March 2021

Experienced and Inexperienced Listeners' Perception of Childrens' /l/ Productions and their Acoustic Correlates

Emily A. Coniglio

Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses

Part of the Phonetics and Phonology Commons

Recommended Citation


https://digitalcommons.lsu.edu/gradschool_theses/5267

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
EXPERIENCED AND INEXPERIENCED LISTENERS’ PERCEPTION OF CHILDREN’S /l/ PRODUCTIONS AND THEIR ACOUSTIC CORRELATES

A Thesis

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Master of Arts in

The Department of Communication Sciences and Disorders

by

Emily Ann Coniglio
B.A., Louisiana State University, 2019
May 2021
ACKNOWLEDGEMENTS

First, I would like to thank my thesis director and mentor, Dr. Hyunju Chung, for her guidance throughout this project. Her expansive knowledge of child speech sound development, acoustics, and linguistics has fostered a desire to become a better researcher and clinician. Her support, feedback, and help through this project and many others has inspired me to come up with new and innovative ways to analyze speech. Without Dr. Chung, I would not have the expansive knowledge of speech sound acquisition required to be a Speech-Language Pathologist. Once again, thank you Dr. Chung for the opportunity to work with and learn from you.

Additionally, I would like to thank Dr. Janna Oetting and Dr. Geoffrey Coalson for serving as members on my committee. My project would not be what it is without their support and input.

Finally, I would like to thank my family and friends for their unrelenting love and support. To my parents, Anthony and Rachel, thank you for your support of my education. Without both of you, I would not be the scholar and person I am today. To my roommate, Laken, thank you for staying up late, bringing snacks, and being a support throughout undergrad and grad school. To Phillip, thank you for your never-ending IT support and help with statistical analysis. To Mattie, thank you for helping me relax and have fun throughout this stressful time. Without all of you, none of this would have been possible and I am forever grateful.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS...........................................................................................................ii

LIST OF TABLES..................................................................................................................iv

LIST OF FIGURES................................................................................................................v

ABSTRACT............................................................................................................................vi

1. INTRODUCTION.............................................................................................................1

2. LITERATURE REVIEW...................................................................................................2
   Articulatory and acoustic characteristics of /l/.................................................................2
   Linguistic and dialectical differences of /l/.................................................................4
   /l/ acquisition and characteristics of child speech...................................................5
   Speech perception........................................................................................................7
   Research questions......................................................................................................8

3. METHODS.......................................................................................................................9
   Participants...................................................................................................................9
   Stimuli.......................................................................................................................13
   Procedures.................................................................................................................15

4. RESULTS......................................................................................................................18
   Perceptual patterns by listener group........................................................................18
   Acoustic Correlates of Ratings................................................................................21

5. DISCUSSION...............................................................................................................26
   General discussion.....................................................................................................26
   Clinical implications..................................................................................................27
   Limitations................................................................................................................28
   Future directions.......................................................................................................29

APPENDIX A. Experiment Feedback Questionnaire.......................................................30

APPENDIX B. Responses to Feedback Questionnaire ...................................................31

REFERENCES..................................................................................................................33

VITA..................................................................................................................................36
LIST OF TABLES

1. Listeners grouped by listener group, participant ID, exposure to children, last articulation test given, and years of experience.................................................................10

2. Child speakers organized by speech group, age, and sex..................................................13

3. Stimuli and number of stimuli per context in parenthesis.............................................15
LIST OF FIGURES

1. Example of Qualtrics experiment VAS presented to listeners..........................15
2. Sample /l/ trajectory of the word lip.................................................................17
3. Histograms of the frequency of the average rating..........................................18
4. Density plot of average ratings of tokens by listener group..............................19
5. Scatter plot of standard deviation by average rating.........................................20
6. SD by listener group..........................................................................................21
7. F2 and rating by listener group........................................................................22
8. Scatter plot of F3 and rating by listener group................................................23
9. F1 and F2 patterns separated by vowel context................................................24
10. F2 TE and TR by ratings and listener groups..................................................25
Abstract

The phoneme /l/ is one of the highly misarticulated sounds for young children. Referrals for articulation are often based on a listener’s perception of the problem. The aim of the current study was to examine three listener groups’ perception of word-initial /l/ produced by young children to understand if level of experience with child speech impacts listeners’ perception on /l/. The three groups were separated based on their years of experience: speech-language pathologists with at least 10 years of experience (SLP group), graduate students in speech-language pathology (GS group), and naive listeners with no clinical phonetics experience (NL group). Specifically, the differences in perception were examined in relation to the productions’ acoustic correlates. Listeners judged productions of children’s word-initial /l/ using a Visual Analog Scale (VAS). The results showed that the mean ratings and standard deviation did not differ based on listener group, and that F2 was the perceptual cue that SLP and graduate student groups, but not NL group, used to determine the ratings.
CHAPTER 1. Introduction

Speech disorders are the most common communication disorder in children with 48.1% of 3- to 10-year old children, and 24.4% of 11- to 17-year old children presenting with speech problems only (Black, Vahratian, & Hoffman, 2015). These demographics demonstrate how common errors in speech production are in children. Of these errors, liquid sounds (lateral (/l/) and rhotic (/ɹ/)) present an additional challenge for children with SSD due to the articulatory configuration of the articulators, that requires both an anterior and posterior constriction (Lin & Demuth, 2015; Smit, Hand, Freilinger, Bernthal, & Bird, 1990). The lateral and rhotic sounds are a part of the later developing sounds within English phonology (Smit et al., 1990). Furthermore, children produce intermediate sounds that do not fit into a phonemic category that is acceptable to adult speakers (Schellinger, Edwards, & Munson, 2010). In addition to intermediate productions, /l/ is of increased difficulty perceptually due to the articulatory and acoustic properties of the phoneme being intermediate between consonants and vowels (McGovern & Strange, 1977).

