The Impact of Divided Attention on the Ganong Effect as a Function of Age

by

Samantha Leigh Bordman

B.Phil. Candidate in Communication Science and Disorders, University of Pittsburgh, 2018

Submitted to the Graduate Faculty of

The School of Health and Rehabilitation Sciences

of the requirements for the degree of

Bachelor of Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This thesis was presented

by

Samantha Leigh Bordman

It was defended on

November 29, 2018

and approved by

Michael Walsh Dickey, PhD., Associate Professor, Communication Science and Disorders

Dawna Duff, PhD., Assistant Professor, Communication Science and Disorders

Lori Leibold, PhD., Director, Center for Hearing Research, Boys Town National Research Hospital

Thesis Advisor: Sheila Pratt., PhD, Communication Science and Disorders

The Impact of Divided Attention on the Ganong Effect as a Function of Age

Samantha Leigh Bordman, B.Phil.

University of Pittsburgh, 2018

The Ganong effect refers to the impact of lexical knowledge on auditory perception of words when stimuli are acoustically ambiguous. When adult listeners are presented with words that include acoustically ambiguous speech sounds they are more likely to shift their perception to be consistent with a real word rather than a non-word. The current study explored the Ganong effect with adults and children aged 7 to 9 years, and examined whether divided attention differentially impacted the Ganong effect across the groups. It was hypothesized that both children and adults would exhibit the Ganong effect, but that adults would show the effect to a greater extent than the children. The adults and children were presented words and syllables beginning with a velar plosive (/k/-/g/) that varied by voice-onset-times (VOT) from 0 to 80 ms and comprised three acoustic continua: /kɪft/- gift/, /kɪs/-/gɪs/, and /kɪ/-/gɪ/. The adults and children were asked to identify the initial consonant in stimuli during undivided and divided attention conditions. Both children and adults exhibited the Ganong effect in all listening conditions, but there was no additional lexical drift observed in the divided attention condition. The adults and children did not differ significantly on the tasks.

Table of Contents

List of Tables

List of Figures

Preface

I would like to offer my most sincere thanks and appreciation to Dr. Pratt for advising my through the Bachelor of Philosophy process throughout my undergraduate career. I would not have been able to complete this thesis without her mentorship and reassurance. Additionally, I would like to thank my family and friends for their support and help. Finally, I would like to thank Dr. Suzik and Falk Laboratory School for allowing me to recruit and run participants at their facilities.

1.0 Introduction

1.1 Lexical Effects Across Levels of Speech

Listeners often ignore erroneous or conflicting acoustic cues in speech so that they hear utterances that are plausible in their language and culture. For example, nonwords are unlikely in most conversational settings so listeners tend to hear acoustically ambiguous or incomplete word productions as real words. This bias against nonsense combinations exists at many linguistic levels from individual phonemes and words (Ganong, 1980) to sentences and passages (Miller, Heise, & Lichten, 1951; Windmann, 2007). The perceptual mechanisms allow listeners to ignore or modify errors and ambiguity that might occur due to speech production differences, misarticulations, dialect, noise and inappropriate word selection.

The ability to make ambiguous, partial or erroneous cues whole is evidenced by the phonemic restoration effect (Warren & Obusek, 1971). When gaps are inserted into a speech sample and are then filled with clicks or coughs, listeners treat the gaps as if they were not present. That is, they filled the gap and often fail to notice the presence of the gap, even when told in advance that gaps would be present. However, phoneme restoration is less effective when phonemes are used to fill the gap, likely because they introduce competition at the phonological and lexical processing levels. Whether the gaps are filled at the phonetic level, probabilistically, or lexically is of debate, but the benefit of filling the gaps is that listeners can recover information in adverse and impoverished acoustic conditions. Extraneous noises like coughing or a car horn may interfere with sensory encoding of elements of speech signals but the perceptual system replaces the noise with the likely missing information. Because most listeners experience variable listening conditions this process occurs frequently. Other evidence of the auditory perceptual system adjusting to optimize correct perceptions is the Ganong effect.

1.2 **The Ganong Effect**

Ganong (1980) hypothesized that voice-onset-time (VOT) categorical boundaries shift in single-word utterances as a function of word knowledge – that the boundaries shift to favor real words over non-words. He constructed word-pairs that varied in seven VOT steps (15-55 ms) where one end of the continuum resulted in a real word and the other ended in a non-word rhyme. For example, one continuum varied from /kɪs/ to /gɪs/, whereas the other varied from /kɪft/ to /gɪft/. Individual words were presented in random order and the participants were asked to identify the initial consonant of each word as either "G" or "K". Ganong found a significant lexical effect in that his listeners produced larger categories for the real words than the non-words. They showed shifts in the VOT categorical boundaries to favor the real words resulting in identification functions with larger area under the curve for the real word category. Furthermore, Ganong observed that the lexical effect was pronounced along the acoustic continuum where ambiguity was greatest (boundary region) and was less evident at the ends of the continua where the VOTs were prototypical for the perceptual categories. This effect has been replicated in many, but not all studies (Mattys & Wiget, 2011; Norris, McQueen & Cutler, 2003; Kingston, Levy, Rysling & Staub, 2016) and the results and interpretations vary substantively by the stimuli, experimental methods, and theoretical perspective (Davis & Johnsrude, 2007; Kingston et al., 2016; Pitt &

Samuel, 1993). A primary use of the Ganong effect procedure has been to identify the roles of top-down and bottom-up processing, and how and where speech acoustics and language interact.

