

Influence of Misarticulation on Preschoolers' Word Recognition

By

Copyright 2013
Breanna I. Krueger

Submitted to the graduate degree program in the Intercampus Program of Communicative Disorders and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Arts.

Chairperson Holly L. Storkel, Ph.D.

Deborah B. Daniels, Ph.D.

Jonathan Brumberg, Ph.D.

Date Defended: 12/17/2013

The Thesis Committee for Breanna I. Krueger
certifies that this is the approved version of the following thesis:

Influence of Misarticulation on Preschoolers' Word Recognition

Chairperson Holly L. Storkel, Ph.D.

Date approved: 12/17/2013

I. Abstract

Previous research has shown that children are sensitive to speech variability in dialect and accent, and can extract information about the speaker. Misarticulated speech is a form of variability that children encounter in social situations with peers. Children are sensitive to the changes found in accented speech, but their perception of misarticulated speech has not been studied. If children do not understand misarticulated speech from their peers, they may experience a decrease in incidental word learning from peers and a reduced quality of social interactions. The purpose of the present study is to investigate if children are sensitive to misarticulations in speech, and if their ability to identify words containing misarticulated speech is affected by the speech sound substitutions being common or uncommon in children's developmental phonology. Twenty preschoolers heard minimal triplets of words that were canonical productions (e.g., leaf), productions with common substitutes (e.g. weaf), and productions with uncommon substitutes (e.g. yeaf). A forced-choice paradigm required children to click on either a real picture or a novel, anomalous picture after hearing each token. Children's mouse movements, selections and reaction times were recorded and analyzed to determine if there is a difference in response between canonical productions and those containing substitutions. Children selected more real objects pictures when they heard a canonical production than a misarticulated production. Reaction time and area under the curve were negatively impacted in substitution conditions. Among the misarticulated productions, children selected more real objects when they heard a production containing a common substitute than when they heard an uncommon substitute, but reaction time and area under the curve were not significantly different. These findings suggest that children's word recognition is facilitated by their experience with words, which supports an exemplar model

of the lexicon. Children are sensitive to substitution types that they have experience with, but this recognition comes at a cost to processing which may affect their overall understanding of rapid speech.

Acknowledgements

Table of Contents

I.	Abstract	iii
	Acknowledgements	v
	Table of Contents	vi
II.	Introduction	1
	A. Phonemic categorization of sounds.....	1
	B. Phonetic categorization of sounds.....	3
	C. Ambiguous Productions	3
	D. Misarticulation in natural speech	5
III.	Participants.....	8
	Figure 1	8
IV.	Stimuli.....	10
	A. Auditory Stimuli	10
	Table 1	11
	Table 2	12
	B. Visual Stimuli	13
	Table 3	14
	Table 4	16
V.	Procedures.....	17
	A. Preliminary testing	17
	B. Presentation of Experimental Stimuli	18
	Figure 2	19
VI.	Data Analysis	19
VII.	Results.....	20
	Figure 3	20
	Figure 4	22
	Figure 5	23
	Table 5	24
VIII.	Discussion	25
IX.	Limitations and Future Directions	29

X.	Conclusion	31
XI.	References.....	32
	Appendix A.....	34
	Appendix B.....	37

II. Introduction

Differences in speech occur as a result of speaker variability and are observed both between talkers and within talkers. There are two categories of speaker variability that affect speech perception: phonetic variability and phonemic variability. The phonetic category provides pragmatic and supralinguistic information about the speaker that may signal to the listener that the speaker is from another location (e.g. dialect). This type of variability does not change the meaning of the word or utterance; it provides additional information about the speaker, but does not affect the linguistic elements of the message. It is simply the result of individual differences among speakers. Conversely, phonemic differences do change the meaning of the word or utterance. For example, the words “bat” and “bad” are different because the final consonants are different phonemes in English. Phonemic differences are perceived as linguistic units of meaning and may affect the efficient perception of speech—especially in children (S. C. Creel, 2012; Jimenez & Creel, 2012). The message that is derived from both phonetic and phonemic differences is driven by cultural differences (phonetic) and linguistic differences (phonemic). All children are born with the capability to assign meaning to these elements, but the exact message is determined later in the child’s life through exposure to L1 and incidental learning.

A. Phonemic categorization of sounds

It is well supported that every production a speaker makes differs in its spectral characteristics, and, in order to accommodate and process this variation, humans are born with the innate ability to group sounds with proximal spectral attributes into phonemic categories that allow us to assign meaning to these sounds (Liberman, Harris, Hoffman, & Griffith, 1957). This ability is present at the earliest stages of life, and continues to be refined throughout the early

stages of development as the first language is acquired. Previous research indicates that children are sensitive to cross-linguistic phonemic boundaries in voicing features within the first year of life (Trehub, 1976; Werker, Gilbert, Humphrey, & Tees, 1981). However, this ability narrows to the categories of the child's first language throughout childhood as phonemic categories of the first language are acquired (Werker & Tees, 1983). Eimas, Siqueland, Jusczyk, and Vigorito (1971) investigated the extent of these categories further through the use of synthesized syllables with systematically controlled voice-onset-time (VOT), and found that infants in monolingual English speaking environments at 1 and 4 months of age distinguished between voiced and voiceless stop pairs at the same VOT duration where the adult phonemic boundary lies for native English speakers. Despite variability within these categories, children and adults are consistent in their perception of a boundary in VOT that cues the identification of a distinct border between categories. This research indicates that although infants are born with the ability to distinguish between many phonemic categories from birth, this ability narrows with exposure and the development of the phonological system to specialize perception in favor of the native language.

Throughout phonological development, children are exposed to varied productions of sounds in their first language that further solidifies their phonemic categories and their rapid identification of words. In a study measuring adults' and infants' ability to distinguish between syllables containing native language contrasts, it was found that both adults and infants are highly accurate in identifying phonemic contrasts; however, only the infants in this study demonstrated the ability to distinguish between phonemically contrasting syllables that were found in languages other than English (Trehub, 1976). This suggests that adults perception is driven by the amount of exposure they have had to the sound system of their first language,

whereas infants, on the other hand, only rely on their innate faculties at a very early age because they do not have the experience to help shape their perceptual categories.

B. Phonetic categorization of sounds

Although some speech sound variability affects the meaning of a word, there is another type of variability that does not: phonetic variability. This can be explained using the categorical perception example discussed in the previous section. There are two distinct categories that can be identified by VOT. Therefore, any variability *within* those categories would be perceived as phonetic variants of the same phoneme. Another example of phonetic variability is aspiration in English. Voiceless stop consonants (e.g. /p, t, k/) are aspirated in word initial position, but when these are preceded by a /s/, the stops become unaspirated. This change is not phonemic in English, but allophonic, and therefore, these consonants are not identified as being voiced (e.g. /b, d, g/).

C. Ambiguous Productions

Children are able to perceive and process variability in speech as effectively as adults, and may share some of the same biases regarding speech that differs from the child's own. Creel (2012) studied the effect of systematic shifts in vowels and consonants that may provide insights into how children process speech variability. An eye-tracking and mutual exclusivity paradigm was used. Mutual exclusivity is a strategy children use to learn new words. For example, if a child is presented with a novel word and sees a known object and an unknown object, the child will respond that the novel word is the name of the unknown object because the child assumes that there cannot be two names for a single object (Merriman, Bowman, & MacWhinney, 1989). This principle is utilized by Creel (2012) to investigate if children perceive shifted productions of words as novel words (indicating that the shifted production is interpreted as phonemic) or as

phonetic variations of a word the child already knows. Creel (2012) conducted three experiments to investigate the flexibility of children's perception of shifted productions of words to determine how children interpret these shifts. The first experiment investigated children's response to shifted vowels across phonemic boundaries and found that children highly preferred the real object regardless of the shift, indicating that the shifted production was interpreted as a phonetic variant of the target production. However, their eye tracking data suggested a longer processing time and more uncertainty with shifted vowels, suggesting that there was a processing cost in accommodating the shifted production. The second experiment employed the real/novel object selection using the same systematic change in vowels and found that the real object, again, was preferred, replicating the findings from the first experiment. The results of these first two experiments demonstrate that children are broadly accepting of vowel changes as a phonetic variant of the canonical word, but this acceptance requires a longer processing time before making a selection indicating a processing struggle between phonemic and phonetic categorization of that variability.

The goal of Experiment 3 was to investigate the effect of changing vowel location or consonant voicing on real object selection and eye gaze as compared with canonical productions to determine the degree of sensitivity to change. To test perceptual sensitivity, the stimuli were divided between "close" and "far" feature distance from the target. Each production was different in either one feature (e.g. "close" voiced→voiceless /b/→/p/) or in two (e.g. "far" voiced bilabial→voiceless velar /b/→/k/). Children's eye movements were tracked when presented with four pictures—three real objects, one novel object. Each participant heard the following: six canonical productions, six vowel-shifted words, six words containing onset consonants where the initial consonant voicing was changed, six words where the initial

consonant voicing and place was changed, six words containing coda consonants with changed voicing, and six words containing coda consonants where the voicing and place were changed (e.g. /greip/→/kreip/, /greip/→/treip/ and /klɔk/→/klɔg/, /klɔk/→/klɔd/). The results of this experiment revealed that children's selection of novel objects increased with feature distance from the canonical production, while eye movements slowed (Creel, 2012). This finding suggests that children are less likely to interpret a shifted production as a phonetic variant of the target production as the difference between the phonetic variant and target increases. That is, when the shifted production is largely different from the target production, the shifted production is interpreted as a new word, entailing a phonemic contrast between the shifted and target production.

