THE NATURE OF PHONOLOGICAL REPRESENTATIONS IN ADULTS AND CHILDREN: EVIDENCE OF MISPRONUNCATION DETECTION

By

ERIN MARIE CHUTE

A Thesis Submitted to the Honors College

In Partial Fulfillment of the Bachelors degree
With Honors in

Speech, Language and Hearing Sciences

THE UNIVERSITY OF ARIZONA

MAY 2011

Approved by:

Dr. Edwin Maas
Department of Speech, Language and Hearing Sciences

The University of Arizona Electronic Theses and Dissertations Reproduction and Distribution Rights Form

Name (Last, First, Middle) Chute, Erin Marie	
Degree title (eg BA, BS, BSE, BSB, BFA): BS	
Honors area (eg Molecular and Cellular Biology, English, Studio Art): Speech, Language and Hearing Sciences	
Date thesis submitted to Honors College: April 27, 2011	
Title of Honors thesis: The Nature of Phonological Representations in Adults and Children: Evidence of Mispronunciation Detection	
:The University of Arizona Library Release	I hereby grant to the University of Arizona Library the nonexclusive worldwide right to reproduce and distribute my dissertation or thesis and abstract (herein, the "licensed materials"), in whole or in part, in any and all media of distribution and in any format in existence now or developed in the future. I represent and warrant to the University of Arizona that the licensed materials are my original work, that I am the sole owner of all rights in and to the licensed materials, and that none of the licensed materials infringe or violate the rights of others. I further represent that I have obtained all necessary rights to permit the University of Arizona Library to reproduce and distribute any nonpublic third party software necessary to access, display, run or print my dissertation or thesis. I acknowledge that University of Arizona Library may elect not to distribute my dissertation or thesis in digital format if, in its reasonable judgment, it believes all such rights have not been secured. Signed: Erin M. Chute Date: 4/26/11

Last updated: Nov 15, 2009

ABSTRACT

This thesis explored the controversial issue of whether children process words via holistic representations, or via segmental representations like adults. The Lexical Restructuring Hypothesis (LRH) states children have whole-word or -syllable representations, whereas the Segmental Hypothesis (SH) claims children have segment-sized representations like adults. Another question this thesis asked was whether children with speech sound disorders have similar representations as their typically-developing peers. The thesis used data from typically-developing children, children with speech sound disorders, and adults. The study used a Mispronunciation Detection task, in which participants indicated whether a word was pronounced correctly or incorrectly. There were five conditions: initial-segment, final-segment, two-segment, foil, and correct. Adults and typically-developing children made more errors in the initial-segment condition than in any other condition. Children also showed more errors in the final-segment condition than in the two-segment condition. The results support the SH because if children were only sensitive to featural differences, they would have made more errors in the two-segment condition, in which there were fewer features on one segment making differences more difficult to detect. Children with speech sound disorders demonstrated similar phonological representations in comparison to their typically-developing peers.

INTRODUCTION

The process of speech production involves the coordination of multiple complex processes. Speaking only a single word implements processes such as lexical retrieval (finding the word), phonological planning (finding the sounds of the word) and speech motor planning (finding the movements to produce those sounds) (e.g., Shiller, 2006). While much research has been conducted in regards to adult speech production, relatively little is known about how children plan their speech. This thesis asks the question of whether children represent words in a holistic manner (the whole word) or if they represent words through individual sounds like adults. It focuses on answering this question using a Mispronunciation Detection (MPD) task, in which children have to indicate whether a spoken word is pronounced correctly or not. Another question this thesis asks is whether children with speech disorders have different representations than typically-developing children.

Phonological encoding in adults

Speech production in a simple naming task begins with visual processing and access to the depicted concept. The speaker then retrieves the word with its semantic and syntactic information, for example, the word “ball” is represented by ideas like “round,” “bounces,” “rubbery,” etc, and the fact that it is a singular noun. From the semantic planning, the word is encoded in the phonological planning stage, that is, the sounds (phonemes) and the stress pattern are retrieved. The process continues to the speech motor planning stage, where appropriate articulatory movements are specified to produce the sounds. Finally, articulators are set into motion and speech is produced.

In adults, phonological encoding is thought to involve phonemes, rather than syllables (e.g., Schiller, 2006). This idea is based on both speech error evidence and evidence from various priming tasks, including auditory picture-word inference (e.g., Seiger-Gardner & Schwartz, 2008) and implicit priming (e.g., Meyer, 1990; Roelofs, 1999).

Phonological representations in children

This thesis project seeks to answer the question of how children represent the sound structure of their words, either holistically or through segments like adults. The Lexical Restructuring Hypothesis (LRH) and the Segmental Hypothesis (SH) both attempt to answer this question. According to the LRH, children begin with holistic representations, consisting of whole words or syllables, and then reformulate those representations into smaller parts corresponding to the phonemes of the language (Walley, 1987). The SH suggests that children already have segments based on phonological representations and do not develop them from larger holistic units, initially (Gerken et al., 1995).