1.3 **Categorical vs. Interactive Language Processing**

Top-down and bottom-up processing are generally described in a binary fashion, but the process is more complex and nuanced. If top-down processing is used listeners apply their linguistic and world knowledge to influence how speech acoustic cues are comprehended within a word or utterances. Some authors argue that top-down processing involves the reshaping of the mental representation of the sounds (McClelland, Mirman & Holt, 2006) through feedback mechanisms that replace the sound representation itself after completing some processing at the sound category selection level (Pitt & Samuel, 1993). In contrast, bottom-up processing is data-driven and considered serial in its architecture. Bottom-up processing dictates analysis of acoustic characteristics of speech before phoneme selection and accessing more superordinate lexical information. It often is attributed to autonomous or general auditory models of speech perception, which rely on transitional probabilities and recurrent networks to account for the correct selection of phonemes given variable and compromised input (e.g., compensation for coarticulation, perception in noise) (Pitt & McQueen, 1998; Norris, 1993). Norris, McQueen and Cutler (2000) proposed that lexical information influences phoneme decisions but the influence does not feedback to the phoneme activation level and does not erase the initial low-level response to the acoustic signal.

The Ganong effect often is described as strong evidence for top-down, interactive processing, although the procedure has been used by proponents of autonomous models to account for location of the effect in the perceptual hierarchy. Ganong (1980) concluded that the results of his study were consistent with lexical influences preceding phonetic categorization, because categorical boundaries shifted to favoring the real-word end of the continua. He argued that his results would not be possible if acoustic-level processing did not take lexical information into account. Ganong also asserted that finding the greatest effects with the most ambiguous stimuli and not at the endpoints served as additional evidence that acoustic and lexical information are present at the same time during speech and language processing. It also suggests that speech processing is not a simple top-down vs. bottom-up process.

The importance of top-down influences on the sub-lexical processing of speech cues varies greatly and remains a source of debate (Kingston et al., 2016; McClelland, Mirman & Holt, 2006; Pitt & McQueen, 1998). Grosjean (1980) argued that both top-down and bottom-up processing mechanisms are necessary to explain the speed and efficiency at which words are perceived, even in isolation. He wrote that, "although we are in agreement that only the inherent phonetic-acoustic characteristics of the word are the important factors for recognition of words out of context, we do believe that top-down information can help restrict the initial cohort of candidates in number and in kind" (pp 274). By presenting participants with similar words varying in syllable length, he concluded that top-down processing allowed listeners to narrow down the number of possible word options to high-frequency words before considering low-frequency words based on acoustic structure.

1.4 **English Voice Onset Time**

Ganong used voicing information in his study to demonstrate lexical influences on perception. For English speech sounds, voicing information (along with place and manner) is used to characterize specific speech-sound categories. Speech sounds produced without vocal fold vibration are considered voiceless, such as /k/, whereas sounds where the vocal folds do vibrate during some portion of the production, such as /g/, are classified as voiced. The sounds /k/ and /g/ are both plosives, meaning that during production, airflow is cut off due to a constriction in the vocal tract and subsequently released as a burst. For the sounds /g/ and /k/, this constriction occurs between the back of the tongue and the velum. The time between the release and the beginning of vocal fold vibration for the subsequent vowel is referred to as VOT.

Voice-onset-time is highly variable across and within individuals, and is influenced by place of articulation, phonetic context, articulation rate, language, word frequency, speaker differences, and dialect (Allen, Miller & Desteno, 2003; Morris, McCrea, & Herring, 2008; Lisker & Abranson, 1967; Macken & Barton, 1980). In English, voiced plosives are considered partially-voice and generally have VOTs of 20 ms or less (typically 0 - 10 ms), whereas voiceless plosives have VOTs greater than 20 ms (typically greater than 40 ms). Children's VOTs for voiceless plosives increase with age (Lowenstein & Nittrouer, 2008), and VOTs vary substantively within and across speakers in conversational speech, with differences as much as 50 ms (Allen, Miller & Desteno, 2003). As a result, ambiguous VOTs are not limited to laboratory manipulations. but are prevalent in daily conversations.

1.5 **Implications of Synthetic Stimuli vs. Natural Stimuli on Perception**

Most early studies looking at lexical influences on the speech perception, including Ganong's (1980) initial study, used synthesized speech to construct their acoustic continua Synthetic speech is perceptually more challenging than natural speech, although it should be noted that the quality of synthetic speech is improving, as is our reliance on it. Synthetic speech is used in rehabilitation technology developed for people with disabilities, such as high-technology augmentative and alternative communication devices and text to speech readers. It is used in educational and treatment settings and on personal technology devices. Many automated communication systems use synthetic speech. However, synthetic speech tends to cause greater listening effort, and increased latencies and reduced accuracy when compared to natural speech (Francis & Nusbaum, 2009). Synthetic speech tends to have fewer perceptual cues available and therefore, has restricted cue redundancy and a higher risk of misperceptions. The difference in accuracy in single-word recognition between natural and synthetic speech can reach as much as 16% (Rajinder, 2003).

Difficulty in perceiving synthetic speech is age-related. Roring, Franklin, and Charness (2007) found that providing context in word-identification tasks helped older adults when natural speech was used, but the benefits were limited for synthetic speech presentations. In addition, their listeners were only 60% accurate when identifying synthesized words presented in isolation but over 90% accurate with natural speech. Accuracy in identifying sentences also was significantly better with natural than synthetic speech. Because synthesized speech is less complete and more difficult to process than natural speech, it is plausible that lexical influences on perception increase when listening to synthetic speech. Listeners likely need to increase their reliance on linguistic knowledge and contextual factors when listening to synthetic speech.

The primary arguments for using synthetic speech is that it allows for fine control of acoustic parameters and limits the number of conflicting cues present in the stimuli. This parameter control, however, may make the stimuli less relevant to the speech processing mechanisms and not reflect the variability and perceptual challenges present in natural settings.

Children's perception of synthetic speech is a well-documented, presumably in consideration of computerized educational and training platforms, and the implications of augmentative and alternative communication devices. It also has been used as a more difficult stimulus and a means of separating phonetic from linguistic- and cognitive-level processing (Coady, Evans, Mainela-Arnold, & Kluender, 2007; Evans, Viele, Kass, & Tang, 2002). In terms of response latencies, children perform more poorly on comprehension tasks when presented with synthetic speech than when presented with natural speech. Increased latencies are found across childhood, but young children produce substantively longer latencies than older children and adults when listening to synthesized speech (Reynolds & Jefferson, 1999). Sound categories also tend to be more defined and accurate with natural than synthetic speech stimuli, especially for young children and children with language and perceptual deficits (Blomert & Mitterer, 2004; Coady et al, 2007; Evans et al., 2002).