The goal of this study was to examine listeners' perception of /l/ produced by young children. Specifically, to understand if the experience level of listeners with child speech impacts their perception of children’s word-initial /l/, perception of three groups of listeners were examined. This includes a) naive listeners with no experience with child speech (NL group), b) graduate level clinicians with moderate degree of experience with child speech (GS group), and c) licensed clinicians with more than 10 years of experience with child speech (SLP group). It is important to understand if differences between listener groups exist in order to better identify children who are in need of therapeutic services or to better understand the development of /l/ in each child.
CHAPTER 2. Literature Review

Articulatory and acoustic characteristics of /l/

In speech production, there are two distinct phonemic categories that speech sounds are classified as: consonants and vowels. Consonants are produced with a constriction made along the vocal tract, and vowels are produced without the constriction in an open vocal tract. Within the category of consonants exists a subcategory called semivowels, which includes the phonemes /r/, /w/, /l/, and /j/ that share acoustic and articulatory characteristics (Espy-Wilson, 1992; Chaney, 1988). Although these sounds are considered consonants because of the constriction made along the vocal tract and ability to act as syllable margins, these sounds are intermediates between consonants and vowels (Espy-Wilson, 1992). Semivowels are produced like vowels with continuous flow of air, but have an oral constriction without frication noise (Espy-Wilson, 1992). The focus of this paper is one type of the semivowel sounds, lateral /l/, produced by young children.

The phoneme /l/ can be articulated with different constrictions. The first distinct way the /l/ phoneme can be produced is called light /l/, which is articulated by tongue tip making alveolar contact (Giles & Moll, 1975). The second distinct way is called dark /l/ and is articulated with a secondary gesture where the tongue root is retracted along with raising of the posterior tongue body resulting in a pharyngeal constriction (Narayanan, Alwan, & Haker, 1997). The terms light and dark are often referred to in the literature as /l/ coloring. /l/ coloring also exists on a continuum that is influenced by surrounding vowel context (Sproat & Fujimura, 1993; Recasens, 2012; Simonet, 2015). /l/ articulation is also heavily dependent upon word position (Espy-Wilson, 1992). Some believe that these factors contribute to overall /l/ lightness and darkness.
(Sproat & Fujimura, 1993; Giles & Moll, 1975). Furthermore, it has been found that tongue shape and configuration is highly variable both within a single speaker and between multiple speakers for both light and dark /l/ (Narayanan et al., 1997). It is also important to note that the light/dark classifications of /l/ do not result in phonemic differences, but rather are representative of allophonic variation (Sproat & Fujimura, 1993). These differences in articulation also result in acoustic differences that can be observed on a spectrogram.

The important acoustic parameters for analyzing the /l/ phoneme are the first (F1) and second formant frequencies (F2), and intensity. F1 values are associated with the vertical height of the tongue, and F2 values are associated with the anterior-posterior advancement of the tongue. In terms of /l/ acoustics, light /l/ exhibits greater F2-F1 distance than dark /l/ (Sproat & Fujimura, 1993). F2 values are lowered by moving the tongue back in the oral cavity, which results in the smaller F2-F1 distance for dark /l/ (Roussel & Oxley, 2010). F2 values for adults\(^1\) range from 750 Hz for dark /l/ to 1800 Hz for light /l/, also resulting in the differences in F2-F1 distance for the /l/ allophones (Roussel & Oxley, 2010). In Midwestern American English, F2 values for adult males range from 760 Hz to 1272 Hz, and 325-500 Hz for F1 depending on vowel context and syllable position (Sproat & Fujimura, 1993; Recasens, 2012). F2 values are the distinguishing feature between /l/ and other semivowels (Epsy-Wilson, 1992; Dalston,1975). Additionally, /l/ can have lower F3 and F4 values dependent on surrounding vowel contexts. The /l/ phoneme also has lower intensity compared to the vowel, which can be observed both on the spectrogram and waveform (Espy-Wilson, 1992).

Furthermore, /l/ has relatively high sonority, or loudness, compared to nasals and obstruents (Johnson & Britain, 2007). These values can differ based on location in the word.

\(^1\) Speaker gender not specified in Roussel and Oxley (2010).
Word-initial /l/, also called onset or prevocalic, occurs before the vowel at the beginning of the word (e.g., lamp, light, look). Contrarily, post-vocalic /l/, also called coda or word-final, occurs after the vowel at the end of the word (e.g., ball, pool, tall). Word-initial /l/ also demonstrates spectral discontinuities that result in abrupt increases in F1 and F2 values (Espy-Wilson, 1992). Since post-vocalic /l/ demonstrates a lesser degree of alveolar contact, this results in a steady formant pattern (Espy-Wilson, 1992).