D. Misarticulation in natural speech

Although Creel (2012) examined the perception of misarticulated words, these misarticulations were not aligned with children's experience with misarticulation. That is, the stimuli were systematically created by changing specific features, but there was no consideration for whether the changed words resembled real variability that occurs in natural speech that children would have heard. Children hear misarticulated speech in their interactions with peers or siblings, however, the results of Creel's (2012) study do not account for how the child's exposure to speech variability affects their response. Perhaps one of the most common types of misarticulated speech that children hear as a result of typical phonological development. The literature surrounding children's perception of phonological differences fails to account for this speech that all children produce themselves and hear from their siblings and peers. It has been long established that biological predisposition and exposure are both key factors in how we

perceive speech, yet the area of children's perception of their peers' phonological patterns is comparatively understudied.

As children acquire language, they produce universally similar articulation error-patterns in their phonology according to the phonological system of their first language. When the child begins to acquire her first word, she may not have the phonological repertoire to make all of the phonemes in the words she is producing. Therefore, certain substitutions are made for sounds. According to normative data, these substitutions are systematically the same for most English-speaking children at various stages of development (Smit, 1993a; Smit, 1993b). For example, if children produce a substitute for /r/, that substitute is most frequently a /w/ and very rarely a /j/. In addition to providing information about the types of substitutions made, these norms also tell us at what age these substitutions occur in children. From that data, we can determine which sounds are "late-acquired" and which sounds are "early-acquired." Those that are late-acquired are likely to be produced with substitutions much longer than those that are early-acquired, and so children hear them from their peers well into the preschool years. The perception of this variability has not been examined in children. Although it is apparent that they do accommodate speech variability to a certain degree (Creel, 2012), we do not have evidence of the impact of experience with misarticulated speech on children's speech perception. It is possible that since children commonly hear a /w/ as a substitute for /r/ (e.g. "wabbit" instead of "rabbit"), they may be more likely to accept this as a phonetic variation rather than a phonemic difference. In complement, they may be more likely to interpret an unattested substitution (e.g., "yabbit" instead of "rabbit") as a novel word.

The present study will investigate preschooler's processing of misarticulated speech to answer the following questions:

- 1) Do children identify words containing speech sound substitutions as real words or as novel words?
- 2) Do children process words containing misarticulated speech in the same way as canonical productions of the same word?
- 3) Do children respond differently to speech sound substitutions that are typically produced by children versus those substitutions that are rarely produced during typical development?

These questions will be investigated through measuring children's responses to words that contain speech sound substitutions and canonical productions of the same words. The first question will be addressed by measuring how many real-object responses are selected by the children when presented with these varied productions of the canonical words. It is predicted that there will be more real-object responses for canonical productions of words than in the misarticulated conditions. The second question will be addressed by evaluating the response time and computer mouse movements of the children in responding to stimuli. In concordance with the findings of Creel (2012), a processing cost is expected for misarticulated productions when compared to canonical productions. The third question will be addressed by looking at the difference in responses when the words containing speech sound substitutions have common substitutions as compared to those with uncommon speech sound substitutions, using the same dependent variables as the first two questions. Common substitutions are those that children have more experience with, so it is predicted that (1) children will select real objects more often for common substitutes than uncommon substitutes, indicating that common substitutions are more likely to be interpreted as phonetic variants than uncommon substitutes; (2) children's processing time will be less for common substitutes than for uncommon substitutes, indicating reduced

processing costs for common over uncommon misarticulations due to past experience interpreting these misarticulations.

III. Participants

Participants were 20 monolingual English-speaking preschoolers aged 4 years to 5 years, 11 months (\bar{X} =5.04 years; SD =0.44, range 4.0--5.92) with 9 females and 11 males. Participants were recruited through fliers distributed to three preschools in Lawrence, Kansas and via a research subject database. Parents of the children were asked to complete a history questionnaire (see Appendix A) describing aspects of the child's development. One field on this questionnaire asked parents to rate the frequency of exposure to children who misarticulate by making a mark on a visual analog scale by the parents on a 10 centimeter line with "never" on the left edge and "frequently" on the right edge of the line (see Appendix A). This provides a way of measuring that is continuous, but objective since just writing down a percentage or number may mean different things for different parents.

Figure 1

How often is your child around children who pronounce words incorrectly, such as "weaf" instead of "leaf?" (Make a mark on the line below)

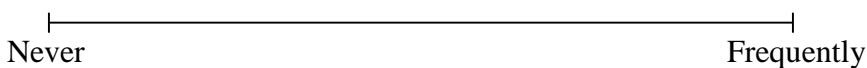


Figure 1: Visual analog scale used to measure the frequency of participants' exposure to children who produce speech sound substitutes.

Parents were asked to draw a dash along the line to indicate their child's experience with misarticulated speech. These responses were then measured to obtain a value from 0-100 mm. The criterion is more concrete for parents with a visual analog scale, and this type of scale has

been shown to be more reliable for the collection of subjective information (McCormack, Horne, & Sheather, 1988). Participants were rated as being around misarticulated speech from other children on average 58.45 (SD=24.85, range 11--100), which approximately corresponds to the midpoint of the scale (i.e., 50 mm), albeit slightly shifted towards the frequent exposure end of the scale. Thus, as expected, all participating children had some current exposure to misarticulated speech, although there were individual differences in the degree of exposure.

Each child was administered the Goldman-Fristoe Test of Articulation, 2nd edition (GFTA-2). The GFTA-2 is a standardized assessment that compares children's production of phonemes in each position (initial, medial and final) at the word level. The number of errors is compared with a normative sample for the child's age and gender. Standard scores for all participating children were within normal limits (\bar{X} =111.55; SD=4.21, range 99--118).

In addition to the GFTA-2, the researcher selected words from the Phonological Knowledge Protocol (PKP) that correspond to the onset phonemes that were used in the study (Dinnsen & Gierut, 2008). In the GFTA-2, only a single vowel context is tested for each phoneme position; therefore, since this study is concerned with the effect of misarticulation on word recognition, a complete picture of the child's phonological production is necessary. The PKP is a picture naming task that includes a total of 198 pictures to elicit phonemes in 5 different vowel contexts for each word position. For the present study, 60 selected items from the PKP were used to explore each child's production of the 12 target phonemes used in the stimuli in the word-initial position. The results indicated that children produced the target phonemes with the following accuracy (listed in order from most to least accurate): /f/ -- 100% accuracy, /dʒ/ -- 99% accuracy, /tʃ/ -- 97% accuracy, /k/ -- 94% accuracy, /g/ -- 94% accuracy, /l/ -- 89% accuracy, /ʃ/ -- 87%

accuracy, /k/ with 84.5% accuracy, /v/ with 80% accuracy, /r/ with 79% accuracy, and /θ/ with 60% accuracy.

Children were also administered the Peabody Picture Vocabulary Test, 4th edition (PPVT-4), to ensure that children's word recognition performance was not influenced by having a low receptive vocabulary. The PPVT-4 is a receptive vocabulary test that requires children to point to pictures spoken by the examiner and choose the correct answer from a field of four. Scores are compared with a normative sample of children the same age. All children earned standard scores that were within normal limits (\bar{X} =114.25; SD=9.55, range 96--129). All participants passed a hearing screening at 20dB HL at 1000, 2000 and 4000 Hz and had normal or corrected-to-normal vision.

IV. Stimuli

A. Auditory Stimuli

The stimuli consisted of twelve monosyllabic real words, the same twelve real words with common speech sound substitutes, and with uncommon speech sound substitutes, so that each canonical word was a minimal triplet with two misarticulated productions (see Table 1). The twelve canonical words were selected based on log frequency in the Child Corpus Calculator (Storkel & Hoover, 2010), to ensure that words would be familiar to the participants, and picturability, as required by the experimental paradigm. The common and uncommon variations of the canonically produced words were also checked in the Child Corpus Calculator (Storkel & Hoover, 2010) to ensure that neither substitute-containing word was a real word that would be in the child's lexicon. The onset phoneme of each word was a late-acquired sound as defined by normative data (Smit, Hand, Freilinger, Bernthal, & Bird, 1990), to make it more likely that

children would have heard misarticulated productions of the word. Each onset-phoneme was then changed to its common substitute by selecting one of the most frequently occurring speech sound substitute based on normative data (Smit et al., 1990). The phoneme for the uncommon substitute was selected to be rarely occurring based on normative data (Smit et al., 1990). Common and uncommon substitutes were selected in tandem so that the feature distance between each substitute and the canonical production could be matched (*see* Table 1).