Ganong's assertion that lexical categorization is interactive with phonetic-level processing was challenged because of his use of synthetically-created speech stimuli. Burton, Baum and Blumstein (1989) argued that natural speech has more voicing cues than synthetic speech; therefore, Ganong's results could not be applied to typical listening situations. To support their argument, Burton et al. completed two experiments – one that replicated the original Ganong study except that natural speech stimuli were used and digitally edited to create /duk/-/tuk/ and /dut/-/tut/ VOT continua. The second study used the same stimuli except that the burst and aspiration amplitudes for each token were adjusted to more natural levels. The first experiment showed a lexical effect but the second experiments produced nearly identical results across the two continua, even in the ambiguous VOT boundary regions. Burton et al. (1989) argued that their results refuted Ganong's claim of phonetic-lexical interaction and that higher-level processing only influenced sublexical level processing when acoustic cues are impoverished or conflicting, which is unlikely with natural speech. The situation can arise with synthetic speech because it is acoustically impoverished. In instances where the lexical effect did seem to control phonetic categorization, the reaction times of participants was significantly slowed. The increased reaction times were considered evidence of participants actively considering lexical information prior to making decisions, rather than having lexical knowledge integrated into the phonetic level of processing.

It should be noted, however, that Burton and Blumstein (1995) repeated their study with corrected stimuli and added noise conditions. They failed to replicate the original study but did see lexical effects when background noise was added. Their conclusion was that stimulus quality and integrity, rather than naturalness, influenced the role of the lexicon in phonetic-level perception. Because children are adversely affected by background noise, especially background speech, they may need to rely more on lexical skills than adults, despite immature linguistic skills (Corbin, Bonino, Buss, & Leibold, 2016; Hall, Grose, Buss, & Madhu, 2002; Leibold & Buss, 2013). In children with language or cognitive deficits this pressure on language processing can reveal linguistic weaknesses not evident under optimal listening conditions (Coady et al., 2007; Evans et al., 2002).

1.6 **Implications of Divided Attention on the Ganong Effect**

It is widely accepted that dual-task performance is generally poorer than performance on a single task. The adverse effects of divided attention on task performance often is considered from the view that central processing mechanisms have access to a finite pool of resources for allocation across tasks occurring at any given time (Just et al., 2001; Just & Carpenter, 1992; Navon & Gopher, 1979). This assumption implies that human processing is based on an "economy" where rapid decisions, like word selection, can be negatively impacted when resources are limited due to multiple demands, immaturity or reduced innate capacity. It should be noted, however, that others have argued that the problem is not one of restricted overall availability of resources but restrictions in the process of resource allocation (Pashler, 1984; Wickens, 2008).

Of importance to the current study is that dual-task performance varies with age with children and older adults having more difficulty in dual- and multi-task situations. Verhaeghen et al. (2003) conducted a metanalysis of 33 dual-task studies comparing young and older adults and confirmed the negative effects of dual-task processing and showed that they were additive, meaning that the negative effects associated with divided attention compounded to make a divided attention task more suffer more in terms of latency and accuracy than two undivided attention tasks separately. The effects were evident in both latency and performance accuracy, and more pronounced with older adults relative to latency. The young and older groups suffered similar reductions in accuracy during multi-tasking activities.

There also has been much attention paid to the impact of divided attention on speech recognition. This attention is well-deserved, as listening environments rarely, if ever, are restricted to an individual sound or isolated word. Important listening activities occur in schools and workplaces, where listeners must consider a range of stimuli from visual stimuli, emotional distractions, variations in content difficult and other acoustic quality. Multi-tasking on laptops and other electronic devices during lectures happen across many academic and language learning environments, decreasing memory of content (Courage, Bakhtiar, Fitzpatrick, Kenny, & Brandeau, 2105; Hembrooke & Gay, 2003). Zhang and Samuel (2014) found that participants only engaged in perceptual learning under optimal conditions, and not when participating in adverse conditions including background noise or under dual-task conditions (Zhang & Samuel, 2014). Learning only occurred at the phoneme level, which emphasizes the need to optimize conditions during learning tasks.

Zhang and Samuel (2014) also found that divided attention impacts the perception of phonemes. As such, it is likely that using a divided attention task would strengthen the Ganong effect. Studies have shown that reliance on linguistic knowledge is especially evident in situations where the ability to focus on an acoustic signals has been compromised (Mattys et al., 2013; Mattys & Widget, 2011). For example, Mattys et al. (2013) found that acute anxiety in listeners results in a heightened reliance on previous linguistic knowledge when categorizing ambiguous initial-word consonants. Situations where divided attention tasks occur during speech processing produce a lexical drift (boundary shift) consistent with the Ganong effect, were "listeners tend to ignore important acoustic details in the speech signal and rely too much on the lexical plausibility of its content" (pp 1606). Mattys et al. considered this finding to be consistent with the interactive perspective offered by Ganong (1980).

Lexical drift also has been studied in situations where cognitive load and attention is manipulated with visual tasks detractors. Mattys and Widget (2011) found that adding a visually-based cognitive load significantly altered phoneme boundaries for the /k/-/g/ continuum, and like Ganong (1980), found that perception favored the real-word end of the continuum over the nonword end. However, Mattys and Palmer (2015) subsequently found that divided attention tasks failed to support the interactive accounts of lexical drift. Their study tested the perception of high- and low-frequency words under noise-degrading conditions. Their decision-making task did not create a bias towards high-frequency word selection, instead, participants relied on acoustic information to guide their decisions and picked phonemically similar options. The Mattys and Palmer's results were more consistent with bottom-up speech processing, but cognitive load and lexical influences cannot be disregarded.