**Linguistic and dialectal differences of /l/*

The /l/ phoneme can be found in a number of languages throughout the world. Within these languages, several dialects can exist. Among these languages and dialects, /l/ productions and acoustics exist along a continuum in terms of articulation. One of the dialectical dependent features of /l/ production is lightness and darkness. Languages and dialects can fall into three main categories: a) light word-initial and final /l/, b) dark word-initial and final /l/, and c) light word-initial and dark word-final /l/ (Recasens, 2012). Examples of languages that have light /l/ are French, Spanish, and Italian. Contrarily, examples of languages and dialects that have dark /l/ are Leeds British English, Portuguese, and Russian (Recasens, 2012). American English /l/ coloring, on the other hand, is dependent on syllable position, dialect, and region, as well as by speaker’s race and age (Sproat & Fujimura, 1976; Oxley, Roussel, & Buckingham, 2007; Van Hofwegen, 2011). For example, in a study of /l/ in African American English, Van Hofwegen (2011) showed an interaction between age and race on /l/ coloring. It was found that younger African American speakers in the Princeville, North Carolina area showed darker /l/s than their older counterparts. This suggests convergence of /l/ coloring with the dark /l/ that European American speakers of the area use. The third category for /l/ coloring, light word-initial and dark word-final, depends on syllable position and is often found in dialects within a language. For
example, Midwestern American English has a light word-initial /l/ and a dark word-final /l/ (Sproat & Fujimura, 1993), whereas Australian English is considered to have a dark /l/ in all positions (Wells, 1982; Newton, 1996). These examples demonstrate how /l/ can vary within English based on dialect spoken in that region.

/l/ acquisition and characteristics of child speech

The age of acquisition of the /l/ phoneme varies in the literature. Word-initial /l/ has been reported to be acquired as early as age 2;0 (Dyson, 1988) and as late as 9;0 (Smit et al., 1990) in American English-speaking children. Smit (1993) found that word-initial liquids are often deleted in young (ages 2;0-3;0) children’s speech and are not mastered until 6;0-9;0 with an average of 96% accuracy. For post-vocalic /l/, according to Smit et al. (1990), it is acquired between 6;0 and 7;0 with 96-98% accuracy, whereas Templin (1957) reported acquisition at the age of 6;0. Smit (1993) also reported that post-vocalic /l/ is often substituted with schwa, and is not mastered until 8;0-9;0 with an average of 97% accuracy. Lin and Demuth (2015) found that for Australian English-speaking children, they acquired word-initial /l/ before age 3;0 and post-vocalic /l/ after the age of 7;11. It was also noted that many of the children’s post-vocalic /l/ productions were perceived as being vocalized and transcribed as /w/ or a low back vowel (Lin & Demuth, 2015). In a systematic review, Crowe and McLeod (2018) reported that /l/ was acquired between 4;0 and 4;11 using the 90% criterion both in the United States and Globally for English speaking children.

SSD is of particular importance due to the prevalence in childhood speech disorders. Since SSD is one of the most common childhood communication disorders (Black et al., 2015), it is important to understand developmental speech milestones, and correctly identify SSD in children. Because of the prevalence and risk factors of SSD, it is important for SLPs to have
accurate perception of child speech. Currently, there are inconsistencies within the literature about /l/ acquisition. In addition, possible dialectal influence on /l/ productions make it even more challenging to understand true acquisition patterns of /l/.

As for the error patterns, /l/ is often substituted with other semivowel sounds, specifically glides, /w/ and /j/, which is referred to as liquid gliding (Hodson & Paden, 1981). In addition to these phonemic substitutions, children often produce intermediate productions between /l/ and /w/ or /j/ (Klein, Grigos, McAllister Byun, & Davidson, 2012). These intermediate productions can be difficult to judge perceptually, causing difficulty in assessment and intervention (Bernstein, 2015). It should also be noted that the age and type of substitution is diagnostically important. According to the American Speech-Language Hearing Association (ASHA), gliding of liquids should be eliminated between the ages of 6;0 and 7;0 (Peña-Brooks & Hedge, 2015; Shipley & McAfee, 2016). Presence of gliding beyond the appropriate age and substitutions that are uncommon in typical development (e.g., alveolar fricatives for /l/) are indicators that SSD is present (Peña-Brooks & Hedge, 2015).

Child speech typically falls into two distinct categories related to speech sound production: typically developing speech (TDS) and speech sound disorders (SSD). Children with SSD are believed to have different underlying phonemic representations than children with TDS and adults (Chaney, 1988). Children with SSD make errors that are not developmentally appropriate resulting in incorrect productions. Incorrect productions that affect the individual speech sounds include omissions (e.g., lamp as amp), substitutions (e.g., lamp as wamp), additions (e.g., lamp as glamp), and distortions (e.g., dentalization and lateralization of phonemes) (Bernthal, Bankson, & Flipsen, 2013).
Speech perception