A common experimental task used to assess phonological representations without needing to tap into the articulation aspect is the MPD task (e.g., Bowey & Hirakis, 2006; Sutherland & Gillon, 2005). In such a task, children typically see a picture and hear a spoken word or nonword. The children then need to decide whether the word for the picture representation was correctly pronounced. A study done by Sutherland and Gillon (2005) looked at phonological representations in children with and without speech impairment using an MPD task. The MPD task required children in the study to determine the correctness in production of words in comparison to their individual phonological representation systems. The children were shown single-referent images on a computer screen (e.g., helicopter, caterpillar, watermelon, etc). The child then heard auditory output to match each word. The child had to select a “happy face” if the word he/she heard was a good production and a “large black cross” if the word he/she heard was a poor production. Children with speech impairment made more inaccurate judgments of mispronounced speech sounds, which suggested difficulties in recalling phonological representations or incorrectly specified representations. The results of the study indicate that phonological judgment tasks are appropriate to show phonological representations in children with and without speech impairments.

Some evidence from this task for the LRH was provided by Walley (1987). Walley (1987), used picture and non-picture mispronunciation tasks that also looked at the receptive input for measuring children’s phonological representations as explained in Sutherland and Gillon (2005). The children ranged in age from 4-5 years old and were tested with words in both isolation and in story-form. In a constrained (story) context, a position effect was detected by both the older and younger children, meaning that word-initial errors were detected more accurately than errors in word-final position. For example, a word such as “tire” with a mispronunciation of /t/ as /k/ (“kire”) was detected more often than a mispronunciation with the word “toast” with /k/ replacing the /t/ final sound (“toask”). Another experiment with isolated words, using picture and no-picture conditions, replicated this position effect for the only children only. The younger children (4-year olds) showed no position effect with or without the picture, meaning the accompanying picture did not affect the child’s awareness to mispronunciations in word-initial or -final positions. For the older children (5-year olds) the effect was only obtained when a picture accompanied the stimuli. They were more sensitive to word-initial mispronunciations, and in both experiments they had a greater position effect than the younger children. It is this emergence of the position effect and the greater magnitude of the position effect for older children that were taken as support for the LRH. Both studies further suggest that children have a greater position effect in constrained contexts, either amidst a story or an isolated word with a picture referent. The position effect is not present in unconstrained contexts, such as a word in isolation with no picture referent.

Although the emergence of the position effect has been used as evidence in favor of the LRH, an alternative explanation for this finding has recently been proposed. In particular, Bowey and Hirakis (2006) suggested that the position effect could be an effect of acoustic clarity rather than restructuring. They mention that sounds in initial position tend to be in stressed syllables, have greater importance in retrieving lexical cues, and are therefore more reliable cues. For example, the sound /t/ in initial position (as in "tea") will always be aspirated, whereas the /t/ in final position (as in "eat") is never aspirated and may even be unreleased. Therefore, greater accuracy at detecting initial mispronunciations may be due to the greater salience rather than restructuring.

Bowey and Hirakis (2006) compared these two alternatives by examining the effects of both position and clarity of acoustic-phonetic information on MPD of singleton onsets in children and adults. The study provided evidence for the clarity hypothesis. Children were more sensitive to the mispronounced stressed word-initial /s/ sound in "circle" than the mispronounced unstressed second syllable onset, /s/, in "whi*stle.*" They were also more sensitive to the mispronounced /s/ in the onsets of second syllables with a clearer acoustic-phonetic position, such as in the word "ca*SSette*," than in onsets of unstressed second syllable onsets. Therefore, decreased sensitivity in mispronounced onsets of second syllables based on position is not a valid argument for lexical restructuring because the position effect in children and adults can be explained as effects of clarity rather than position.

Some positive evidence for segmental representations comes from an experiment by Gerken Murphy, and Aslin (1995). In this study, 3-4 year old children were asked to signal their recognition of a pre-specified target word, “lick,” by pressing a button. This task is similar to the MPD, except it maintains the same phonological form in memory to match the incoming sounds instead of retrieving unprepared sounds from a picture. The study by Gerken et al. (1995), examined whether or not the children detected a foil that differed from the target word (“lick”) by two features on one segment (e.g. “gik”) more often than a foil that differed by one feature on two segments (e.g. “zig”). The study found that children made more errors when the foil differed by one segment than when it differed by two segments (even though the number of differing features was the same). The findings suggest that children can and do use information about segments to evaluate mispronunciations and not just through representation of phonetic features. The research shows outcomes contrary to the LRH and in favor of the SH, in that children may represent their words in terms of segments (as well as features).

Limitations in the procedure of this study are that Gerken et al. (1995) used only one target word as a fixed target. The use of a fixed target reduces the need to retrieve phonological representations because the child does not need to plan a different word for each trial. Thus, although these findings are promising, they should be replicated using a larger set of items in a task that requires children to retrieve output phonological representations (e.g., from a picture). The present study was specifically designed to accomplish this task.

The Present Study

This study was designed to assess the segmental versus holistic nature of phonological representations in children with and without speech disorders and in adults. Based on the findings from the study by Gerken et al. (1995), we developed an MPD task, in which children have to retrieve a new phonological representation on each trial. We constructed a task that requires matching the incoming signal (spoken word) to an existing representation (retrieved from memory and represented by a picture) in order to judge correctness (Sutherland & Gillon, 2005). This allowed us to obtain data related to the child's encoding processes. Our study aims to produce more evidence related to these hypotheses (LRH and SH) using a wider range of items. In particular, this thesis used the same logic developed by Gerken et al. (1995) and compared two-feature mismatches on either a single segment (initial or final consonant) or distributed across two segments (initial and final consonant).

METHODS

Participants

The experiment was conducted with four typically-developing children ages 5-7 (mean age 6;8), three typically-developing children ages 8-10 (mean age 9;1), and 24 adult participants (mean age 21;4). The study also included three children with speech sound disorders all in the younger age group (mean age 5;6). Clinical diagnosis of speech sound disorder was made by a certified and experienced Speech-Language Pathologist. All participants were monolingual native English speakers.