1.7 **Developmental Considerations**

Most of the aforementioned research on the impact of various listening conditions on the Ganong effect was based on adult participants. Yet, children commonly listen in unfavorable environments (e.g., preschools and classrooms are full of visual, acoustic, and emotional distractions) and learn language, social skills and academic content in those environments (Crandell & Smaldino, 2000; Jamieson, Kranjc, Yu, & Hodgetts, 2004). Therefore, it is important to understand children's reliance on lexical knowledge when they listen in background noise or conditions of divided attention. Dependence on lexical skills by children in these types of conditions assumes that they have sufficient lexical knowledge and understand the linguistic/phonemic rules to make decisions that overrule phonetic-level information. If the Ganong effect is present in children, they must present with a substantial amount of phonemic and lexical skills.

Although language gains are rapid in early childhood, school-aged children show substantive vocabulary growth, some owing to the written and oral language input in school settings. For example, in grades 1-3 it is estimated that children learn approximately 3,000-5,000 new vocabulary words a year, although the amount is influenced by various factors such as socioeconomic status and maternal education. In third grade, when students are aged 8 – 9 years, their vocabularies increase by approximately 3,000 words a year (Nagy & Anderson, 1984). These increases in the early grades are added evidence that primary school-aged children have considerably smaller vocabularies than adults, and therefore may be less inclined to use an interactive perceptual process than adults.

It also can be assumed that children have heard specific words less often than adults, and possibly with less frequency, so that linguistic factors such as word-frequency are less potent than for adults. So too, words that are high-frequency in adults communication environments might not be high-frequency in child communication environments. This is relevant as some adult studies have shown that the effects of lexical drift increase when the word/nonword pair includes a particularly high-frequency word (Ratcliff et al., 2016).

Apart from lexical knowledge, the presence of the Ganong effect relies on the listener's ability to know what phonemes are allowed to exist in any specific order based on the rules of a given language. As such, children's phoneme awareness (Yopp & Yopp, 2000) and understanding of phonotactic rules may be a factor in lexical influences on speech perception. Phonemic awareness has gained a great deal of attention in education because it corresponds with reading success (Ehri et. al, 2011). For example, Badian (1993) assessed young school-aged children with the *Test of Auditory Analysis Skills* and found that adequate readers scored higher than poor readers. Rosner and Simon (1971) also found significant differences between kindergarteners and third graders on the test. The difference in phonemic awareness between kindergarteners and third graders reflects perceptual maturation and phonemic skills gained in the early school years. To that end, many teacher training programs have responded by encouraging expanded phonemic awareness instruction for pre-literate children and early readers (Brady & Shankweiler, 2013; Lundberg, Olofsson, & Wall, 1980; Metsala & Walley, 2001).

There is a paucity of research regarding lexical drift and the Ganong effect in pediatric populations; however, children appear to have flexible categorical boundaries when perceiving vowels. When presented with a vowel continuum with a native vowel at one end and a nonnative at the other, young children perceived words further toward nonnative vowels on the continuum than older children and adults (Walley & Flege, 1999). This pattern suggests that children may be more accepting of the acoustic properties of speech and less susceptible to supra-phonetic factors and influences. That is, they may rely more on bottom-up than top-down processing. Supporting this argument is that young children exhibit the phonetic context effect as consistently as adults. Although the phonetic context effect does not rely on the same linguistic knowledge as the Ganong effect, these perceptual abilities show that children are attuned to small acoustic differences in speech utterances (Utz, 2009).

2.0 Experimental Questions

Because there is limited study of the Ganong effect in children, especially under divided attention conditions, the following questions were addressed.

1. Do children show the Ganong effect for words containing initial /k/-/g/ VOT continua (/kɪss/-/gɪss/ vs. /kɪft/-/gɪft/)?

2. Do children differ from adults in their presentation of the Ganong effect?

3. Does a simultaneously presented visual task impact the Ganong effect?

4. Does a simultaneously presented visual task impact children differently than adults?

The Ganong effect is expected to be present in children and adults but greater in adults given their greater lexical skills. That is, it is predicted that VOT boundaries for children will not shift as drastically for adults. The VOT boundaries are expected to drift even further away from the word-creating end of the continua for the adults and children in divided attention tasks. This would suggest that visual-auditory divided-attention tasks interfere with speech processing tasks due to limitations on central processing resources. However, if the dual task also interferes with lexical access, children may show a reduced or no boundary shift.

3.0 Methods

3.1 **Participants**

This study was approved by the University of Pittsburgh IRB and all participants signed a consent and/or assent form before beginning study procedures. The participants were divided into two groups: 5 typically developing children aged 7-9 years and 14 young adults aged 19-22 years. The age range of the children was selected to ensure they could attend to both the visual and auditory tasks, as the Ganong effect has not been assessed in children previously. Additionally, children aged 7-9 have at least some level of formal schooling, which would assist them with attending to the attentional requirements of the task. The adult participants were recruited through Pitt+Me, flyers posted throughout the University of Pittsburgh, and announcements in undergraduate courses in the Department of Communication Science and Disorders at the University of Pittsburgh. The children were recruited through Pitt+Me, flyers posted around the University of Pittsburgh. The participants recruited through these measures were compensated with $10 per session. In addition, children recruited through these methods were compensated with stickers or a small toy.

One child was recruited through the Falk Laboratory School. All children aged 7-9 years at the Falk Laboratory School received a letter detailing the procedures of the study and consent forms. A follow-up phone call was made before beginning study activities to answer questions regarding the study and complete the background questionnaire. Verbal consent from both the parent and child was given on this phone call. The child recruited through the Falk Laboratory School completed the study activities in the school facilities. The child was required to provide assent again before the study began. Per school policy, this child was not paid but received stickers and a small toy.

All participants spoke English as their first and primary language and had no significant history or evidence of speech, language, reading, learning, or hearing difficulties.