Table 1

Canonical Word	Log Frequency	Common Substitute	Feature Distance from Canonical	Uncommon Substitute	Feature Distance from Canonical	Feature Difference of Common vs. Uncommon
Chick	1.48	Shick	1	fick	4	-3
leaf	2.04	Weaf	6	yeaf	5	1
thumb	2.11	Fumb	2	shumb	3	-1
comb	2.36	Tomb	4	pomb	3	1
Jar	2.49	Dar	4	gar	4	0
safe	2.57	Tafe	3	pafe	4	-1
van	2.58	Ban	3	dan	4	-1
shirt	2.81	Sirt	2	firt	3	-1
clock	2.91	Kwock	6	kjock	5	1
rope	3.08	Wope	3	yope	2	1
fish	3.34	Pish	3	tish	4	-1
girl	3.95	Dirl	4	birl	3	1
Mean	2.643		3.417		3.667	-0.363
SD	0.619		1.441		0.850	1.226

Table 1: Canonical stimuli were chosen based on frequency, picturability and age of acquisition of onset phoneme. Common and uncommon substitutes were chosen based on normative data and feature distance from the canonical production.

These thirty-six words were recorded by the researcher, a female native English speaker with a Midwestern dialect of American English, in an anechoic chamber with a Marantz PMD671 solid-state digital recorder. Words were recorded three times each in the carrier phrase “Look at the _____” to ensure similar intonation and to control for listing effects. The words were then extracted from the carrier phrase using Praat sound editing software, with 250ms of silence

embedded at the onset and offset of each word (Boersma & Weenink, 2013). The mean intensity of all items together was measured to be 65dB, then each word was scaled to 65 dB to match. The duration of the words were then measured and compared to find tokens that were similar in duration across the three stimuli conditions (canonical, common substitute, uncommon substitute). These durations can be seen in Table 2. The difference in duration between the canonical word and the common substitute and the uncommon substitute were compared using a three-way ANOVA. The analysis of the canonical words and the common substitute words was not significant [$F(1,2)=2.100, p=0.15, \eta_p^2=0.16$]. A naïve judge who was unfamiliar with the stimuli and the experimental conditions transcribed the stimuli to verify that they were heard as intended.

Table 2

Canonical Word	Duration (ms)	Common Substitute	Duration (ms)	Uncommon Substitute	Duration (ms)
chick	0.560	shick	0.596	fick	0.587
leaf	0.653	weaf (6)	0.652	yeaf (5)	0.610
thumb	0.575	fumb	0.577	shumb	0.569
comb	0.501	tomb	0.507	pomb	0.495
Jar	0.520	dar	0.503	gar	0.451
safe	0.687	tafe	0.633	pafe	0.658
van	0.572	ban	0.481	dan	0.515
shirt	0.620	sirt	0.561	firt	0.562
clock	0.555	kwock	0.541	kjock	0.548
rope	0.471	wope	0.519	yope	0.508
fish	0.623	pish	0.626	tish	0.631
girl	0.411	dirl	0.395	birl	0.387
Mean	0.562		0.549		0.543
SD	0.079		0.074		0.077

Table 2: The duration of each token in milliseconds (ms) and summary statistics.

B. Visual Stimuli

Thirty-six black and white line drawings were selected to match each production of the canonical words. Twelve of the pictures depicted the canonical words and were selected from Microsoft Clipart and the Snodgrass and Vanderwart (1980) standardized set of pictures. Twenty-four of the pictures were anomalous objects from the Kroll and Potter (1984). Characteristics of the nonobjects are available from Kroll and Potter (1984) and Storkel and Adlof (2009). Storkel and Adlof (2009) collected data from adults for all of the nonobjects and from children for only a subset of the nonobjects. Since the current study is with children, only the subset of nonobjects with child data were used. From this subset, nonobjects were eliminated based on adult object-likeness ratings from Kroll and Potter (1984). This value was obtained by asking adults and children to rate on a 7-point scale the degree to which the non-object resembled a real object (where 1 indicated “nothing like a real object” and 7 indicated “looks like a real object”). Nonobjects with ratings of 6 and above were removed so that the remaining nonobjects would be less likely to be confused with real objects. Child semantic set size and child strength of the first neighbor from Storkel and Adlof (2009) were then used to further reduce the pool and create two sets of matched nonobjects. Child semantic set size is the number of different neighbors reported by at least two child participants (Storkel & Adlof, 2009). Child strength of the first neighbor was calculated by dividing the number of children who responded with this particular neighbor by the total number of child participants. Nonobjects with many semantic neighbors (i.e., 9 or more) and relatively weak first neighbors (i.e., strength of 0.20 or less) were selected so that children would not have a strong real object interpretation of the nonobjects. The 24 selected nonobjects were then divided into two sets of 12 that were matched on all three nonobject characteristics as shown in Table 3 below. Each of these pictures were

cropped and resized to 144 x 144 pixels to control for the size of the selection area in the mousetracking software.

Table 3

Picture Set 1	Object-Likeness Rating	Semantic Set Size	Frequency of Strongest Neighbor	Picture Set 2	Object-Likeness Rating	Semantic Set Size	Frequency of Strongest Neighbor
nobj 75	3.3	9	0.09	nobj79	3.6	9	0.04
nobj31	5.2	9	0.11	nobj81	5.1	10	0.05
nobj39	3.1	10	0.13	nobj58	3.1	10	0.16
nobj22	3.8	10	0.05	nobj26	3.6	10	0.20
nobj29	3.3	11	0.20	nobj46	3.2	11	0.09
nobj17	5.7	11	0.03	nobj61	5.7	11	0.08
nobj02	5.9	11	0.07	nobj11	5.2	11	0.14
nobj38	3.9	12	0.05	nobj80	4.8	11	0.05
nobj52	3.4	12	0.11	nobj47	4.3	13	0.11
nobj27	5.2	13	0.16	nobj53	3.4	14	0.08
nobj88	3.2	15	0.07	nobj01	3.6	15	0.05
nobj59	4.3	17	0.11	nobj13	3.6	16	0.05
Mean	4.2	11.7	0.10		4.1	11.8	0.09
SD	1.0	2.4	0.05		0.9	2.2	0.05

Table 3: Set 1 and Set 2 of nonobject pictures with their object-likeness rating, semantic set size and frequency of strongest neighbor with summary statistics.

Eight total versions of the study were created to control for potential effects of nonobject picture bias and to ensure that each token was played once per block in the first half and second half of the experiment. In Block 1, each canonical word was presented 1 time. Also in Block 1, either the common or uncommon substitute word was presented one time. For example, as shown in Table 4, “chick” occurs in first block and its uncommon substitute was “fick.” In contrast, the canonical word “leaf” occurs in Block 1 with the common substitute “weaf” also in Block 1. Then in the second block, each real word is presented a second time, but now with a different visual display. For example, in Block 1, “chick” occurred with the real picture of a “chick” and “nonobject 75.” In Block 2, “chick” occurs again but now with the “chick” picture and “nonobject 79.” This ensured there was no bias toward choosing one nonobject over another

in canonical production conditions. Also in block 2, whichever misarticulation was not presented in Block 1 is presented in Block 2 to balance misarticulation types across the totality of the experiment. So for example, the uncommon substitute for “chick,” “fick,” occurred in Block 1 but now the common substitute “shick” occurs in Block 2. So, in each block, the proportion of canonical productions and misarticulated productions is equivalent (i.e. 50%). Within the misarticulated productions, 50% are common, and 50% are uncommon in a given block. The location of the real object and the nonobject picture was randomized across trials so that on 50% of the trials, the real object was on the left, and the nonobject was on the right, and on the other 50% of the trials, the real object was on the right, and the nonobject was on the left. Furthermore, the location of the pictures in Block 2 was reversed from that in Block 1. For example, “chick” in Block 2 the nonobject is on the left, real object is on the right, whereas it was the opposite array for Block 1 (*see* Table 4).

Table 4

Block 1				Block 2			
Stimulus	Condition	Response Left	Response Right	Stimulus	Condition	Response Left	Response Right
chick	CanonU	chick	nobj75	chick	CanonC	nobj79	chick
fick	Uncommon	nobj75	chick	shick	Common	chick	nobj79
thumb	CanonU	thumb	nobj39	thumb	CanonC	nobj58	thumb
shumb	Uncommon	thumb	nobj39	fumb	Common	nobj58	thumb
jar	CanonU	nobj29	jar	jar	CanonC	jar	nobj46
gar	Uncommon	jar	nobj29	dar	Common	nobj46	jar
van	CanonU	nobj02	van	van	CanonC	van	nobj11
dan	Uncommon	nobj02	van	ban	Common	van	nobj11
clock	CanonU	clock	nobj52	clock	CanonC	nobj47	clock
cjock	Uncommon	nobj52	clock	cwock	Common	clock	nobj47
fish	CanonU	fish	nobj88	fish	CanonC	nobj01	fish
tish	Uncommon	fish	nobj88	pish	Common	nobj01	fish
leaf	CanonC	nobj81	leaf	leaf	CanonU	leaf	nobj31
weaf	Common	leaf	nobj81	yeaf	Uncommon	nobj31	leaf
comb	CanonC	nobj26	comb	comb	CanonU	comb	nobj22
tomb	Common	nobj26	comb	pomb	Uncommon	comb	nobj22
safe	CanonC	safe	nobj61	safe	CanonU	nobj17	safe
tafe	Common	nobj61	safe	pafe	Uncommon	safe	nobj17
shirt	CanonC	shirt	nobj80	shirt	CanonU	nobj38	shirt
sirt	Common	shirt	nobj80	firt	Uncommon	nobj38	shirt
rope	CanonC	nobj53	rope	rope	CanonU	rope	nobj27
wope	Common	rope	nobj53	yope	Uncommon	nobj27	rope
girl	CanonC	nobj13	girl	girl	CanonU	girl	nobj59
dirl	Common	nobj13	girl	birl	Uncommon	girl	nobj59

Table 4: Block 1 (left) and Block 2(right) of Version 1. From left to right, audio stimulus, type of stimulus, and the array of the left and right pictures for each visual display.