The typically-developing children were categorized based on their a) articulatory skills no more than one standard deviation below the mean on the Goldman-Fristoe Test of Articulation-2 (GFTA-2; Goldman & Fristoe, 2000) and clinical judgment of typical speech development based on nonstandardized tasks, including elicitation of a conversational speech sample; b) adequate oral structure and function, based on clinical judgment from an oral mechanism exam; c) normal nonverbal intelligence based on subtest T-scores between 40 and 60 on the nonverbal portion of the Reynolds Intellectual Assessment Scales (Reynolds & Kamphaus, 2003); d) expressive and receptive language abilities within normal limits as determined by a standard score of 85 or greater on the Expressive Vocabulary Test – Second Edition (EVT-2; Williams, 1997) and the Peabody Picture Vocabulary Test – Fourth Edition (PPVT-4; Dunn & Dunn, 2007), respectively; and e) normal hearing indicated by passing a hearing screening at 500, 1000, 2000, and 4000 Hz at 20 dB (ASHA, 1997). Participant test scores are presented in Table 1.

Children with speech sound disorders met all the criteria above, except that they received a primary diagnosis of a speech sound disorder not due to structural abnormalities, indicated by a score below the 16th percentile on the GFTA-2 (Goldman & Fristoe, 2000), or clinical judgment of vowel errors and prosodic abnormalities, which are not captured by the GFTA-2 (relevant in the case of participant 300). Participants were recruited from the University of Arizona and Tucson communities, respectively. Children with speech sound disorders were recruited from the University clinic, the University-affiliated Child Language Center, local schools and Speech-Language Pathology practices.

TABLE 1. ***Participant Information***

Group	Part.#	Sex	Age	EVT	PPVT	GFTA
SSD	300	M	5;6	120 (91st)	117 (87th)	96 (29th)
	303	F	5;5	93 (32nd)	96 (39th)	91 (14th)
	304	M	5;7	91 (27th)	105 (63rd)	44 (<1st)
TD5	100	F	7;8	107 (68th)	116 (86th)	105 (56th)
	101	M	5;11	111 (77th)	109 (73rd)	100 (25th)
	102	F	7;2	99 (47th)	100 (50th)	107 (>60th)
	305	F	6;7	116 (86th)	142 (99.7th)	108 (>65th)
TD8	200	M	10;10	104 (61st)	112 (79th)	102 (24th)
	201	M	8;1	112 (79th)	150 (98th)	86 (6th)
	202	F	8;4	112 (79th)	127 (96th)	104 (>48th)
Adults	N=24	21F;3M	21;8 (3;7)	113 (14)	110 (14)	N/A

Materials

The task included (1) a picture of an alien (Ohala, 1999; see Figure 1), (2) pictures of 19 objects and animals familiar to children, and (3) recordings of mispronounced and correctly pronounced names of the pictures. Pictures were scaled to fit in the same size and location on the computer screen (see Figure 1 for example). These pictures were also used in a separate naming experiment (not discussed in this thesis) conducted with these same participants prior to the MPD task. Thus, participants were familiar with these pictures and their names.

High quality audio recordings were made by a female native speaker of American English. Recordings were verified for accuracy by an independent listener and edited to remove any silent periods before or after the utterance (using Adobe Audition v 1.5). Mispronunciations involved two-feature errors, with the two phonetic feature mismatches either on a single segment or distributed over two segments. For example, mispronunciations of the target "fish" included "bish" (initial segment differs in place and manner of articulation), "fip" (final segment differs in place and manner of articulation), and "shitch" (two segments: initial differs in place, final differs in manner of articulation).

The explanation for this thesis was that if children have segmental representations, then a mispronunciation involving two different segments should be easier to detect than a mispronunciation in which only one segment is incorrect. If children do not have segmental representations, then they will detect mispronunciations based on featural differences. If that is the case, there should be no differences between the one-segment and two-segment conditions because the number of incorrect features is the same (Gerken et al., 1995). There were two reasons for including both initial-segment mispronunciations and final-segment mispronunciations. One reason was to potentially replicate previous studies that have observed a position effect, which supports the LRH (e.g., Walley, 1987; but see Bowey & Hirakis, 2006). The second reason was to encourage participants to listen to the entire word before making a decision of correctness.

In addition to these conditions, there was also a "foil" condition in which all segments, including the vowel, were different from the target (e.g., "boz" for "fish"). The foil conditions were used to ensure participants were following task instructions. The final condition included correct pronunciations to ensure "yes" responses within the task. Thus, in all, there were five conditions: initial-segment, final-segment, two-segment, foil and correct.

All mispronunciations were nonwords to avoid any influence from existing lexical representations, and were matched across conditions for several variables known to affect speech perception and speech production. These included length in syllables and phonemes (all were monosyllabic with CVC forms), spoken word duration (in milliseconds), phonotactic probability and neighborhood density (e.g., Storkel & Hoover, 2010; Vitevitch & Luce, 2004). Two phonotactic probability variables and two neighborhood density variables were controlled. The first controlled phonotactic probability variable included a summed phonotactic segment probability, which refers to the probability that a given segment occurs in a given word position, and includes the three segments of each nonword (e.g. /z/, /ɛ/, and /g/ for /zɛg/). The second consisted of the summed phonotactic sequence probability, which refers to the probability that two segments occur together, and include the two sequences of two segments of each nonword (e.g. /zɛ/ and /ɛg/ for /zɛg/). The first neighborhood density variable controlled included the neighborhood density, which refers to the number of real words that differ by only one segment from the nonword. The second included the mean logarithmic frequency of all the neighbors.