Table 1. Adult Demographic Information

Participant	Age	Gender
A1	21	Male
A2	21	Female
A3	24	Male
A4	21	Female
A5	22	Male
A6	20	Female
A7	21	Female
A8	19	Female
A9	19	Female
A10	22	Male
A11	21	Female
A12	20	Female
A13	21	Female
A14	22	Male

Table 2. Pediatric Demographic Information

Participant	Age	Gender
C1	7	Male

C2	8	Female
C3	8	Female
C4	7	Male
C5	9	Male

3.2 **Screening Procedures**

All participants completed a background questionnaire, tympanometry, otoscopy, a pure-tone hearing screening, the Words-in-Noise Test (Wilson, Carnell & Cleghorn, 2007), the five-item listening and reaction time versions of the Computerized Revised Token Test (McNeil et al., 2015), the Snellen vision screening, and the Peabody Picture Vocabulary Test (Dunn, 1997). Participants with scores more than three standard deviations from the norm on any portion of the screening measures would have been excluded, but this was not the case for any participants. Participants or their guardian completed a background questionnaire to ensure participants met inclusion criteria. Otoscopy was completed to ensure the participant's tympanic membrane anatomy was typical and that their external ear canals were unobstructed. Tympanometry was completed using a 226 Hz probe-tone to ensure typical middle ear function. The pure-tone screen was completed with insert headphones with 500, 1000, 2000, 4000, and 8000 Hz tones presented at 25 dB HL through a diagnostic audiometer. The Words in Noise Test assessed word recognition in babble with a decreasing signal to noise ratio. The listening version of the Computerized Revised Token Test assessed auditory language processing and working memory, and the reaction time version of the test screened for motor response difficulties. The Snellen vision screening was administered to confirm typical vision with appropriate corrective lenses. The Peabody Picture Vocabulary Test was completed to document age-appropriate receptive vocabulary skills.

The participants were required to successfully complete the training procedures for the experimental procedures. Training consisted of three sections: acoustic only, acoustic and visual at a slow rate, acoustic and visual at the experimental rate. Participants were presented with the instructions “You will now be presented with six sounds. Please press “K” on the keyboard if you hear a word beginning with “K”, and press “G” if you hear a word beginning with “G”. Press any key to continue”. The speech stimuli were recordings of a Pittsburgh native female speaker saying “kiss”, “giss”, “gift”, and “kift” (described below). After each presentation, the computer screen prompted the participant to press the key corresponding to the initial sound in the stimulus by displaying “K or G”. This sequence was repeated six times until all the original stimuli were presented. The screen then displayed, “You will now be presented with sounds and images. While listening to the sound, search the image for a red square (See Figure 1 below). Please press “K” if you hear a word beginning with “K”, and press “G” if you hear a word beginning with “G”. Press “N” if there is no square in the image, and press “S” if there is a square in the image. Press any key to continue.” The participant was then randomly presented with the same unaltered acoustic stimuli paired with a random visual stimulus. There were ten presentations of simultaneous stimuli in this section. Each visual stimulus was presented for 1.5 seconds. After each presentation, the screen displayed “G or K”. Once the participant selected a key, the screen displayed “square or no square”, prompting participants to choose “S” for square or “N” for no square. The next section repeated the same instructions as the previous, but it informed participants that the visual stimuli would be presented for a shorter amount of time. In this section, visual stimuli were presented for 560 ms. There were 10 presentations of simultaneous stimuli in this section. In order to be included in the study, participants needed to meet 80% accuracy cumulatively in the second and third sessions on both identification of initial sound of endpoint stimuli on both ends of the continua and red square.

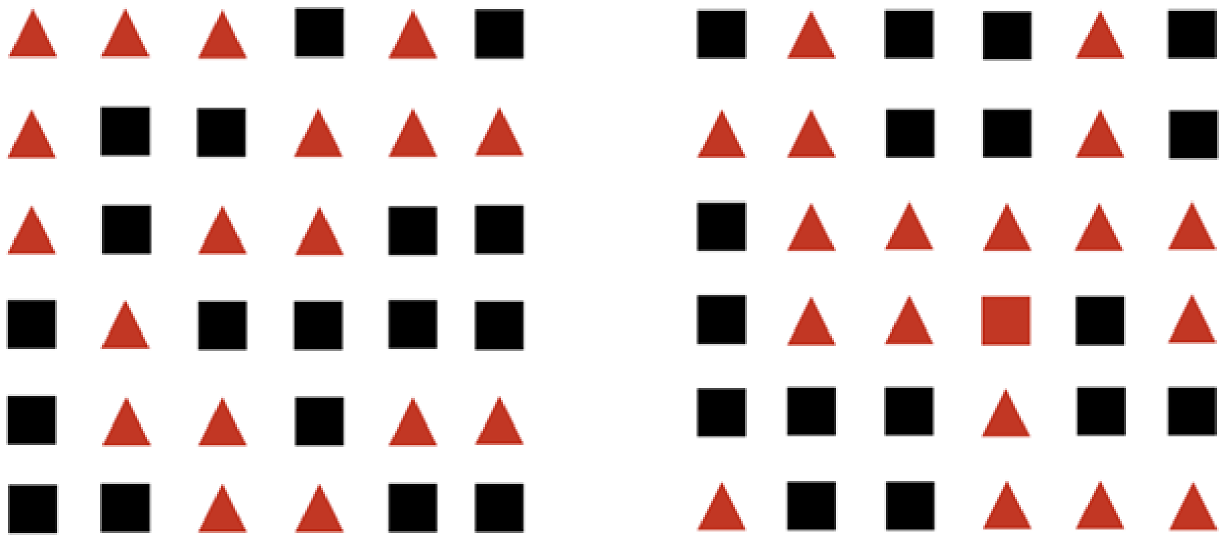

Figure 1. Examples of control and red square visual arrays

Table 3. Adult Screening Measures Information

Participant	WIN Average	PPVT III Percentile Rank	Training Percentage Score
A1	3.07	88	93
A2	-0.4	45	89
A3	4.73	47	81
A4	1.73	91	96
A5	0.4	68	100
A6	-1.2	87	96
A7	-0.6	61	100
A8	-1.73	91	96
A9	0.13	68	93
A10	-0.13	82	98
A11	-1.2	93	96
A12	-0.67	66	96
A13	-0.67	66	96
A14	-0.67	73	98

Table 4. Child Screening Measures Information

Participant	WIN Average	PPVT III Percentile Rank	Training Percentage Score
C1	9.73	37	85
C2	1.47	87	85
C3	2.53	96	83
C4	-0.4	77	93
C5	-0.4	94	87

3.3 **Stimuli**

The acoustic and visual stimuli used in this study were based on those used in Mattys and Widget (2011). The stimuli consisted of 8-step continua from /g/ to /k/ with the endings /ɪft/, /ɪ/, and /ɪs/. Two of the continua formed non-word and real-word endpoints and the other was a consonant-vowel syllable and served as a control stimulus.