This first version of the experiment was systematically manipulated to create the alternate 7 versions as follows:

- 1) Block 1 exchanged with Block 2
- 2) Pictures moved to opposite side of the display (pictures on the left are now on the right)
- 3) Pictures moved to opposite side of the display *and* Block 1 exchanged with Block 2
- 4) Pictures from Set 1 assigned to common productions; pictures from Set 2 assigned to uncommon productions

- 5) Pictures from Set 1 assigned to common productions; pictures from Set 2 assigned to uncommon productions, *and* Block 1 exchanged with Block 2
- 6) Pictures from Set 1 assigned to common productions; pictures from Set 2 assigned to uncommon productions, *and* pictures moved to opposite side of the display (pictures on the left are now on the right)
- 7) Pictures from Set 1 assigned to common productions; pictures from Set 2 assigned to uncommon productions, *and* pictures moved to opposite side of the display *and* Block 1 exchanged with Block 2.

V. Procedures

Written consent was obtained from the parent (*see* Appendix B), and then each child participated in 1 session of testing and computer practice lasting 30-45 minutes, and 1 session of experimental testing. Before each session, verbal assent was obtained from each child (*see* Appendix B), and the child was informed that he or she could stop the experiment at any time without consequence. Each session took place in the preschool, in a quiet room apart from the classroom. In the first session, children were administered preliminary testing in articulation, receptive vocabulary, hearing and competence in computer use. In the second session, children completed any remaining preliminary testing and were administered the experimental protocol.

A. Preliminary testing

Each subject participated in 1 session of preliminary testing that lasted 30-45 minutes. A hearing screening was conducted as well at 20dB HL at the frequencies of 1000, 2000 and 4000 Hz. In order to assess articulation and receptive vocabulary, the researchers employed two standardized norm-referenced assessments that are commonly used in the field of speech-language pathology—the Goldman-Fristoe Test of Articulation, Version 2 and the Peabody

Picture Vocabulary Tests, Version 4, respectively. In addition, a supplemental articulation probe (i.e., Phonological Knowledge Protocol, (Dinnsen & Gierut, 2008)) was administered to better understand the context in which the child's speech sounds occur.

All children also participated in a computer training task. For this task, participants were seated at a Dell Latitude D610 PC Laptop with external speakers, a 17.25" x 13.25" mousepad, and a single-button optical mouse. The picture and audio stimuli were presented using MouseTracker software (Freeman & Ambady, 2010). Each child heard ten trials of real words and saw real object pictures with each word repeated once. Children were asked to click on the picture that they heard. This practice was visually identical to the experimental protocol but allowed the researcher to provide feedback. . This feedback included: "Be sure to move straight and fast to the picture!" and "Click to hear the next word." These stimuli were recorded and prepared in the same way as described above (*See Stimuli*).

In the second session, children participated in 5 practice trials and 48 experimental trials. The 5 practice trials contained the same real words as the training task, and feedback was provided using the same method as in the first training task. For experimental trials, items were presented randomly and feedback was not provided. Each session lasted approximately 45 minutes.

B. Presentation of Experimental Stimuli

All 48 experimental trials were presented in random order using the Mouse Tracker software program (Freeman & Ambady, 2010). Each trial consisted of two screens: a "start" screen and the "stimuli screen" (*see* Figure 2). The start screen had a "start" button in the bottom center of the screen, then after clicking "start" the following screen, the stimuli screen, appeared with visual stimuli located in the top right and left-hand corners of the screen (*see* Figure 2). When the start button was clicked, the pictures were presented simultaneously with the onset of

the auditory stimuli, and timing began. Once the picture selection was made, the screen would return to the “start” screen to await the next trial and the picture choice, reaction time, and mouse trajectory were recorded by the software. Participants were permitted to respond at their own pace, but were instructed to move the mouse as quickly and as directly as possible to their selection. No feedback was provided.

Figure 2

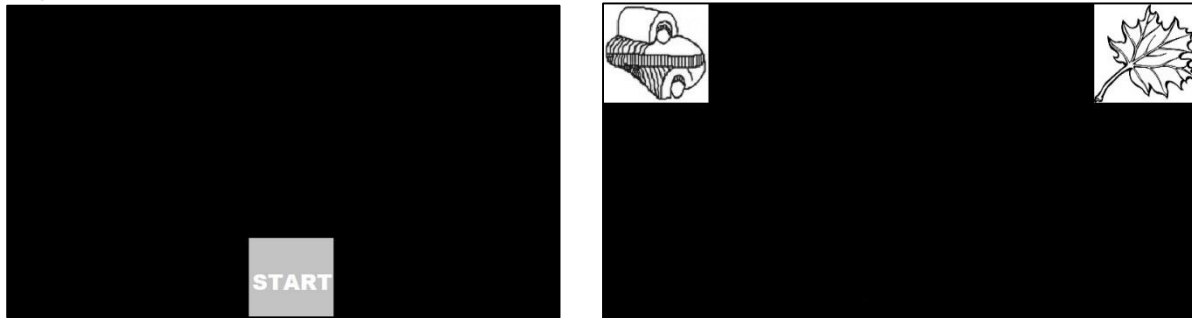


Figure 2: Participants first see the "Start" screen (left). After clicking the "Start" button, the screen changed to present the object choices (right).

VI. Data Analysis

Each output file was processed using the Mousetracker Analyzer software (Freeman & Ambady, 2010). The researcher extracted the proportion of real object choices, the mean reaction time to the real object, and the mean area under the curve when selecting the real object into a spreadsheet for statistical analysis. These measurements were separated by independent variable types: real word (paired with common speech sound substitute nonobject), real word (paired with uncommon speech sound substitute nonobject), common speech sound substitute, and uncommon speech sound substitute. The percentage of real object selections was calculated for each condition. Reaction time was measured by the software as the time between the “start” click and the response click of the mouse. The mean area under the curve is the mean of the area beneath the actual response trajectory (on the monitor screen), and the most direct path to that

response choice. Reaction time and area under the curve were only analyzed for trials where the child selected the real object. As shown in the results, the most prevalent response choice in all conditions was selection of the real object. Thus, restricting analysis to real object choices allows comparison of speed and mouse trajectory to the same response outcome.

VII. Results

Figure 3

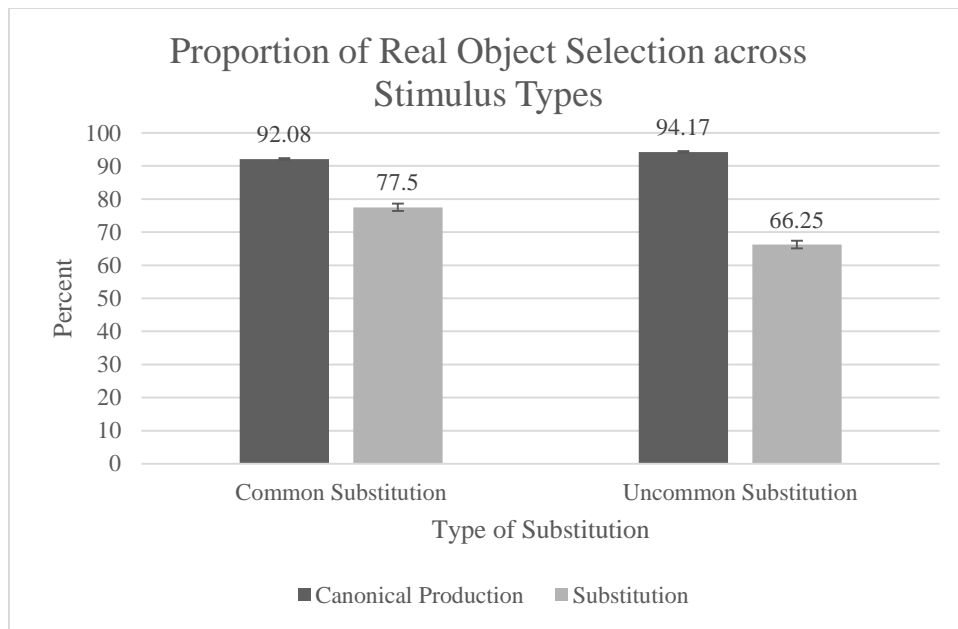


Figure 3: Percent of real object selections by canonical production and substitution type, with bars showing the standard error of the mean.