The information from phonotactic probability and neighborhood density was matched, or in other words, had no significant differences between conditions ($p > 0.05$) when using both adult and child norms. This is an important factor because the size of the lexicon is critical in determining these variables and is expected to be different between children and adults. We used existing databases with web-based resources to obtain this information (Storkel & Hoover, 2010; Vitevitch & Luce, 2004). All targets and nonword mispronunciations are presented in Table 2; the condition matching information is presented in Table 3.

TABLE 2. ***Items and Conditions***

Targets	Correct	1 Segment (Initial)	1 Segment (Final)	2 Segments	Foil
bed	bɛd	ðɛd	bɛð	mɛb	kiʤ
mouse	maʊs	ʤaʊs	maʊʧ	waʊf	bɔɪ
comb	koʊm	ʃoʊm	koʊʒ	toʊb	zɛg
wall	wɑl	zɑl	wɑm	jɑd	tæʃ
tail	teɪl	θeɪl	teɪj	keɪz	væp
knife	naɪf	ʤaɪf	naɪp	ʤaɪθ	zoʊt
pear	pɛr	gɛɹ	pɛm	kɛl	hɪn
moon	mun	zun	mug	nul	gʌʧ
doll	dɑl	vɑl	dɑð	tɑz	neɪg
road	roʊd	zoʊd	roʊð	joʊb	miθ
witch	wɪʧ	ðɪʧ	wɪʒ	vɪʃ	tɑʒ
net	nɛt	tɛt	nɛn	lɛk	bɪm
bowl	boʊl	loʊl	boʊs	poʊd	weɪm
hat	hæt	zæt	hæb	ʃæd	ðid
goal	goʊl	zoʊl	goʊm	koʊz	pæʤ
can	kæn	tæn	bɛð	pæz	ʤoʊd

TABLE 3. ***Condition Matching***

Means (SDs) of mispronunciation conditions for phonotactic probability, neighborhood density, and spoken word duration. There were no significant differences for any of the variables between any of the mismatch conditions for either the adult values or the child values (all p>0.05).

	Correct	Initial-Segment	Final-Segment	Two Segments	Foil
Child S-Sum	0.179 (0.043)	0.141 (0.035)	0.140 (0.036)	0.141 (0.040)	0.136 (0.045)
Child B-Sum	0.011 (0.006)	0.007 (0.004)	0.007 (0.003)	0.005 (0.003)	0.005 (0.004)
Child Neighbors	16.0 (5.7)	9.1 (5.7)	8.6 (2.5)	9.0 (5.0)	9.0 (4.7)
Child LF Mean	2.760 (0.312)	2.711 (0.262)	2.651 (0.440)	2.631 (0.359)	2.546 (0.298)
Adult S-Sum	0.167 (0.043)	0.130 (0.034)	0.133 (0.038)	0.133 (0.046)	0.123 (0.046)
Adult B-Sum	0.007 (0.006)	0.005 (0.004)	0.006 (0.004)	0.005 (0.003)	0.005 (0.004)
Adult Neighbors	24.5 (7.4)	14.8 (6.6)	12.6 (5.1)	15.3 (8.0)	13.9 (6.3)
Adult LF Mean	2.137 (0.222)	2.141 (0.219)	2.071 (0.222)	2.053 (0.268)	1.970 (0.282)
Word Duration (ms)	794 (132)	821 (104)	825 (140)	838 (152)	817 (164)

Child: data based on child corpus (Storkel & Hoover, 2010); Adult: data based on adult corpus (Vitevitch & Luce, 2004); S-Sum: sum of the frequency of occurrence of each segment; B-Sum: sum of the biphone probability; Neighbors: number of neighbors; LF Mean: mean of log frequency of neighbors

Task and Procedures

The MPD task required participants to indicate whether a spoken utterance was the correct pronunciation to match a target picture. This task was used because it requires speech perception as well as access to phonological output representations from a picture input, without requiring overt articulation of the picture name.

To create an interactive element for the task, especially for the children, participants were introduced to an alien (Ohala, 1999), a “Pollywolly,” named Molly, who was from another planet and was visiting earth to learn to speak English. Participants were told that the Pollywolly would name things around her, and that sometimes she said the word wrong. They were asked to help the Pollywolly learn English by pressing a “yes” button or a “no” button (labeled mouse buttons of a laptop) to communicate with the Pollywolly.