Forty-eight natural acoustic stimuli were used. The natural acoustic stimuli were created by manipulating recordings of a 19-year-old female native Pittsburgh, PA speaker producing the utterances “say kift again”, “say gift again”, “say kiss again”, “say giss again”, “say gi again”, and “say ki” again. Using Adobe Audition, utterances were manipulated to have VOTs ranging from 0 to 80 ms at 11.42 ms intervals. The vowel durations were adjusted to account for overall duration and average RMS intensity was matched. This resulted in sixteen acoustically controlled stimuli for each ending (-/ɪs/, - /ɪft/, - /ɪ/).

The visual stimuli for the dual task condition consisted of 96 six-by-six arrays made up of black squares and red triangles. Each array appears as 7-inch x 7-inch on a 13-inch Macbook Air. Using number randomizer software, each shape was assigned a location on each array. In half of the arrays, a red square was randomly assigned a location using the same software and took the place of the black square or red triangle previously in that location. Each visual stimulus was paired with an acoustic stimulus randomly every time the experiment was run.

3.4 **Experimental Procedures**

The experimental procedures were similar as those used for the training task except that the full range of stimuli along each continuum were presented. The study procedures were consistent with those used by Mattys and Widget (2011).

The experiment was run using SuperLab 5 on a 13 inch MacBook Air. The acoustic signals were routed through a 4-channel tabletop amplifier to Radioshack circumaural headphones and calibrated to present the words at 65 dB SPL. The procedures were conducted in a sound booth in Forbes Tower at the University of Pittsburgh and a quiet room at the Falk Laboratory School. The experiment consisted of two blocks (undivided attention, divided attention). The order of these blocks was counterbalanced to prevent order effects. Each block contained 3 presentations of each acoustic stimulus from each ending group. In all, each block contained 144 events. In undivided attention blocks, only acoustic stimuli were presented. In divided attention blocks, acoustic stimuli were randomly paired with the visual array described above (Figure 1), which appeared for 560 ms. It was ensured that each block consisted of 50% visual stimuli with a red square and 50% without a red square. There was a 2-second inter-stimulus interval between each event presentation.

Participants were reminded of the instructions before beginning the experimental session and had the option to review the training. Each block lasted approximately 15 minutes, and participants were instructed to take breaks between blocks. In some cases, participants needed to spread the blocks across several sessions due to time and attention constraints. This was particularly true for the children.

After each screening and experimental session, participants were compensated $10 (or stickers and toys) and both the participant and primary investigator signed a receipt confirming the participant had been compensated.

4.0 Results

Table 5 and the figures below illustrate the results. Percent correct was calculated for each participant for each point along the VOT continua per condition. Then area under the curve and boundary location was calculated for each individual identification function. The boundary location was determined with a log probit analysis. Then a linear mixed model analysis was applied to the area under the curve and probit results with secondary post-hoc comparisons.

Figures 6-15 illustrate the results for the undivided and divided attention tasks for the five child participants in the study. These figures were included due to the high variability in the percent “K” responses of each child, and the low number of child participants, which may have contributed to the variability in the overall child results. Additionally, the individual figures illustrate the differences seen between children, as the participants differed in age, gender, and scores on the screening measures.

Only main effects for VOT continuum were found, with significant lexical effects present for both adults and children in both undivided and divided listening conditions (p<.0001) and was most pronounced for the /kɪs/-/gɪs/ continuum. Area under curve was greater and VOT boundary was less for the /kɪs/-/gɪs/ continuum than the other two continua in both listening conditions (p<.0001). Area under curve and VOT boundary for the /kɪft/-/gɪft/ continuum were similar to that of the /kɪ/-/gɪ/ continuum. The adults and children did not differ significantly, although the small number of children likely reduced the ability to see a difference. Divided attention did not produce added lexical drift for the adults, and although there does appear to be a difference for the children it did not reach significance. The /kɪs/-/gɪs/ continuum did show lexical effects into the /gɪs/ end-point region for the children, which was not the case for the adults. As a result, the children’s perception for that continuum could be considered less categorical than in the undivided listening condition.

Accuracy on the divided attention tasks suggested that the children and adults attended to the visual task (Figures 6 and 7). The adults correctly identified the presence or absence of the red square 87.65% of the time, and children completed the task with an average of 73.61% accuracy. A difference in competency between the two groups was expected but the results suggested that both groups completed the tasks as instructed. Accuracy for adults on the visual task was similar to that found in the Mattys and Widget (2011) study.

Table 5. Area Under Curve and VOT Boundary for all Conditions

Condition	**Area Under Curve**	**VOT Boundary (ms)**
Adult Undivided /kɪs/-/gɪs/	3.10	32.46
Adult Undivided /kɪft/-/gɪft/	2.36	40.69
Adult Undivided /kɪ/-/gɪ/	2.64	38.00
Adult Divided /kɪs/-/gɪs/	3.00	33.08
Adult Divided /kɪft/-/gɪft/	2.54	29.15
Adult Divided /kɪ/-/gɪ/	2.72	37.00
Children Undivided /kɪs/-/gɪs/	3.17	28.20
Children Undivided /kɪft/-/gɪft/	2.30	41.40
Children Undivided /kɪ/-/gɪ/	2.73	36.40
Children Divided /kɪs/-/gɪs/	3.67	25.40
Children Divided /kɪft/-/gɪft/	2.53	37.40
Children Divided /kɪ/-/gɪ/	2.77	35.40