The percentage of real object selections were compared using a 2 Word Type (Canonical vs. Substitute) x 2 Substitute Type (Common vs. Uncommon) repeated measures ANOVA. A significant main effect of Word Type was found [$F(1,19)=20.43, p<.001, \eta_p^2=0.52$] (see Figure 3) with children choosing real objects significantly more for canonical productions ($\bar{X}=93.1\%$, $SD=1.3\%$, range 66.66%-100%) than for substitutions ($\bar{X}=71.9\%$, $SD=5.1\%$, range 16.67%--100%). Furthermore, a significant main effect for Substitute Type was found [$F(1,19)=4.430, p=0.049, \eta_p^2=0.189$]. However, this main effect is qualified by a significant interaction between

Word Type and Substitution Type [$F(1,19)=6.204$ $p=0.022$ $\eta_p^2=0.246$]. Four paired samples t-tests were conducted to investigate the differences between canonical productions and misarticulations. A Bonferroni correction for multiple comparisons was applied yielding a critical p-value of 0.0125 for significance for each comparison. Recall that each real word was paired with two different displays; (1) the real object (e.g. rope) and the nonobject from the common substitute condition (e.g. nobj75); (2) the real object (e.g., leaf) and the nonobject from the uncommon substitute condition (e.g., nobj31). As was expected, canonical productions that were matched with common misarticulation visual displays ($\bar{X}=92\%$, $SD=8\%$, range 75%-100%) were not significantly different than canonical productions that were matched with uncommon misarticulation visual displays ($\bar{X}=94\%$, $SD=9\%$, range 66.67%-100%) [$t(19)=-0.77$, $p=0.45$]. In contrast, there was a significant difference between canonical productions ($\bar{X}=92\%$, $SD=8\%$, range 75%-100%) and common substitutes ($\bar{X}=78\%$, $SD=22\%$, range 16.67%--100%) [$t(19)= 3.68$, $p=0.002$]. Likewise, a significant difference was observed between canonical productions ($\bar{X}=94\%$, $SD=9\%$, range 83.3%-100%) and uncommon substitutes ($\bar{X}=66\%$, $SD=28\%$, range 50%-100%) [$t(19)=4.26$, $p<.001$]. Finally, children chose real objects significantly more with common substitutes ($\bar{X}=78\%$, $SD=22\%$, range 16.67%--100%) than in uncommon substitutes ($\bar{X}=66\%$, $SD=28\%$, range 50%-100%) [$t(19)=2.77$, $p=0.012$].

Figure 4

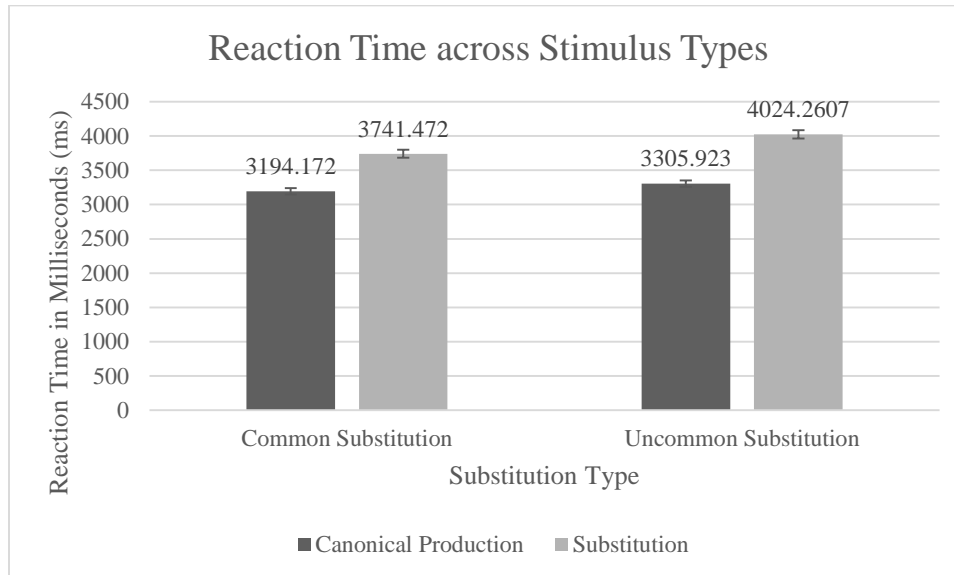


Figure 4: Reaction time in milliseconds (ms) by substitution type, including bars showing standard error of the mean.

Mean reaction time was compared using a 2 Word Type (Canonical vs. Substitute) x 2 Substitute Type (Common vs. Uncommon) repeated measures ANOVA. A significant main effect of Word Type (canonical vs. substitution) was found [$F(1,19)=30.61, p<.001, \eta_p^2=0.62$]. As shown in Figure 4, children responded more quickly to canonical productions ($\bar{X}=3250.05, SD=202.50, \text{range } 1814.09\text{--}6195.63$) than to misarticulated productions ($\bar{X}=3882.87, SD=265.38, \text{range } 1866.36\text{--}7544.5$). No significant interaction between Word Type and Substitution Type was observed [$F(1,19)=0.442, p=0.514, \eta_p^2=0.023$]. The paired samples t-test revealed that children did not chose real objects any faster in canonical productions that were visually matched with common misarticulations ($\bar{X}=3194.17, SD=867.54, \text{range } 1814.09\text{--}5112$) than canonical productions that were visually matched with uncommon misarticulations ($\bar{X}=3305.92, SD=1017.79, \text{range } 1951.6\text{--}6195.63$). Children chose real objects faster in canonical words ($\bar{X}=3194.17, SD=967.54, \text{range } 1814.09\text{--}5112$) than in common substitutions (\bar{X}

=3741.47, SD=1083.46, range 1866.36–6251.36) [$t(19)=-4.819$ $p<.001$] and faster in canonical words ($\bar{X}=3305.92$, SD=1017.79, range 1951.6–6195.63) than in uncommon substitutions ($\bar{X}=4024.26$, SD=1409.48, range 2296.82--7544.5) [$t(19)=-3.34$ $p=.003$]. In contrast, the difference between substitute types was not significant [$t(19)=-1.53$ $p=0.14$]. Children did not respond more quickly to common misarticulations than to uncommon misarticulations.

Figure 5

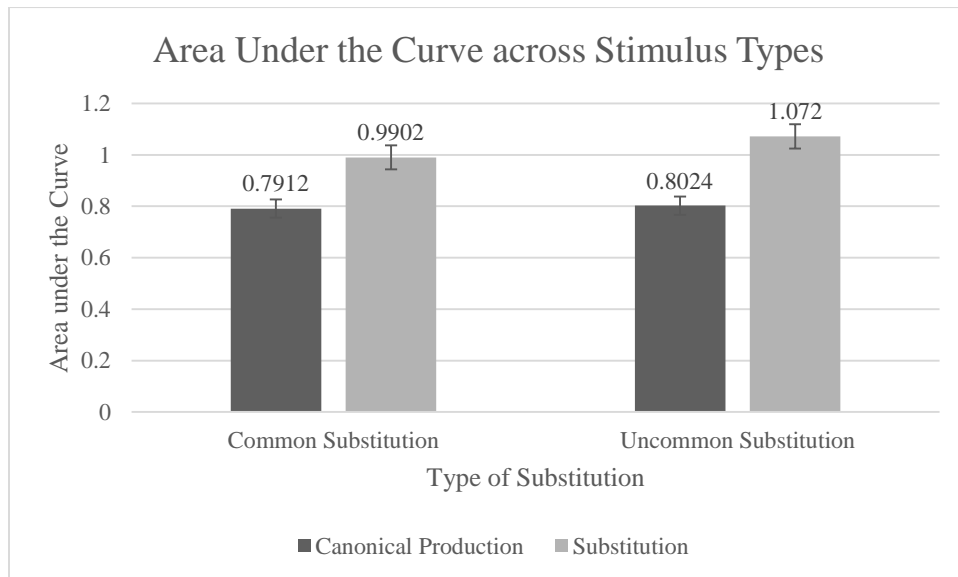


Figure 5: Area under the curve (AUC), with bars showing the standard error of the mean, demonstrates the amount of deviance from the most direct mouse trajectory, where lower values suggest less deviation.

Area under the curve (AUC) was also compared using a 2 Word Type (Canonical vs. Substitute) x 2 Substitute Type (Common vs. Uncommon) repeated measures ANOVA. A significant main effect for Word Type was found [$F(1,19)=5.11$, $p=0.04$, $\eta_p^2=0.21$] (see Figure 5). These results indicated that the AUC was significantly lower on canonical word trials ($\bar{X}=0.8$, SD=0.16, range 0.003--2.65) than on trials containing substitutions ($\bar{X}=1.03$, SD=0.21, range -0.40--4.66). This value indicates that the mouse trajectory for canonically produced words was more direct than for misarticulated words. Moreover, no significant interaction between Word Type and Substitution Type was observed [$F(1,19)=0.12$, $p=0.73$, $\eta_p^2=0.006$].