Speech perception is the ability to transform an acoustic signal into a chemical one that holds a linguistic representation and is meaningful (Holt & Lotto, 2010). This ability is paramount to SLPs, because of the decision making that revolves around diagnosis and treatment of SSD. Previous studies reported that speech perception is impacted by the degree of experience of the listener. In a study focused on /s-θ/, /t-k/, and /d-g/ contrasts in child speech, Munson, Johnson, and Edwards (2012) found that experienced listeners (i.e. SLPs) demonstrated higher intra-rater reliability than their inexperienced counterparts. Intra-rater reliability was measured using clicks on the visual analog scale (VAS) closer to the ends of the /t-k/ continuum, and ratings were more closely related to acoustic parameters of “correct” vs “incorrect” productions than the inexperienced group. Additionally, this study found that the experienced listeners’ ratings were more closely related to the acoustic parameters than the inexperienced listeners. In a similar study, Klein et al., (2012) examined inexperienced (i.e. graduate students in speech-language pathology) listeners’ perception of children’s /r/ production and found that inexperienced listeners showed lower sensitivity and specificity in accurately assessing incorrect productions. Additionally, Harel, Russo Hitchcock, Szeredi, Ortiz, and McAllister Byun (2016) examined reliability and validity of non-experienced listeners through Amazon Mechanical Turk using VAS. This study found that clicks were highly correlated to F2 and F3 values, and that reliability was highly dependent on the individual participants. This study also found that “correct” productions had higher accuracy than “incorrect” productions. Again, the ratings of “incorrect” and “intermediate” productions by inexperienced listeners were variable.

Although studies on speech sound perception exist, there has not been a study that has focused on perception of /l/ sound, especially listeners’ perception of /l/ sounds in child speech.
Listeners’ perceptions of /l/ are of particular interest due to children’s difficulty in the acquisition of /l/ and persistence of errors during /l/ production by young children. The current study focused on children’s word-initial /l/ productions in real words. These productions were judged by three listener groups with different levels of experience with child speech: a) naive listeners with minimum to no experience with child speech or clinical phonetics (NL), b) graduate students in speech-language pathology (GS), and c) speech-language pathologists with 10 years of experience (SLP).

**Research questions**

The purpose of the current study is to understand the ratings of each listener group, and if years of experience with clinical phonetics affect perception of /l/ sounds produced by young children. Additionally, the acoustic correlated of the productions were also examined. The specific questions are as follows:

a. Do listeners’ experience with clinical phonetics affect their perception of children’s word-initial /l/ productions?

b. How do each listener groups’ ratings of the productions correlate with the acoustic parameters of the token?

It was hypothesized that a) the SLP listener group would demonstrate more agreement within the productions by showing a smaller standard deviation compared to the other groups, and b) that the SLP’s ratings would be more closely correlated to the acoustic parameters of the productions.
CHAPTER 3. Methods

All the recruitment and data collection and analysis procedures were approved by the Institutional Review Board of Louisiana State University.

Participants

Participants included a total of 35 listeners. They included 12 NL, 10 GS, and 13 SLPs. All listeners were recruited via word of mouth, flyers, and social media posts. All GS were graduate students obtaining a master’s degree in speech language pathology. All listeners were from the Southern United States except two GS (GS03: Oregon and GS08: Indiana) and one SLP (SLP12: California). All listeners had normal hearing except GS10 who indicated having a bilateral sensorineural hearing loss, who was included due to not showing any deviation from the rest of the GS group. All listeners had normal speech and language development except SLP11, GS05, GS06, and GS10 who had articulation therapy as children. Listeners ranged in age from 18-69 with a mean age of 32. The mean years of experience for the SLP group was 16 years. Table I lists participant ID, degree of exposure to children, and time of latest articulation administered for the GS and SLP group.
<table>
<thead>
<tr>
<th>Listener Group</th>
<th>Participant ID</th>
<th>Exposure to Children</th>
<th>Last Articulation Test</th>
<th>Years of Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>NL</td>
<td>NL01</td>
<td>Infrequently</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NL02</td>
<td>Never</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NL03</td>
<td>Infrequently</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NL04</td>
<td>Infrequently</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NL05</td>
<td>Infrequently</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NL06</td>
<td>All of the time</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NL07</td>
<td>Frequently</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NL08</td>
<td>Infrequently</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NL09</td>
<td>Never</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NL10</td>
<td>Infrequently</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NL11</td>
<td>Infrequently</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NL12</td>
<td>Never</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL (12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>GS01</td>
<td>Infrequently</td>
<td>A year before the experiment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GS02</td>
<td>About half of the time</td>
<td>6 months before the experiment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GS03</td>
<td>Frequently</td>
<td>2 months before the experiment</td>
<td></td>
</tr>
</tbody>
</table>

(table cont’d.)
<table>
<thead>
<tr>
<th>Listener Group</th>
<th>Participant ID</th>
<th>Exposure to Children</th>
<th>Last Articulation Test</th>
<th>Years of Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS04</td>
<td>Frequently</td>
<td>A week before the experiment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS05</td>
<td>All of the time</td>
<td>5 months before the experiment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS06</td>
<td>Frequently</td>
<td>11 months before the experiment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS07</td>
<td>Frequently</td>
<td>A month before the experiment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS08</td>
<td>Frequently</td>
<td>A month before the experiment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS09</td>
<td>Infrequently</td>
<td>A month before the experiment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS10</td>
<td>Frequently</td>
<td>6 months before the experiment</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL (10)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLP01</td>
<td>All of the time</td>
<td>Month of experiment</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>SLP02</td>
<td>About half of the time</td>
<td>2 years before the experiment</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>SLP03</td>
<td>All of the time</td>
<td>12 months before the experiment</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>SLP04</td>
<td>Frequently</td>
<td>2 years before the experiment</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>SLP05</td>
<td>Infrequently</td>
<td>5 years before the experiment</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>SLP06</td>
<td>All of the time</td>
<td>Week of the experiment</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>SLP07</td>
<td>All of the time</td>
<td>3 weeks before the experiment</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>SLP08</td>
<td>Frequently</td>
<td>1 week before the experiment</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>SLP09</td>
<td>Frequently</td>
<td>1 year before the experiment</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>