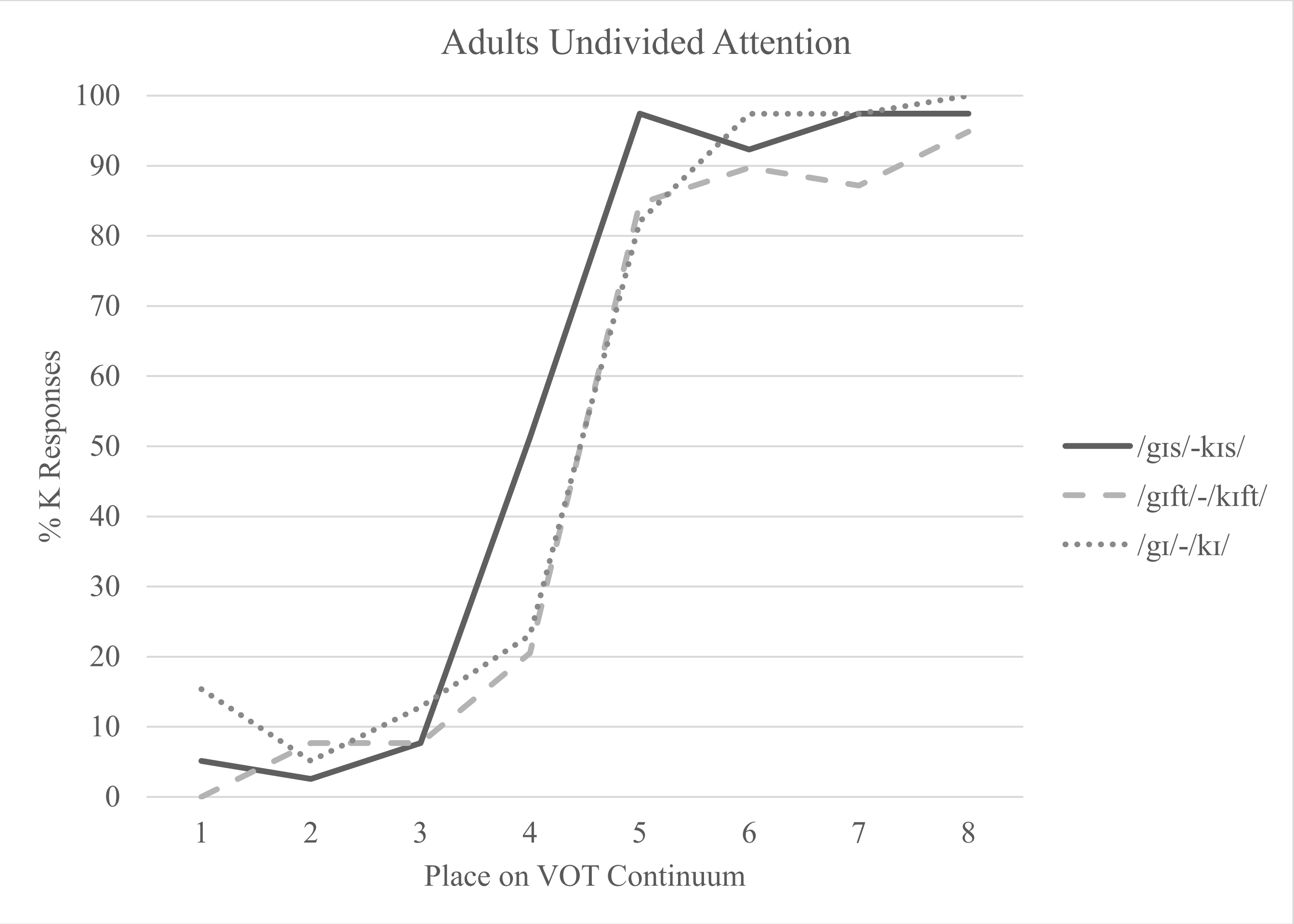

Figure 2. Percent "K" Responses in Adult Participants: Undivided Attention Conditions for all Continua