The analysis of paired samples t-tests revealed that there was no difference in mouse trajectory between canonical productions that were visually matched with common productions ($\bar{X}=0.79$, $SD=0.84$, range 0.003--2.73) than canonical productions that were visually matched with uncommon productions ($\bar{X}=0.08$, $SD=0.68$, range 0.037--2.65) [$t(19)=-0.081$, $p=0.94$]. Children were not more direct in their mouse path in canonical productions ($\bar{X}=0.79$, $SD=0.84$, range 0.003--2.73) than in common misarticulations ($\bar{X}=0.99$, $SD=0.872$, range -0.07--3.13) [$t(19)=-2.00$, $p=0.060$]. Additionally, children were not more direct in their mouse path in canonical productions ($\bar{X}=0.80$, $SD=0.68$, range 0.037--2.65) than in uncommon misarticulations ($\bar{X}=1.07$, $SD=1.17$, range -0.40--4.66) [$t(19)=-1.52$, $p=0.15$]. The results from these last two pairwise comparisons is counter to the significant main effect of Word Type and is likely due to lack of power for the two pairwise comparisons and the large standard deviations for this dependent variable. Finally, children were not more direct in their mouse path in common misarticulations than in uncommon misarticulations [$t(19)=-0.436$, $p=0.668$].

Table 5

Word Type x Substitution Condition	Age (in months)	Frequency around other children who misarticulate	GFTA-2 Standard Score	PKP Proportion correct
Proportion of Real Object Selection in Canonical productions with common misarticulations	$r=0.138$ $p=0.561$	$r=-0.298$ $p=0.202$	$r=0.019$ $p=0.938$	$r=-0.338$ $p=0.145$
Proportion of Real Object Selection in Canonical productions with uncommon misarticulations	$r=0.197$ $p=0.405$	$r=0.099$ $p=0.679$	$r=-0.110$ $p=0.645$	$r=-0.239$ $p=0.311$
Proportion of Real Object Selection in Common misarticulations	$r=0.049$ $p=0.838$	$r=-0.237$ $p=0.315$	$r=0.003$ $p=0.989$	$r=-0.26$ $p=0.268$
Proportion of Real Object Selection in Uncommon misarticulations	$r=0.139$ $p=0.558$	$r=-0.361$ $p=0.118$	$r=0.094$ $p=0.693$	$r=0.055$ $p=0.818$

Table 5: Correlations between proportion of real object selections when hearing each Word Type x Substitution condition and age, exposure to children who misarticulate, GFTA-2 scores, and PKP scores.

Correlations between the phonological production accuracy of the participants and the dependent variables were explored. When comparing percent real object selections for common misarticulations with independent variables, no significant correlations were revealed (*see* Table 5). Examination of the other dependent variables produced similar nonsignificant results.

VIII. Discussion

The purpose of the present study was to determine if children interpreted words containing speech sound substitutions as phonetic variants of known real words or as unknown novel words, and if processing is affected by these substitutions more than in canonical productions of words. Furthermore, the present study aimed to investigate if the child's experience with the substitution pattern affected children's interpretation and processing. Namely, this study attempted to tease apart whether preschoolers treat words containing speech sounds substitutions that are commonly made in childhood phonological development as phonetic variants of known real words more often than those substitutions that are rarely made.

The results showed that children selected real objects for canonical productions more often than for misarticulated productions. This suggests that children have very clear representations of the phonological makeup of the canonical words. This result supports the findings of Creel (2012), and further establishes that preschool-aged children have developed a representation for words based on the most commonly occurring production: the canonical form. Results of the paired samples t-tests revealed that between substitution types, children chose more real objects for misarticulated productions with common substitutions than for those with uncommon substitutions, which further supports the hypothesis that the misarticulated productions that children hear most frequently are more often recognized as phonetic variants of real words. Recall that parents reported, on average, that their children were around children who

misarticulate approximately 58% of the time. Therefore, this frequency of hearing misarticulated words with common substitutions from peers broadens their representations for words and increases their awareness of phonological patterns that occur in younger children. This supports the exemplar theory for word recognition proposed by Goldinger (1996). This theory states that having more exposure to variability facilitates faster processing, and that acoustic idiosyncracies are not “normalized” prior to processing, but are retained as potential variations (Goldinger, 1996).

Despite the vast preference for choosing real objects over nonobjects, this came at a cost to processing time. Preschoolers made their selections significantly slower in misarticulation conditions than in canonical conditions, and children deviated more from the most direct response path. This suggests that children were not as certain of their perception and wavered in their response, which indicates they may have approached the nonobject before finally making their decision. Their longer processing of other options may also suggest that children consider other options when hearing variability in the speech signal. This hesitation in recognition likely affects the efficiency of speech perception. Since children in this case only heard each substitution type once, there was no opportunity for familiarization to variability type. Previous research suggests that when provided with increased experience to speech variability, such as that found in dialectal variation, infants as young as 12 months are able to recognize words in the presence of variation and this ability generalizes to novel voices (Schmale, Cristia, Seidl, & Johnson, 2010; Schmale & Seidl, 2009). Thus, if children were given greater experience with the specific stimuli in this study, their processing efficiency might improve as they learn this speaker’s production patterns. In addition, more constrained experience might be helpful in this regard. That is, there were no cues about the intended meaning of each production, making the

production ambiguous. If training had been provided where a misarticulated production was clearly paired with an intended meaning, the child potentially could rapidly learn the speaker's production patterns (McQueen, Tyler, & Cutler, 2012). This learning of the speaker's production patterns might in turn reduce processing time for misarticulated productions. This hypothesis awaits empirical testing.

The difference in real object selections between common and uncommon misarticulation conditions was significant; however, there was not a difference in reaction time nor in AUC. It is apparent that having experience in misarticulated speech does assist in word recognition, but does not reach the same advantage at the level of processing as canonical productions, which are the most frequently heard by the child. This provides another piece to the puzzle regarding children's recognition of speech variability. The number of exemplars that the child hears (common vs. uncommon) affects the interpretation of the misarticulated word, but it does not seem to affect temporal processing. For example, if a child hears the word "weaf" many times from peers, the interpretation becomes understood as a phonetic variant of the canonical word, "leaf." This same recognition does not occur for an uncommonly misarticulated production (i.e. "yeaf"). Therefore, it seems that although children's recognition is facilitated by having an increased amount of exemplars, the steps required that lead up to recognition are not significantly reduced. The child still must go through the identification of variability, and then search through potential entries before settling on whether the word is known or new. In perceiving the misarticulated speech of other children, preschoolers rely on contextual information to assist in their processing time deficits. Creel (2012) found a small effect for recognition of phonetic variants of words within predictable contexts, but suggested that this effect was so small due to a ceiling effect that resulted from children's bias to select real objects.

Further research should investigate if children's recognition of phonetic variants of words are affected by context. This would provide more information about how children perceive the speech of their peers and how they compensate for slowed processing time of these variants.

The present findings may provide insight into the observation that children with speech and language impairments face social struggles within their peer groups (Hirschfeld & Gelman, 1997; Rice, Hadley, & Alexander, 1993). Studies of children's social reaction to speech variability indicate that children do indeed identify members of their own social group based on speech patterns, prosodic cues, accentedness and language (Hirschfeld & Gelman, 1997). Preschoolers do not base their social preference of playmates on intelligibility, but on the similarity of the speech pattern to their own (Kinzler, Shutts, DeJesus, & Spelke, 2009). With the present findings, it is possible that children are responding to the misarticulation itself, and not necessarily word recognition, since children readily accept varied productions as real words the majority of the time. Typically developing children will produce speech sound errors, similar to what other typically developing children will make, and may be able to use their discrimination abilities and the increased exposure to common misarticulations for word recognition. Children who have speech sound disorders will have a larger amount of misarticulations, and they may not be the same as those that are made by peers. Therefore, it may be more difficult for a typically developing child to understand a child with a speech sound disorder, and the child with the speech sound disorder may have more difficulty due to their own large number of substitutions as well as potentially having difficulty in phonemic discrimination. Further research should investigate the social effect of misarticulation by asking the children what they think about what they hear, or by asking them to choose the "best" token of a word. Moreover it would be

beneficial to tease apart if children are responding negatively to the increased processing time that it takes to receive the intended message.

IX. Limitations and Future Directions

The present study investigated whether children perceived real words containing speech sound substitutions as phonetic variants of real words or as novel words. The speaker for the present study is an adult, and therefore, may yield more hesitated results than if child productions were used. That is, misarticulations are associated with child speech rather than adult speech. The obvious vocal cues that the speaker was an adult may have been at odds with the misarticulated speech. If a child speaker had been used, these types of conflicting cues would be removed, and the stimuli would have been more consistent with the participant's past listening experience (i.e., child speakers misarticulate in certain ways). Future research should investigate the effect of children's naturally misarticulated speech on children's perception of words. This could provide more information about what cues children rely upon for speech perception and word learning. Past research suggests that there may be additional subtle phonetic cues when a misarticulated word is intended as a phonetic variant of a target word (e.g., Macken & Barton, 1980; Edwards, Schellinger, Beckman & Meyer, 2010). That is, children may produce an acoustic distinction between a sound being used as a target appropriately (e.g., [t] for /t/) and a sound being used as a substitute for a different target (e.g., [t] for /d/), where the substitute version (e.g., [t] for /d/) will have acoustic characteristics closer to the target (e.g., /d/). For example, a child may produce [t] for target /t/ with the appropriate long voice onset time (VOT) of an adult voiceless stop. That same child may then produce [t] as a substitute for target /d/ but the production may have a longer VOT than when [t] is produced for target /t/. Thus, the [t] for /d/ production moves closer to the typical VOT for target /d/ but crucially not close enough to be

perceived as the target /d/. This acoustic variation, if present, may further add interpretation that a misarticulated production is actually intended as a phonetic variant of a target word, and may also facilitate processing.