(table cont’d.)
<table>
<thead>
<tr>
<th>Listener Group</th>
<th>Participant ID</th>
<th>Exposure to Children</th>
<th>Last Articulation Test</th>
<th>Years of Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLP10</td>
<td></td>
<td>Frequently</td>
<td>6 months before the experiment</td>
<td>25</td>
</tr>
<tr>
<td>SLP11</td>
<td></td>
<td>All of the time</td>
<td>Day of the experiment</td>
<td>14</td>
</tr>
<tr>
<td>SLP12</td>
<td></td>
<td>All of the time</td>
<td>Month of the experiment</td>
<td>15</td>
</tr>
<tr>
<td>SLP13</td>
<td></td>
<td>All of the time</td>
<td>Day before the experiment</td>
<td>12</td>
</tr>
</tbody>
</table>

TOTAL (35)
Stimuli

Stimuli included word-initial /l/ in monosyllabic words produced by children between the ages 2;11 and 6;5, who were either TDS or SSD. Children were grouped as either TDS or SSD based upon parent report and Goldman-Fristoe Test of Articulation Third Edition test scores (Goldman & Fristoe, 2015. These children were from a larger study on child vowel sound (Chung, 2020; Chung & Weismer, 2021). Table II summarizes child speakers by speech group and age. There were a total of 128 productions (14 children x between 2 and 13 productions x 1 repetition per production = 128), including productions that were transcribed as correct /l/ and those that were transcribed as incorrect /l/. Only productions that were free from background noise and not distorted, were included in the stimuli. For all children, some tokens were missing due to non-compliance during testing or distortion of the token. See Table III for a summary of target productions by vowel context. For more detailed recruitment activities, data collection, and analysis procedures, please refer to Chung (2020). The stimuli were downsampled prior to being inserted into the experiment using Praat (version 6.0.21; Boersma & Weenick 2016)’s intensity normalization feature.

Table II. Child speakers organized by speech group, age, and sex.

<table>
<thead>
<tr>
<th>Speech group</th>
<th>Participant ID</th>
<th>Age</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>TDS</td>
<td>CL28</td>
<td>2;3</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>CL14</td>
<td>2;11</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>CL21</td>
<td>3;8</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>CL25</td>
<td>3;8</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>CL32</td>
<td>4;7</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>CL24</td>
<td>5;2</td>
<td>F</td>
</tr>
<tr>
<td>TOTAL (7)</td>
<td>CS34</td>
<td>5;9</td>
<td>F</td>
</tr>
</tbody>
</table>

(table cont’d.)
<table>
<thead>
<tr>
<th>Speech group</th>
<th>Participant ID</th>
<th>Age</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSD</td>
<td>CS36</td>
<td>3;0</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>CS35</td>
<td>4;6</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>CS37</td>
<td>4;7</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>CS40</td>
<td>4;10</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>CS39</td>
<td>5;2</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>CS38</td>
<td>5;7</td>
<td>F</td>
</tr>
<tr>
<td>TOTAL (7)</td>
<td>CS24</td>
<td>6;5</td>
<td>M</td>
</tr>
<tr>
<td>TOTAL (14)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table III. Stimuli and number of stimuli per context in parenthesis.

<table>
<thead>
<tr>
<th>Vowel Context</th>
<th>High Front</th>
<th>Low Front</th>
<th>High Back</th>
<th>Low Back</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lip (12)</td>
<td>Laugh (13)</td>
<td>Loop (12)</td>
<td>Long (12)</td>
</tr>
<tr>
<td></td>
<td>Lick (13)</td>
<td>Lamp (5)</td>
<td>Lose (12)</td>
<td>Lock (10)</td>
</tr>
<tr>
<td></td>
<td>Lake (9)</td>
<td>Lamb (10)</td>
<td>Look (11)</td>
<td>Log (9)</td>
</tr>
</tbody>
</table>

The child productions were first randomized and placed within the experiment in two equal blocks, with half of the listeners received block A then block B and the other half of listeners received block B then block A. Listeners were asked to rate the child productions on a scale from -1 to 1, where -1 represented /j/, 0 represented /l/, and 1 represented /w/. Listeners were able to click anywhere along the scale. If the listener felt that the production did not fall along the scale, they were able to select “I heard something else” and write in their response. The target word for each stimulus was provided. An example of the experiment’s VAS can be found in Figure1.
Figure 1. Example of Qualtrics experiment VAS presented to listeners.