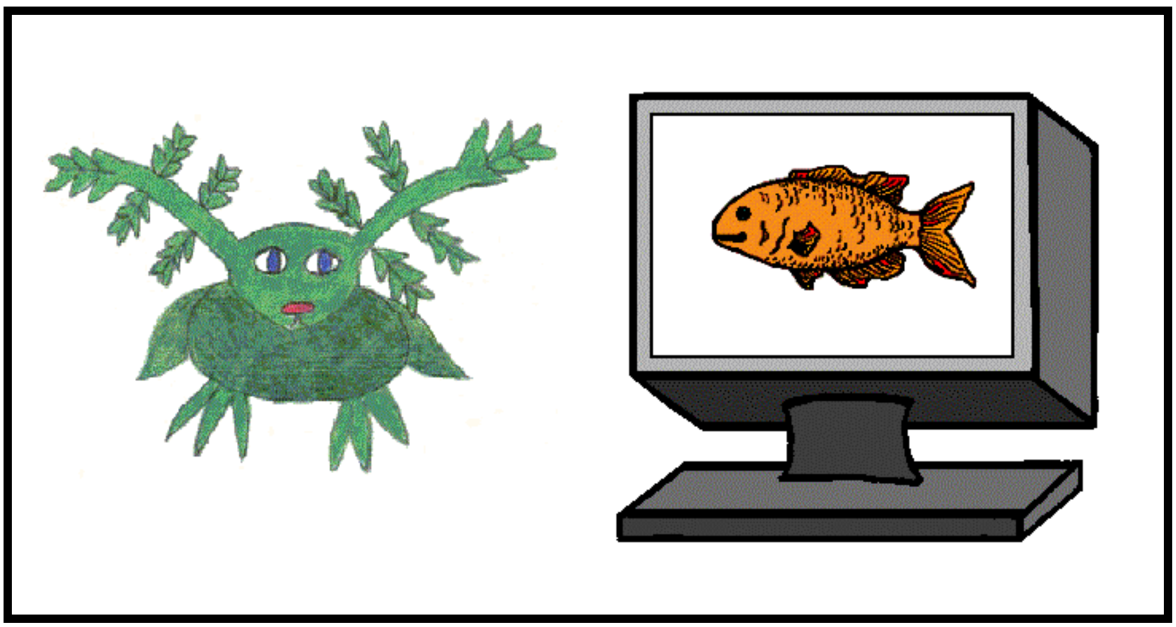
The experiment used E-Prime software (v.2; Psychology Software Tools, Inc.) and conducted on a Toshiba laptop. Sounds were presented over headphones (Sennheiser HD 280Pro for adults; Caliphone 2810 animal headphones for children) at a comfortable listening level determined by the participant during a sound check prior to the experiment. The computer was connected to a button box (Psychology Software Tools, Inc.) that allowed the experimenter to control the experiment. The experimenter pressed a button to present each next trial, but only when the participant was paying attention. This was determined by whether or not the participant looked at the screen when the picture appeared. If there was an external noise coinciding with the presentation of the sound (e.g. planes flying overhead), the experimenter pressed a button that signaled that particular trial to rerun at the end of the block. The experimental software automatically collected responses from the participant.

Each trial involved the following events: First the Pollywolly and a target picture appeared (see Figure 1), and the Pollywolly said the name (correctly or incorrectly). Sound file onset was simultaneous with picture onset. Next, the participant pressed one of the two buttons ("yes" for correct, "no" for incorrect), which removed the picture from the Pollywolly's screen. Finally, the experimenter pressed a button to advance to the next trial.

The five mispronunciation conditions (initial-segment, final-segment, two-segment, foil and correct) were counterbalanced across five blocks of 19 trials each (16 experimental items and 3 filler items). While each block contained 3 or 4 items in each of the five conditions, no block contained the same picture more than once, and trials were presented in random order in each block. The five blocks were divided over three experimental sessions to minimize decreasing responsiveness. Two sessions contained two blocks and one session contained one block. One extra block was added at the end in which the Pollywolly pronounced all picture names correctly.

This was inserted for the interactive purpose of showing the Pollywolly had learned from the participant. Overall, there were two sessions with three blocks and one session with two blocks. Rest blocks were provided between blocks as needed. Only the experimental blocks were used for analysis.

Figure 1. Example Screen of Mispronunciation Detection Task



Design, Analysis and Predictions

The main dependent variable was percent correct. The independent variable was condition, with five levels: initial-segment, final-segment, two segment, foil and correct. Data were analyzed separately for adults and children because of the different sample sizes. Data from adult participants were analyzed using a one-way Analysis of Variance (ANOVA), with post-hoc Tukey tests to identify the source of any significant effects. Data from children was analyzed using a 3 x 5 ANOVA, with one between-subjects factor with three levels (Group: 5-7 year olds, 8-10 year olds, children with SSD). It also used one within-subject factor with five levels

(Condition: initial-segment, final-segment, two-segment, foil and correct). Follow-up Tukey tests were used to identify the source of any significant effects. All statistical tests were performed with an alpha level of 0.05. The two hypotheses examined, the LRH and SH, will be explained in the results. If the LRH is supported, it will determine young children do not have segmental representations. Thus, they will show no difference between the one-segment conditions and two-segment condition because both these conditions have a mismatch of two phonetic features. In contrast, the SH will be supported if there is a one-segment vs. two-segment mismatch difference for all children and adults.

With respect to children with SSD, the pattern is expected to depend on the nature of their disorder. If the speech disorder is primarily structural and not phonological in nature (as in the case of 303), then normal performance is expected. For children with Childhood Apraxia of Speech (CAS) (as in the case of 300 and 304), phonological processing may also be intact, but there may be subtle phonological deficits present that would be difficult to uncover based on a speech production task.

RESULTS

Adults

In the adult subjects, there was a significant main effect of Condition, $F(4, 115) = 46.21$, $p < .0001$. Follow-up Tukey tests revealed that the initial-segment condition had significantly more errors than any other condition ($p < .0001$), as seen in Figure 2. No other comparisons reached statistical significance. As shown in Figure 2, the two-segment condition resulted in no errors.