The present study only investigated children's perception and processing of words with speech sound substitutions. It did not set out to investigate children's opinion of the speech or their tacit awareness of what they were hearing. Future studies should include a component that asks children to rate the productions they are hearing to determine if they are aware of the accuracy and what they think they are hearing. It has been established that children do make social decisions based on how peers speak, but what they are aware of, and what their opinion is, is uncertain. This has obvious ramifications for understanding how misarticulations influence social communication.

Finally, this study reflects the perception of words by children who are typically developing. Future research should investigate the effect of misarticulated speech in children who have phonological impairments and in children who have language impairments in an attempt to isolate if any differences exist in the perception of speech variability as it relates to word learning in these groups. It is possible that children who have difficulty discriminating between phonemes may behave very differently from typically developing children. Moreover, children who have difficulty learning new words, such as those with Specific Language Impairment, may show different patterns of processing than children who are typically developing. This would reveal if there is an underlying mechanism that is present or absent in these populations in the perception of words or sound patterns that may slow their acquisition of vocabulary or their development in speech perception.

X. Conclusion

The present study investigated children's perception of misarticulated speech. The intent was to determine if children accept misarticulated speech as phonetic variants of real words, or if they interpret them as novel words, and the effect that misarticulation had on reaction time and mouse path trajectory. Further, the present study examined whether children were globally accepting of words containing speech sound substitutions, or if they were sensitive to the prevalence of the substitution in typically developing speech. It was found that children generally perceive misarticulated speech as phonetic variants of real words over novel "unnamed" objects but there was a processing cost in accommodating these variations from the canonical production. Moreover, the frequency of the misarticulation (i.e., common vs. uncommon) influenced children's interpretation of the misarticulation, but not their processing. Overall, this pattern of findings fits well with an exemplar model of the lexicon whereby the child's relative experience with canonical productions, common misarticulated productions, and uncommon misarticulated productions directly influences the likelihood that a given production will be interpreted as a particular target word. Although there is a clear processing cost in accommodating misarticulated productions, this processing cost is not influenced by experience with the misarticulation. This finding suggests that the processing steps involved in accommodating misarticulations are not influenced by general experience, although they could be influenced by experience with a particular speaker.

XI. References

- Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer.
- Creel, S. C. (2012). Phonological similarity and mutual exclusivity: on-line recognition of atypical pronunciations in 3–5-year-olds. *Developmental Science, 15*(5), 697-713.
- Dinnsen, D. A., & Gierut, J. A. (2008). Optimality theory, phonological acquisition and disorders.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171*(3968), 303-306.
- Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods, 42*(1), 226-241.
- Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1166.
- Hirschfeld, L. A., & Gelman, S. A. (1997). What young children think about the relationship between language variation and social difference. *Cognitive Development, 12*(2), 213-238.
- Jimenez, S., & Creel, S. (2012). Differences in talker recognition by preschoolers and adults. *Journal of experimental child psychology*.
- Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition, 27*(4), 623.
- Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning and Verbal Behavior, 23*(1), 39-66.
- Lieberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of experimental psychology, 54*(5), 358-368.
- Macken, M. A., & Barton, D. (1980). The acquisition of the voicing contrast in English: a study of voice onset time in word-initial stop consonants. *Journal of Child Language, 7*(01), 41-74.
- McCormack, H. M., Horne, D. J., & Sheather, S. (1988). Clinical applications of visual analogue scales: A critical review. *Psychol Med, 18*(4), 1007-1019.
- McQueen, J. M., Tyler, M. D., & Cutler, A. (2012). Lexical retuning of children's speech perception: Evidence for knowledge about words' component sounds. *Language Learning and Development, 8*(4), 317-339.
- Merriman, W. E., Bowman, L. L., & MacWhinney, B. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the society for research in child development, i*-129.

- Munson, B., Edwards, J., Schellinger, S. K., Beckman, M. E., & Meyer, M. K. (2010). Deconstructing phonetic transcription: covert contrast, perceptual bias, and an extraterrestrial view of Vox Humana. *Clinical linguistics & phonetics*, 24(4-5), 245-260.
- Rice, M. L., Hadley, P. A., & Alexander, A. L. (1993). Social biases toward children with speech and language impairments: A correlative causal model of language limitations. *Applied Psycholinguistics*, 14, 445-445.
- Schmale, R., Cristia, A., Seidl, A., & Johnson, E. K. (2010). Developmental changes in infants' ability to cope with dialect variation in word recognition. *Infancy*, 15(6), 650-662.
- Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: Flexibility of early word representations. *Developmental Science*, 12(4), 583-601.
- 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.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental psychology: Human learning and memory*, 6(2), 174.
- Storkel, H. L., & Adlof, S. M. (2009). The effect of semantic set size on word learning by preschool children. *Journal of Speech, Language and Hearing Research*, 52(2), 306.
- Storkel, H. L., & Hoover, J. R. (2010). An online calculator to compute phonotactic probability and neighborhood density on the basis of child corpora of spoken American English. *Behavior Research Methods*, 42(2), 497-506.
- Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child development*, 466-472.
- Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child development*, 349-355.
- Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, 37(2), 278-286.

Appendix A

Child History Questionnaire

Child's Name: _____

Child's Date of Birth: _____

Child's School: _____

Child's Teacher: _____

Parent's Name: _____

Parent's Address: _____

Parent's Home Phone: _____

Parent's Other Phone Numbers (that you would like us to contact you at):

Parent's E-mail (if you would like to receive scheduling information by e-mail):

How did you hear about the Word and Sound Learning Lab?

Research Identification Number (this will be filled in by staff): _____

Child's Current Age: _____ Child's Current Grade in School: _____

Child's Race (The National Institutes of Health requires this information):

American Indian/Alaska Native Asian Black/African American Native Hawaiian/Other Pacific Islander White

Child's Ethnicity (The National Institutes of Health requires this information):

Hispanic/Latino Not Hispanic/Latino

Child's Gender: MALE FEMALE

Speech/Language Development

1. Do you have any concerns about your child’s speech/language development? YES NO

If yes, please describe your concerns below:

2. Has your child received any special services for his speech/language development? YES NO

If yes, please describe type of services and general years/ages when received, including current services:

3. Have members of the child’s immediate (e.g., biological parent/sibling) or extended family (e.g., biological grandparents/aunts/uncles/cousins) received speech/language services? YES NO

If yes, please describe the speech/language services received by family members (specific family members should be identified only by their relationship to the child. For example, child’s aunt received services for reading)

4. How does your child make his/her needs known (circle)

sentences phrases 1-2 words sounds gestures

5. How well does your child understand what is said to him/her? (circle)

all most some none of the time

6. Please list all languages spoken in your home:

7. Please list all languages your child speaks:

8. Does your child have siblings? YES NO

a. If so, list the ages of your child’s siblings in years and months (e.g. brother 2 years 3 months)

9. How often is your child around children who pronounce words incorrectly, such as “weaf” instead of “leaf?” (Make a mark on the line below)

Never |-----| Frequently

Appendix B

Approved by the Human Subjects Committee University of Kansas, Lawrence Campus (HSCL). Approval expires one year from 12/7/2012. HSCL #20562
--

KU WORD AND SOUND LEARNING CONSENT STATEMENT *Influence of Misarticulation on Preschoolers' Word Recognition*

Your child is invited to participate in a research program on speech production and word recognition conducted by the Word and Sound Learning Lab at the University of Kansas. The Department of Speech-Language-Hearing at the University of Kansas supports the practice of protection for human subjects participating in research. The following information is provided for you to decide whether you wish your child to participate in the present study. You may refuse to sign this form and not allow your child to participate in this study. You should be aware that even if you agree to allow your child to participate, you are free to withdraw at any time. If you do withdraw your child from this study, it will not affect your relationship with this unit, the services it may provide to you, or the University of Kansas.

PURPOSE

The purpose of the research project is to better understand the relationship between speech that children hear and produce and their ability to understand words.

Your child is asked to participate in the task(s) marked below at his or her preschool. For preschool programs, the classroom teacher or leader will be consulted for scheduling issues.

PROCEDURES

Preliminary Evaluation: Prior to or during the preliminary evaluation, you will be asked to complete a questionnaire related to your child's development and environment. During the preliminary evaluation, your child will be given several tests commonly used by speech therapists to examine hearing, articulation, language comprehension, a supplemental speech probe, and to pre-expose children to the test materials. It is anticipated that this testing will require one to two 30-60 minute sessions. The following tests may be given the evaluation session to determine whether your child is eligible to participate in the research project.

The *hearing test* requires that your child wear headphones and listen to tones. This task usually takes 15 minutes.