**Procedures**

The experiment was performed using LSU Qualtrics (Version September 2020 Qualtrics, Provo, UT). This was done to accommodate the COVID-19 pandemic since research activities could not be completed in person due to LSU COVID policies and CDC guidelines. Listeners were able to complete the experiment from anywhere, as it has been previously found that ratings did not differ significantly when listeners participated in uncontrolled environments versus research laboratories (Munson, 2013). Completing the experiment in real-world environments also simulated actual speech sound production. Listeners were first asked to consent to the experiment and then completed a short questionnaire about their number of years of experience with child speech, and speech, language, and hearing history (Appendix A). After completing the questionnaire, and before the actual experiment, listeners completed three practice questions in order to familiarize themselves with the format of the experiment. Listeners were provided with
instructions in both written and audio formats prior to beginning the experiment. After completing the practice block, the experiment began. After completing the experiment, participants were asked to provide feedback by answering three questions. Feedback questions included “What was the most challenging part of the experiment?,” “Was it difficult to make decisions about the production? If so why?,” and “Any other comments about the experiment?” See Appendix B for responses to the feedback questions.

Acoustic Analysis

Two variables were included for the acoustic analysis. This included the first three formant frequencies at the /l/ constriction interval (/l/ midpoint) and trajectory patterns from the /l/ constriction interval to the following vowel midpoint. All acoustic measures were extracted using Praat (Boersma & Weenick 2016). Figure 2 depicts the boundaries for /l/ trajectories. Formant frequencies for all productions were extracted using a customized script at 10 different measurement time points (5, 10, 20, 35, 50, 65, 80, 90 and 95%). From the depicted /l/ trajectories, F2 transition extent (TE) and transition rate (TR) were calculated. TE of F2 was calculated for all tokens by taking the difference between F2 at the 95 and 5 time points. TR was calculated by taking the difference of F2 at 95 and 5 and dividing by the difference in duration at those points. TE reflects the amount of vocal tract shape change and TR reflects the speed of changes in vocal tract configuration (speed of articulator movements) from the /l/ midpoint (constriction interval) (Chung & Weismer, 2021; Stevens, 2000; Weismer & Berry, 2003).
Figure 2. Sample /l/ trajectory of the word *lip* produced by a 5-year-old child (CL24), where onset is the /l/ midpoint and offset is vowel midpoint.
CHAPTER 4. Results

Perceptual patterns by listener groups

First, the average rating for each token in each listener group was examined. The mean rating was calculated by dividing the sum of all the ratings of each token by the total number of ratings. The majority of the average ratings fell between -0.25 and 0.25. This means that the listeners perceptually judged most of the tokens as /l/ or /l/-like (−1 represents /j/, 0 represents /l/ and 1 represents /w/). See Figure 3 for a histogram of ratings by listener group.

Figure 3. Histograms of the frequency of the average rating of the tokens for each listener group.

Upon visual inspection, there seems to be no difference in the histograms between listener groups. A one-way ANOVA was conducted to examine if ratings differed by the listener group. The results showed that ratings did not differ significantly by the listener groups (see Figure 4).
Figure 4. Density plot of average ratings of tokens by listener group, where red represents NL group, green represents GS, and blue represents the SLP group.

Second, in order to examine differences in rating across the listeners, the standard deviation (SD) of each token’s rating was calculated for each listener group. The results of a one-way ANOVA with SD as a dependent variable and listener group as an independent variable showed that SD values did not differ significantly by the listener groups.

When patterns of the SD values were examined in relation to those of ratings, the result showed that as the ratings approached toward the intermediate rating between /l/ and /w/ (toward 0.5), SD values increased (see Figure 5). This indicates that the tokens that were rated as definitely /l/ or definitely /w/ showed more agreement within listeners than tokens rated as intermediate, between two phonemes. This pattern was similar across the three listener groups.
A correlation analysis was conducted to further examine the relationship between listeners’ rating and SD patterns. This analysis was conducted for tokens a) rated between 0 and 0.5 and b) between 0.5 and 1. For tokens rated between 0 and 0.5, the Kendall’s correlation was used. This was to accommodate the non-parametric nature of the data that did not meet the assumption of normality based on the Shapiro-Wilk normality test. This test was conducted for both rating and SD values. Kendall’s correlation revealed a significant relationship between ratings and SD patterns, $r_\tau = 0.58$, $p < .0001$. This indicates a significant positive relationship between rating and SD, as rating become closer to 0.5, an intermediate production between /l/ and /w/, so did SD.

For tokens rated between 0.5 and 1, Pearson’s correlation was used as the data met the normality assumption based on the Shapiro-Wilk normality test. Pearson’s correlation revealed a
significant negative relationship between ratings and SD patterns, \( r = -0.79, p < .0001 \). This indicates that as the rating approached 1 (closer to /w/) the standard deviation decreased. This pattern is depicted in Figure 6.