Children

Results for children are presented in Figure 3. There was also a significant effect of Condition, $F(4,28) = 4.79$, $p = .0045$, in the children. The Group effect was not significant, $F(2,7) = 2.06$, $p = .1984$, nor was the Group x Condition interaction, $F(8,28) = 1.60$, $p = .1708$. Follow-up Tukey tests showed that the initial-segment condition produced more errors than any other condition ($p = .0256$); the final-segment condition produced numerically more errors than the two-segment condition, but this difference did not reach significance ($p = .1240$). Observation of the condition means for the different groups shows that that only the younger children and the speech-disordered children made substantial errors on the final-segment condition.

Further inspection of the results of the individual SSD children indicates child 300 and 303 reveal the Group pattern relative to that of the typically-developing children. That is, they elicited fewer errors in the two-segment condition than the initial-segment conditions (see Appendix). Child 304 shows an equal number of errors in these critical conditions.

Figure 2. Adult Results of MPD Task

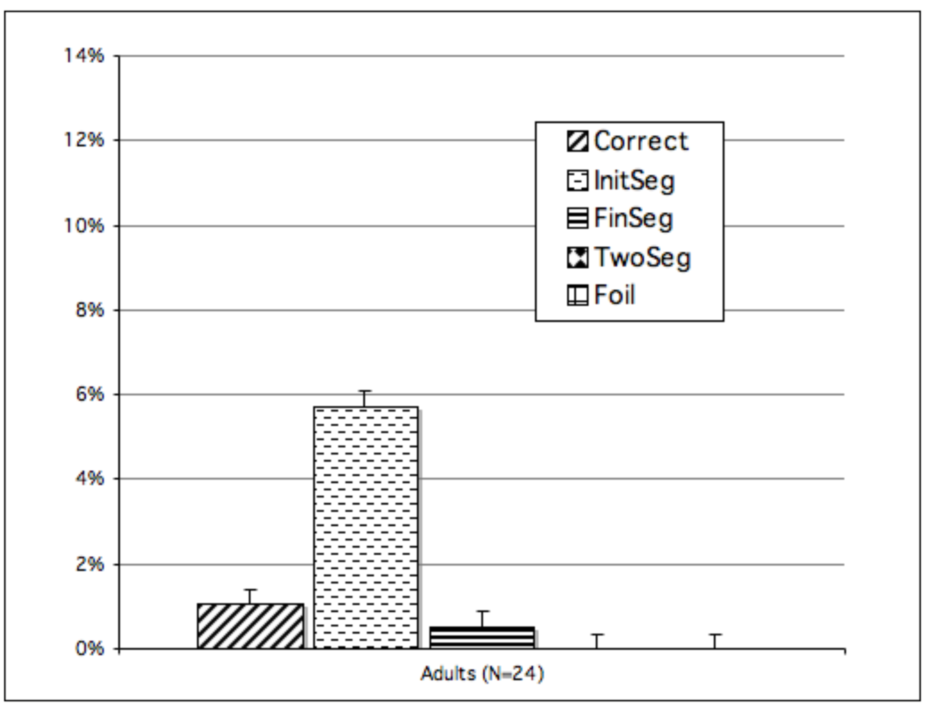
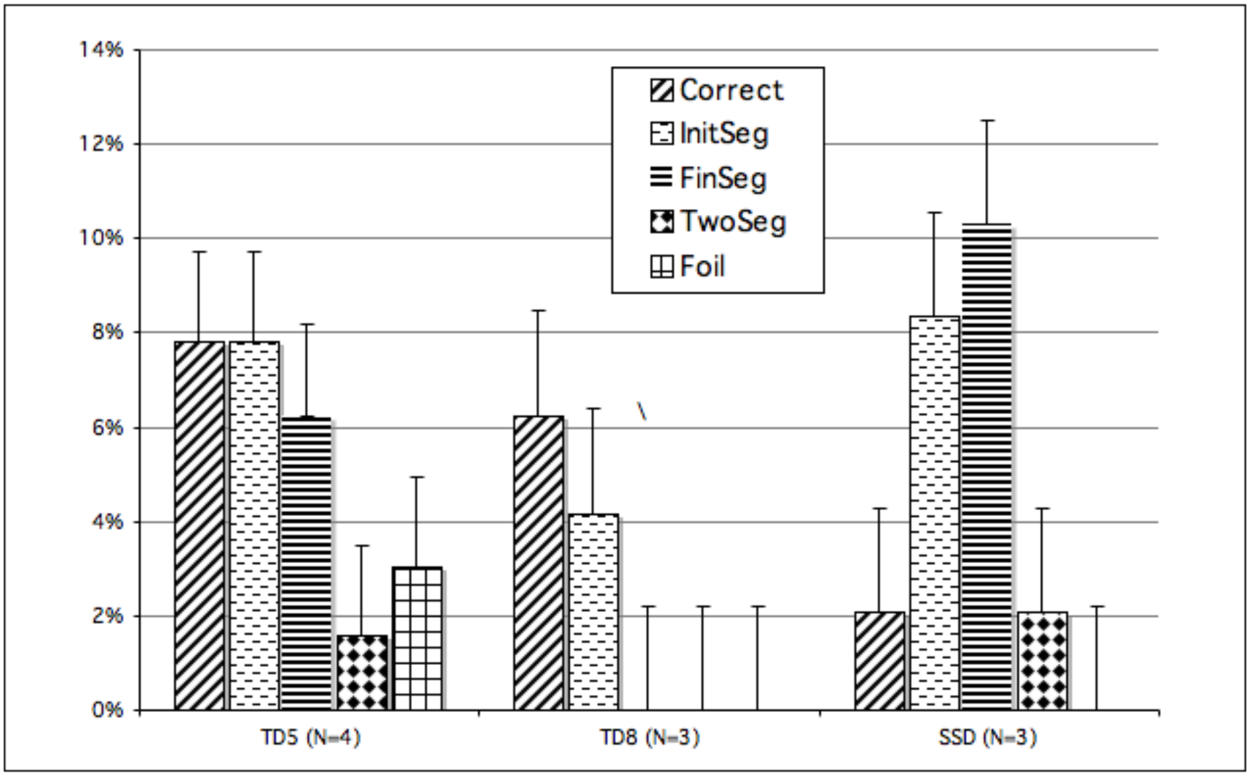