The *vocabulary comprehension test* requires your child to listen to words spoken by the investigator and point to the correct picture from four choices. This task usually takes 15 minutes.

The *articulation test* requires your child to name common pictures such as “house” or “cup.” This task usually takes 15 minutes

Some children will be given a *supplemental speech probe* that requires them to name pictures. This task usually takes 10 minutes.

The *pre-exposure task* requires your child to listen to the names of pictures and use a mouse to point to pictures after they hear them. This task usually takes 15 minutes.

The articulation test will be video recorded and used to transcribe your child’s production of each word. All tests will be video recorded and used to score your child’s performance. Video recordings may be used to train research assistants and undergraduate/graduate students in research or speech-language pathology procedures. If these recordings are used for teaching purposes, your child’s name will be deleted from the recording. All video recordings will be stored in a locked cabinet in a room where access is limited to research personnel.

Word Recognition task: In this study, your child will see a screen with two pictures: one real object and one non-object. Your child will hear “Click on the _____.” followed by one of three pronunciations of the real object: a standard pronunciation (e.g. “chick” for “chick”), a pronunciation with a typical speech sound substitution (e.g. “shick” for “chick”), or a pronunciation with an atypical speech sound substitution (e.g. “fick” for “chick”). This should take approximately one to two sessions of 30-60 minutes each.

The word recognition task will be video recorded. These recordings will be used to identify your child’s responses. All video recordings will be stored in a locked cabinet in a room where access is limited to research personnel.

RISKS

The potential for risks, stress, or discomfort is largely non-existent. It is possible that your child may become bored or tired during testing sessions. Children will be given breaks if this occurs.

BENEFITS

In terms of direct benefits, you will be given a written report detailing your child’s performance on all standardized clinical tests administered during the preliminary evaluation which may be useful in educational planning. In addition, participation in this research program may help to improve your child’s vocabulary because he/she may learn some of the 12 words used in the study. In terms of indirect benefits, the findings from this study will be used to better understand the word recognition skills in typically developing children and the interaction between word recognition and speech development.

PAYMENT TO PARTICIPANTS

You will not be paid for your child’s participation in any aspect of this research.

INFORMATION TO BE COLLECTED

To perform this study, researchers will collect information about your child. This information will be obtained from a history questionnaire completed by you during the preliminary evaluation. Also, information will be collected from the study activities that are listed in the Procedures section of this consent form. Your child's name will not be associated in any publication or presentation with the information collected about your child or with the research findings from this study. Instead, the researcher(s) will use a study number or a pseudonym rather than your child's name.

The information collected about your child will be used by: Breanna Steidley (principal investigator), members of the research team, KU's Center for Research and officials at KU that oversee research, including committees and offices that review and monitor research studies. In addition, Breanna Steidley and her team may share the information gathered in this study, including your child's information, with: collaborating researchers and colleagues. The purpose of these disclosures is to facilitate completion of the research (e.g., analysis of the collected data) and safety monitoring. Again, your child's name would not be associated with the information disclosed to these individuals. Some persons or groups that receive your health information as described above may not be required to comply with the Health Insurance Portability and Accountability Act's privacy regulations, and your health information may lose this federal protection if those persons or groups disclose it.

The researchers will not share information about you with anyone not specified above unless (a) it is required by law or university policy, or (b) you give written permission.

Permission granted on this date to use and disclose your information remains in effect indefinitely. By signing this form you give permission for the use and disclosure of your child's information, excluding your child's name, for purposes of this study at any time in the future.

PARTICIPANT CONFIDENTIALITY

Your child's name will not be associated in any publication or presentation with the information collected about your child or with the research findings from this study. Instead, the researcher(s) will use a study number or a pseudonym rather than your child's name. Your child's identifiable information will not be shared unless (a) it is required by law or university policy, or (b) you give written permission.

Permission granted on this date to use and disclose your information remains in effect indefinitely. By signing this form you give permission for the use and disclosure of your child's information, excluding your child's name, for purposes of this study at any time in the future.

REFUSAL TO SIGN CONSENT AND AUTHORIZATION

You are not required to sign this Consent and Authorization form and you may refuse to do so without affecting your right to any services you are receiving or may receive from the University of Kansas or to participate in any programs or events of the University of Kansas. However, if you refuse to sign, your child cannot participate in this study.

CANCELLING THIS CONSENT AND AUTHORIZATION

You may withdraw your consent to allow participation of your child in this study at any time. You also have the right to cancel your permission to use and disclose further information collected about your child, in writing, at any time, by sending your written request to: Dr. Holly L. Storkel, at Department of Speech-Language-Hearing, 1000 Sunnyside Avenue, 3001 Dole Center, Lawrence, KS 66045-7555 (785-864-0497; 785-864-4873; hstorkel@ku.edu).

If you cancel permission to use your child's information, the researchers will stop collecting additional information about your child. However, the research team may use and disclose information that was gathered before they received your cancellation, as described above.

QUESTIONS ABOUT PARTICIPATION

Questions about procedures should be directed to the researcher listed at the end of this consent form.

PARTICIPANT CERTIFICATION

I have read this Consent and Authorization form. I have had the opportunity to ask, and I have received answers to, any questions I had regarding the study and the use and disclosure of information about my child for the study. I understand that if I have any additional questions about my rights as a research participant, I may call (785) 864-7429 or write the Human Subjects Committee Lawrence Campus (HSCL), University of Kansas, 2385 Irving Hill Road, Lawrence, Kansas 66045-7568, e-mail irb@ku.edu.

I agree to allow my child to take part in this study as a research participant. By my signature I affirm that I have received a copy of this Consent and Authorization form.

_____ Date

_____ Type/Print Participant's Name

_____ Parent/Guardian Signature

Researcher Contact Information
Breanna I. Steidley, M.A.
Graduate Student

Holly L. Storkel, Ph.D.
Associate Professor

Department of Speech-Language-Hearing
1000 Sunnyside Avenue
3001 Dole Center
University of Kansas
Lawrence, KS 66045
785 864 - 0497

KU WORD AND SOUND LEARNING CONSENT STATEMENT

Influence of Misarticulation on Preschoolers' Word Recognition

Assent Procedures for Preschool Children

Because there are several different types of tasks that children are asked to do, assent procedures will differ across sessions to accurately reflect the tasks that a child is being asked to do in a given session. Thus, assent must be obtained at the beginning of each session.

Preliminary Evaluation

My name is [insert PI/GRA/staff name], and I am learning about how kids listen to words and ways that they understand what they hear. I would like you to work with me today for 1/2 an hour to 1 hour. I would like you to play [insert number of tasks] games. The [insert number of tasks] games are: [insert description of tasks from below depending on session]. If you don't feel like working with me, you don't have to. You can stop at any time and that will be all right. Do you want to do this with me?

The *hearing test*: “wear headphones and listen to tones”

The *vocabulary comprehension test*: “listen to words and point to a picture”

The *articulation test*: “name pictures”

The *supplemental probe*: “name pictures”

The *familiarization task*: “listen to words and click on a picture”

Word recognition task

My name is [insert PI/GRA/staff name], and I am learning about how kids listen to words and ways that they understand what they hear. I would like you to work with me today for 30-60 minutes. Today we will listen to words. After you hear a word, you get to use this mouse to click on the picture you hear. If you don't feel like working with me, you don't have to. You can stop at any time and that will be all right. Do you want to play on the computer with me?

Responses

Positive Responses: (1) Verbal responses such as “yes;” (2) Nonverbal gestures such as a head nod; (3) Nonverbal signals such as putting away whatever the child was working on and getting up to accompany the investigator.

Negative Responses: (1) Verbal responses such as “no” or “I’m busy;” (2) Nonverbal gestures such as head shake; (3) Nonverbal signals such as crossing hands over chest and pouting or crying or refusing to put work away and accompany the investigator.

Consequences

Positive Response: Child and examiner will move to research space and begin research tasks.

Negative Response: Research will not be initiated. The refused task or set of tasks will be attempted up to two more times. After a total of three refusals of a given task/set of tasks, the task/set of tasks will not be attempted again. However, the next task/set of tasks in the research protocol will be attempted one time. If the next task/set of tasks is refused one time, the research will be discontinued. If the next task/set of tasks is assented to, research will continue as planned and any future refusals will follow the same protocol (i.e., up to 3 attempts before moving to next task and discontinuation if refusal on the following task).

The rationale for this protocol is that children may refuse tasks for a variety of reasons (e.g., tired, more interested in an ongoing activity). Thus, multiple attempts are needed to determine whether the child doesn’t want to do the task *at all* versus doesn’t want to do the task *at this moment*. Likewise, a second task is attempted to determine whether the child doesn’t want to participate in research *at all* versus doesn’t want to participate in *certain research tasks*.

Withdrawal of Assent: A child can withdraw assent during a procedure as indicated by the negative responses already outlined (e.g., crying in the middle of the task) as well as additional responses indicating that the child does not wish to continue (e.g., “I want to go back to class;” “I’m tired of this;” “I don’t want to do this anymore.”) Withdrawal of assent will follow the same procedures as a negative response to the assent question. That is, the withdrawal is counted as one negative response and the task can be re-attempted up to two more times.