![Figure 6. SD by listener group where the first plot is ratings between 0 and 0.5 and the second plot is between 0.5 and 1.0.](image)

**Acoustic Correlates of Ratings**

First, the relationship between F2 and F3 values extracted at the /l/ midpoint (5% time point) and the listeners’ ratings was analyzed. F2 and F3 values differentiate /l/ from /w/, especially F2 differentiating /l/ from /w/ (Epsy-Wilson, 1992). The results of the linear regression showed that there was a significant effect of F2 values on rating for the SLP group, \( R^2 = 0.06197, F(1, 81) = .0001749, p = .013 \), and the GS group, \( R^2 = 0.07779, F(1, 81) = .0001985, p = .0056 \), but not the NL group \( R^2 = 0.02834 F(1, 76) = -.0001216, p = .076 \). See Figure 7 for a scatter plot of F2 by rating separated by listener group.
The same analysis was conducted for F3 values. The results of the linear regression showed a significant effect for the NL group, but not the SLP and GS groups, SLP group, $R^2 = .03722$, $F(1,81) = -0.0001435$, $p = .013$, GS group, $R^2 = 0.01562$, $F(1,82) = -0.0001055$, $p = .0056$, and NL group $R^2 = .05553$, $F(1,76) = -.000167$, $p = .076$. See Figure 8 for the scatter plot of F3 rating by listener group.
TE and TR values by rating was also calculated to understand if the change in F2 transition into the following vowel provided an additional perceptual cue for listeners. To do this, first, F1 and F2 patterns were analyzed by vowel context. Figure 8 shows patterns of F1 and F2 across the 10 time points by vowel context (high front (/i, ɪ, eɪ/), low front (/æ/), high back (/u, ʊ/), and low back (/ɔ, ɑ/)). As can be observed in Figure 9, front vowels demonstrated more change in F1 and F2 patterns from the /l/ midpoint to the vowel midpoint as compared to those of the back vowels.
A one-way ANOVA was conducted to understand if TE values differed by vowel context. The results showed that TE values differed significantly by vowel context \( [F(3, 3479) = 453.9, p < .0001] \). A post-hoc Tukey HSD analysis showed that /æ/ had the greatest TE followed by /i ɪ/, \( p < .0001 \), then /o u/, and finally /o a/. A one-way ANOVA was conducted to understand if TR differed significantly by vowel context. The results showed that TR also differed significantly by vowel context, \( F(3, 3479) = 703.8, p < .0001 \). A post-hoc Tukey HSD analysis showed that /æ/ has the greatest TE followed by /i ɪ/, then /o u/, and finally /o a/.

Since high vowels /æ/ and /i ɪ/ showed significantly higher F2 TE and TR than the other vowel contexts, these contexts were analyzed to examine if TE and TR have an effect on rating.
As can be observed in Figure 10, the SLP group’s rating increased as F2 TE decreased, while the NL and GS groups remained relatively stable. For F2 TR, the SLP group’s rating increase again while F2 decreased, while the NL and GS groups rating increased as F2 TR increased. The results of simple linear regressions one with TE as a variable and the other with TR as a variable, however, showed no significant effect of TE and TR on ratings for all three groups. See Figure 10 for TE and TR by rating and listener group.

Figure 10. F2 TE and TR by ratings and listener groups.
CHAPTER 5. Discussion

General Discussion

The aim of the current study was to understand if differences in /l/ perception exist between listener groups based on their level of experience with child speech and the acoustic correlates of these productions. The first research question examined if perception differed due to experience with clinical experience. The second research question examined what the acoustic parameters for the tokens and how these related to the ratings. For the first research question, it was hypothesized that the SLP groups would demonstrate more agreement than the other groups. The results showed that there was no significant difference among listener groups for the mean rating or SD. There was, however, a meaningful relationship between rating and SD. As ratings approached /l/ (0) and /w/ (1), the SD of the groups decreased, with a significant positive correlation for tokens rated between 0 and 0.5 and a negative correlation for the tokens rated between 0.5 and 1 for all three listener groups.

These findings for /l/ are different from the findings of previous research which found greater consensus and differing ratings for the SLP groups. Klein et al., (2012) found that the experienced listeners demonstrated more consensus compared to their inexperienced counterparts when rating children’s productions of /l/. The current findings also differ the finding of Munson et al., (2012), which examined listeners’ perception of fricatives and found that the SLP group’s ratings differ from the inexperienced listeners. The current study showed that SD did not significantly differ among the listener groups, therefore, no one listener group demonstrated more agreement than the other. It was observed, however, that as the rating approached toward /l/ or /w/, the agreement about the production increased in all three listener
groups. This indicates that intermediate productions between /l/ and /w/ were more difficult to judge.

The second purpose of the current study was to understand the acoustic correlates of the tokens and how these parameters relate to the ratings. The acoustic cue that had a significant relationship with rating was the midpoint of F2 in the /l/ segment. The SLP and GS groups demonstrated a significant effect of F2 midpoint on rating. Epsy-Wilson (1992) and Dalston (1975) also found that the differentiating acoustic factor of /l/ and /w/ was F2 of the semivowel portion. Given this, the findings of the current study support their studies as the SLP and GS groups used F2 to differentiate between /l/ and /w/. Recall that F2 /l/ midpoint was used to compare if the semivowel portion alone, or in combination with the following vowel affect perception more. Since the results of the TE and TR indicated no effect, the significant results of F2 midpoint demonstrate that the /l/ portion alone was enough to create a perceptual cue for the SLP and GS groups.