Figure 3. Child Results of MPD Task



TD5 refers to typically-developing 5-7 year old children; TD8 refers to typically-developing 8-10 year old children; SSD refers to children diagnosed with speech sound disorder

DISCUSSION

The purpose of this study was to determine whether children represent words holistically or through individual phonological segments. The thesis examined two existing hypotheses, the LRH and the SH. The LRH claims children are sensitive to featural differences but not segment differences, while the SH claims children form words based on segmental representations of the word. Walley (1987) determined that a position effect, greater sensitivity to word-initial segments, was only present when referents were contextualized (i.e. when pictures accompanied the stimuli or within a story). This finding is inconsistent with the results found in this study. Based on the results from the adult data, more errors were present in the initial-segment condition than in any other condition. Also, there were no errors in the two-segment condition. Therefore, adults were more sensitive to the final-segment condition and two-segment condition than one segment difference in the word-initial position. The same effects were present in the child data.

Some theories exist to explain why this reverse position effect occurred. One explanation was that the simultaneous presentation of picture and auditory stimuli was a distraction to the initial sound. However, if this were true, the data would reflect results from only the last two phonemes of each word. According to the LRH, the two-segment condition should then elicit more errors than the final-segment condition because there is only one feature difference in the two-segment condition, but two feature differences in the final-segment condition. The SH is supported, based on these results, in that there is a greater sensitivity to the two-segment condition than the final-segment condition. These findings are consistent with previous studies showing children are more sensitive to foils of a target word differing by two segments than one segment, despite the same number of different features (Gerken et al., 1995). Another theory for this effect is that if the last sounds of the word override the first sounds, the subject’s working memory will only remember the last sounds. If the final sounds are correct, the subject may indicate the whole word as correct, even with an incorrect word-initial sound.

There was a considerable amount of variability within the child data. Both the typical and speech-disordered children had a greater number of errors in the final-segment condition than the two-segment condition. These data support the SH in that the children were more sensitive to differences in two segments than in one segment with a two-feature difference. While all three groups of children showed significant errors in the initial-segment condition, the speech-disordered children had a greater number of errors in the final-segment condition. However, when viewing individual data from the three SSD children, child 303 carried the group's final-segment condition outcome with 18.75% errors, while the other two subjects only had 5.88% errors and 6.25% errors. These percentage errors were the same or lower than percentage errors of the typically-developing children. The overall error rate was higher in the children with speech disorders, while the pattern of results across conditions was similar. This suggests these children had similar phonological representations to their typically-developing peers (i.e. segmental representations). These findings also support the use of a mispronunciation detection task to investigate phonological representations in children with speech sound disorders (Sutherland & Gillon, 2005).

Another noteworthy finding in the child data was the high percentage of errors on correct words in both typically-developing groups. There are a few theories as to why this occurred. For example, the sound files consisted of highly enunciated words to avoid speech inconsistencies in the speaker. However, for developing children, these words may have sounded unfamiliar in their enunciated form and been considered incorrect. Another theory may be inattention. If the child was not focused on the task it may have caused inconsistencies in the data. A third theory is that the children anticipated incorrect word forms due to the greater proportion of incorrect pronunciations overall. Additional data is needed to support any of these theories.

It is also relevant to note that data from the eight-year-old group of typically-developing children closely matched the adult data. These results are consistent with the growth in child development. The transition in data from the typical five-year-old group to the eight-year-olds to adults displays a developmental growth in segmental processing. Segments appear to be used in eight year olds more frequently, but the effect is also seen in the five year olds. While this study focused on whether or not segments are present in children's speech processing, the question that comes next is at what age do segments appear in a child's processing. It may be beneficial for future research to study children at a younger age.

Bowey and Hirakis (2006) also found evidence against the position effect but in favor of an effect in clarity of acoustic-phonetic information. There was greater sensitivity to stressed syllables with clearer acoustic-phonetic information than in unstressed syllables. For the purposes of avoiding unrelated errors in multisyllabic words, this thesis only used one-syllable words in its dataset. The outcome of fewer mispronunciation detections on the initial-segment condition did not support the study by Bowey and Hirakis (2006). Had the outcomes been consistent with this study, the data would have shown greater sensitivity to the onset sounds. Bowey and Hirakis (2006) had discussed the idea that sounds in initial position are usually more reliable in retrieving lexical cues because these cues are more consistent and clear. However, the data from this thesis suggest that either such a clarity difference was not present in the stimuli or that word-final sounds may be more critical than word-initial sounds.

