASK

1 dot/tot
2 buy/guy
3 cub/cup
4 hiss/hit
5 peas/peace
6 dad/bad
7 back/bag
8 leave/leaf
9 care/pair
10 lake/wake
11 fan/pan
12 pair/bear
13 sew/toe
14 walk/rock
15 tough/cuff
16 bride/bride
17 yuck/luck
18 leash/leech
19 shock/sock
20 bed/beg
21 van/fan
22 rug/rub
23 kneel/near
24 win/bin
25 red/lead
26 nine/dine
27 chip/ship
28 cuff/cup
29 zoo/sue
30 top/pop
31 bed/bet
32 gash/gas
33 gap/cap
34 dig/gig

VITA

Amanda Koehn Good

Candidate for the Degree of

Master of Science

Thesis: DISENTANGLING FREQUENCY AND TALKER VARIABILITY IN A STATISTICAL LEARNING TASK INVESTIGATING CHILD SPEECH ACQUISITION

Major Field: Communication Sciences and Disorders

Biographical:

Education:

Scheduled to complete the requirements for the Master of Science in Communication Sciences and Disorders at Oklahoma State University, Stillwater, Oklahoma in May, 2017.

Completed the requirements for the Bachelor of Arts in Linguistics at McGill University, Montreal, QC in 2014.

Experience:

Conducted a sociolinguistic survey of Native American English (Montreal, QC; 2013)

Presented preliminary findings of this thesis at the Symposium on Research in Child Language Disorders (Madison, WI; 2016)

Interned at Perkins-Tryon Elementary School conducting speech and language therapy (Perkins, OK; 2016)

Interned with Therapy Specialists conducting cognitive therapy (Stillwater, OK; 2017)

Scheduled to intern with Lions Voice Clinic conducting voice therapy (Minneapolis, MN; 2017)

Professional Memberships: ASHA, NSSHLA