**Clinical Implications**

Overall, the findings of the current study have an impact on clinical practice for the field of speech-language pathology. This study found that the SLP and GS groups used the determining perceptual cue in making their decisions. The F2 midpoint of /l/ demonstrated a significant effect on SLP and GS groups, although this did not result in overall different ratings. The dependency of the SLP and GS groups on F2 as the perceptual cue for child speech demonstrates that training in clinical phonetics does increase a listener’s ability to discriminate speech sounds. The GS group’s more stringent ratings also indicate that the more recent exposure in an instructional manner also increases sensitivity to rating child productions. The NL group, however, did not utilize F2 as a perceptual cue.
Since the SLP and GS groups demonstrated more sensitivity to F2 midpoint than the NL group, this indicates a higher level of training in clinical phonetics. This sensitivity to F2 midpoint would be beneficial to the SLP and GS groups when assessing and treating speech sound disorders. Since the GS group had the more stringent ratings compared to the other two groups, it would be beneficial for SLPs to take continuing education courses on assessing and treating speech sound disorders to maintain the skills they develop in graduate school.

**Limitations**

There were limitations to the current study. The first limitation was that participants relied on perception alone, which is not reflective of real-world speech therapy environment. When assessing and treating speech sound disorders, SLPs and graduate students rely on multiple cues to judge a child’s speech accuracy. These cues can include observing the movement of the articulators, secondary behaviors such as facial grimacing, acoustic and auditory cues. In the current study, listeners relied on perception alone. This is one of the main difficulties that listeners listed on the survey feedback, as one listener said, “Without seeing the mouth, it was hard to tell!” In a future study, a video of the child articulating the sound would be beneficial to the SLP and graduate students, since they are trained to rely on more than perception alone.

Additionally, the target word being presented during the children’s /l/ productions also influenced listeners’ perception of the token. This factor was also listed as a barrier by the listeners on the feedback survey. One respondent indicated: “I think that since I saw the target word it made it easier for some words to decide but harder for others since I knew what it is supposed to sound like.”
The directions to the listeners also may have been unclear. Although the directions were listed before every production, listeners noted on the feedback survey that the directions were unclear.

Finally, the current study was limited to listeners’ perceptions of /l/ initial only. A future study should include /l/ in all positions of words to understand if there are differences in /l/ perception based on word position. The use of /l/ productions that were considered “clear” and “undistorted” also may have influenced the listeners’ perceptions. This ultimately led to the majority of the productions being rated as 0, which may have contributed to the null findings between the three groups of listeners.

Future Directions

A future study should include more productions that are incorrect and distorted so that ratings can be more evenly distributed. This would allow for the study to become more balanced and possibly produce group differences in perceptions, which would support the findings of previous studies. In addition to adding a variety of productions, /l/ in all positions of words could be added to create a more balanced experiment. This would allow for an exploration of word position effects on perception. Another addition to the current study could include videos of the children producing the target words so that the listeners can use other cues when deciding a rating. These additions would allow for a more in-depth analysis of listener perception.
Appendix A. Experiment Feedback Questionnaire

How challenging was the survey to complete?

<table>
<thead>
<tr>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Easy</td>
</tr>
<tr>
<td>Easy</td>
</tr>
<tr>
<td>Neither Easy or Difficult</td>
</tr>
<tr>
<td>Difficult</td>
</tr>
<tr>
<td>Very Difficult</td>
</tr>
</tbody>
</table>

Was it difficult to make decisions about the production? If so why?

Any other comments about the survey?
Appendix B. Responses to Feedback Questionnaire

How challenging was the survey to complete?

- Very Easy-5
- Easy-14
- Neither Easy or Difficult-10
- Difficult-2

Was it difficult to make decisions about the production? If so why?

- No-9
- Several of the recordings seem to start during the production which made it more difficult to discern the sound and several were also fast.
- Somewhat, sometimes hearing changes when physically looking at the written word The productions made by some of the children were hard to distinguish. It sounded like it could have been a mix between 2 sounds.
- Some words made sounds other than Y,W
- Sometimes if the audio had background noise
- Sometimes the audio recording almost cut off the first sound.
- at times, because it was just recordings! would be easier to determine to watch their lips.
- Harder when unable to see the productions being made as far as placement goes
- I think that since I saw the target word it made it easier for some words to decided but, harder for others since I knew what it is supposed to sound like.
- a bit. Longer audio would be helpful as well as video of the child's face/mouth during production
- Wasn’t sure if the ending of the word was relevant.
- Instructions were confusing did you just want L sound or vowel and ending ??
- Sometimes. Kids had other error patterns that made it more difficult to tease out. Some kids were doing it right but had it exaggerated tongue placement or residual W lip rounding which would slightly distort production
- Without seeing the mouth, it was hard to tell!
- No visual input
- Y vs l was difficult
- Yes because it is easier to distinguish between a child's /l/, /w/, and /y/ when you can see their mouth make the production. It was hard just listening.
- Some, wasn't sure where to put on sliding scale

Any other comments about the survey?

- Great job!
- Need clearer directions
- It was fun to flex the speech muscle!
References


Templin, M. C. (1957). Certain language skills in children; their development and interrelationships.


VITA

Emily A. Coniglio is a native of LaPlace, Louisiana, where she was born and raised. Emily received her undergraduate degree from Louisiana State University in Communication Sciences and Disorders and Religious Studies with a minor in Linguistics. As a graduate student, she discovered her passion for working with children with speech and language disorders, and began working in the LSU Child Speech Lab under the direction of Hyunju Chung. Emily will receive her graduate degree in May 2021. After completing her master’s degree, Emily plans to find a clinical fellowship in the public school system.