In conclusion, the findings from the present thesis do not support the LRH but instead support the SH. This conclusion is based on the finding that children were more sensitive to two-feature mispronunciations when the two mismatching features were distributed across two segments than when they were located on a single segment. As a final point, this thesis provides further support for using the mispronunciation detection task in investigating phonological representations in children with speech sound disorders.

REFERENCES

American Speech-Language-Hearing Association (ASHA). (1997). Guidelines for Audiologic Screening [Guidelines]. Retrieved 10/05/2008 from www.asha.org/policy.

Bowey, J.A. and Hirakis, E. (2006). Testing the protracted lexical restructuring hypothesis: The effects of position and acoustic-phonetic clarity on sensitivity to mispronunciations in children and adults. *Journal of Experimental Child Psychology, 95,* 1-17.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test (4th Edition)*. AGS/Pearson.

Gerken L., Murphy, W. D., & Aslin, R. N. (1995). Three- and four-year-olds' perceptual confusions for spoken words. *Perception & Psychophysics, 57*, 475-486.

Goldman, R., & Fristoe, M. (2000). *The Goldman–Fristoe Test of Articulation (2nd Edition)*. Austin, TX: Pro-Ed.

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language, 29*, 524-545.

Ohala, D. K. (1999). The influence of sonority on children's cluster reductions. *Journal of Communication Disorders, 32*, 397-422.

Reynolds, C. R., & Kamphaus, R. W. (2003). *Reynolds Intellectual Assessment Scales (RIAS)*. Lutz, FL: Psychological Assessment Resources, Inc.

Roelofs, A. (1999). Phonological segments and features as planning units in speech production. *Language and Cognitive Processes, 14*, 173-200.

Seiger-Gardner, L., & Schwartz, R. G. (2008). Lexical access in children with and without specific language impairment: A cross-modal picture-word interference study. *International Journal of Language and Communication Disorders, 43* (5), 528-551.

Schiller, N. O. (2006). Phonology in the production of words. In K. Brown (Ed.), Encyclopedia of language &linguistics (pp. 545-553). Amsterdam et al.: Elsevier.

Storkel, H. L. & Hoover, J. R. (2010). An on-line calculator to compute phonotactic probability and neighborhood density based on child corpora of spoken American English. Behavior Research Methods, 42(2), 497-506.

Sutherland, D., & Gillon, G. T. (2005). Assessment of phonological representations in children with speech impairment. *Language, Speech, and Hearing Services in Schools, 36*, 294-307.

Vitevitch, M.S. & Luce, P.A. (2004) A web-based interface to calculate phonotactic probability for words and nonwords in English. Behavior Research Methods, Instruments, and Computers, 36, 481-487.

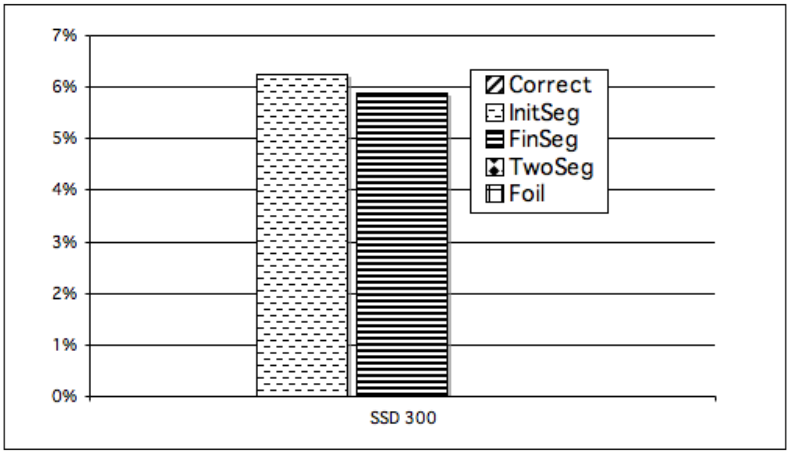
Walley, A. C. (1987). Young children's detection of word-initial and –final mispronunciations in constrained and unconstrained contexts. *Cognitive Development, 2*, 145-167.

Williams, K. (1997). *Expressive Vocabulary Test*. Circle Pines, MN: American Guidance Services.

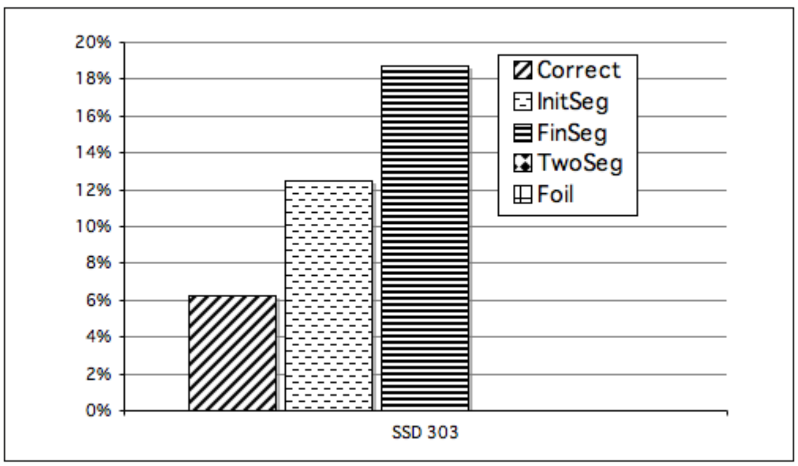
APPENDIX

Individual results of MPD tasks for SSD children

SSD Participant 300



SSD Participant 303



SSD Participant 304