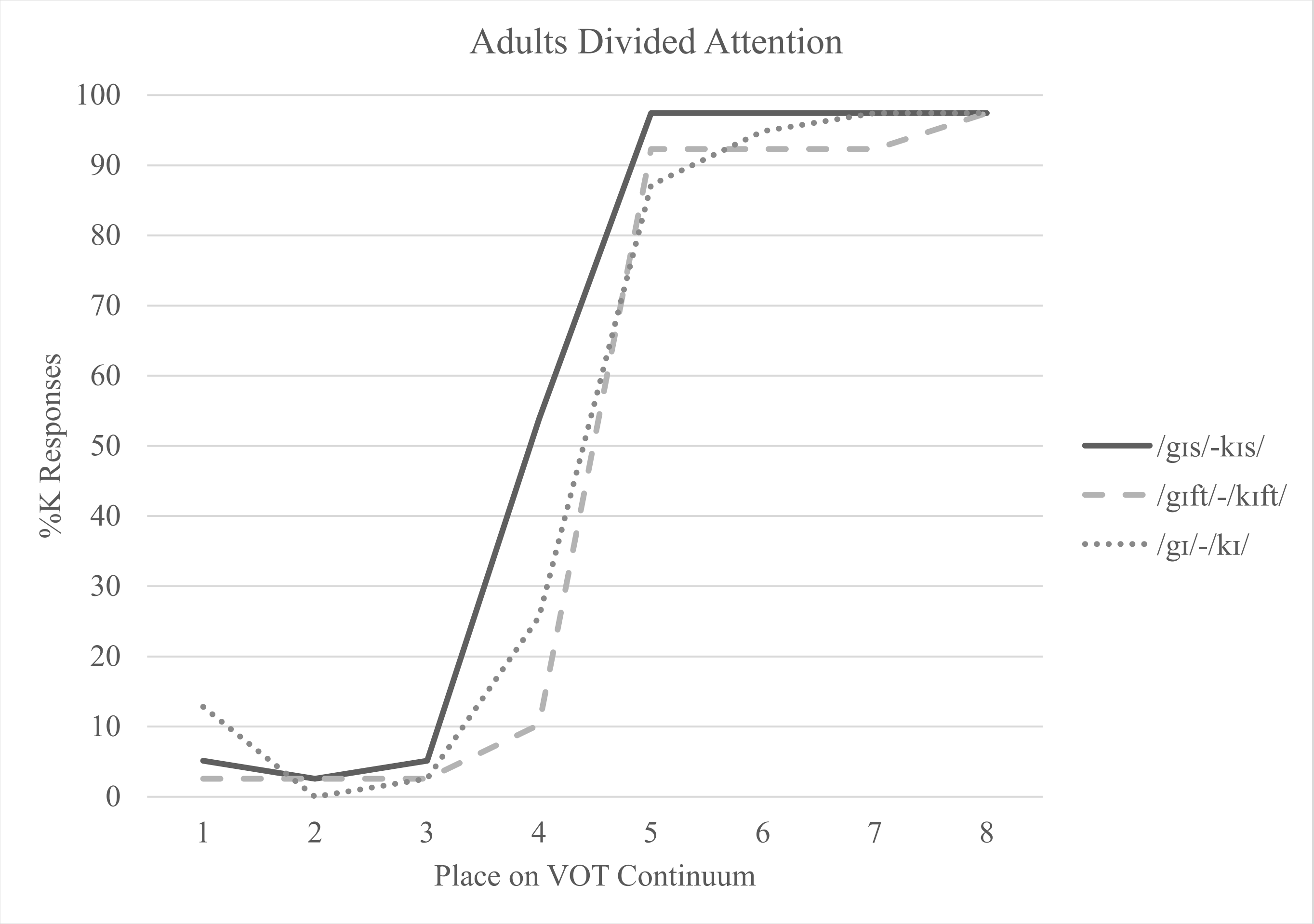

Figure 3. Percent "K" Responses in Adult Participants: Divided Attention Conditions for all Continua

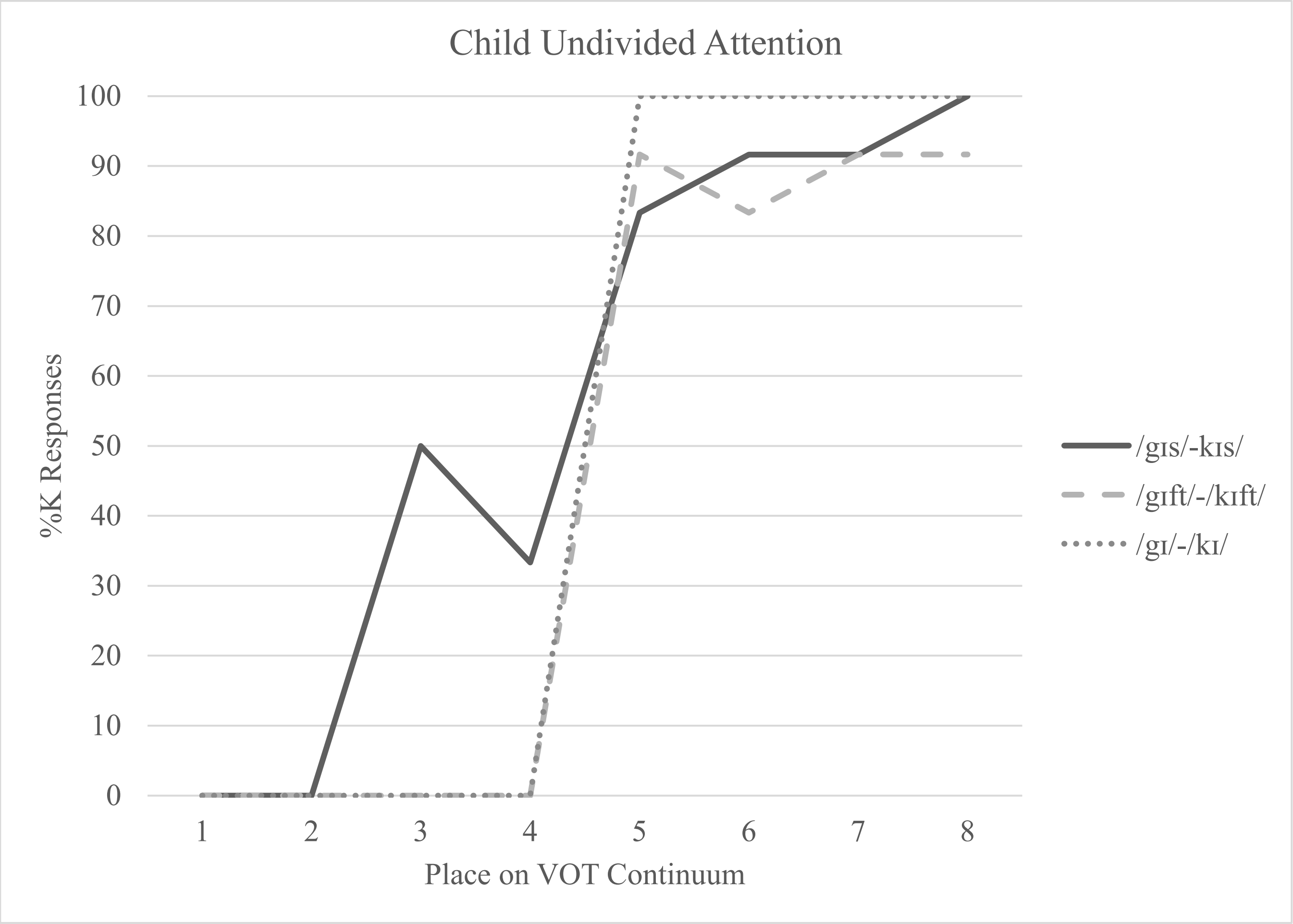

Figure 4. Percent “K” Responses in Child Participants: Undivided Attention Conditions for all Continua

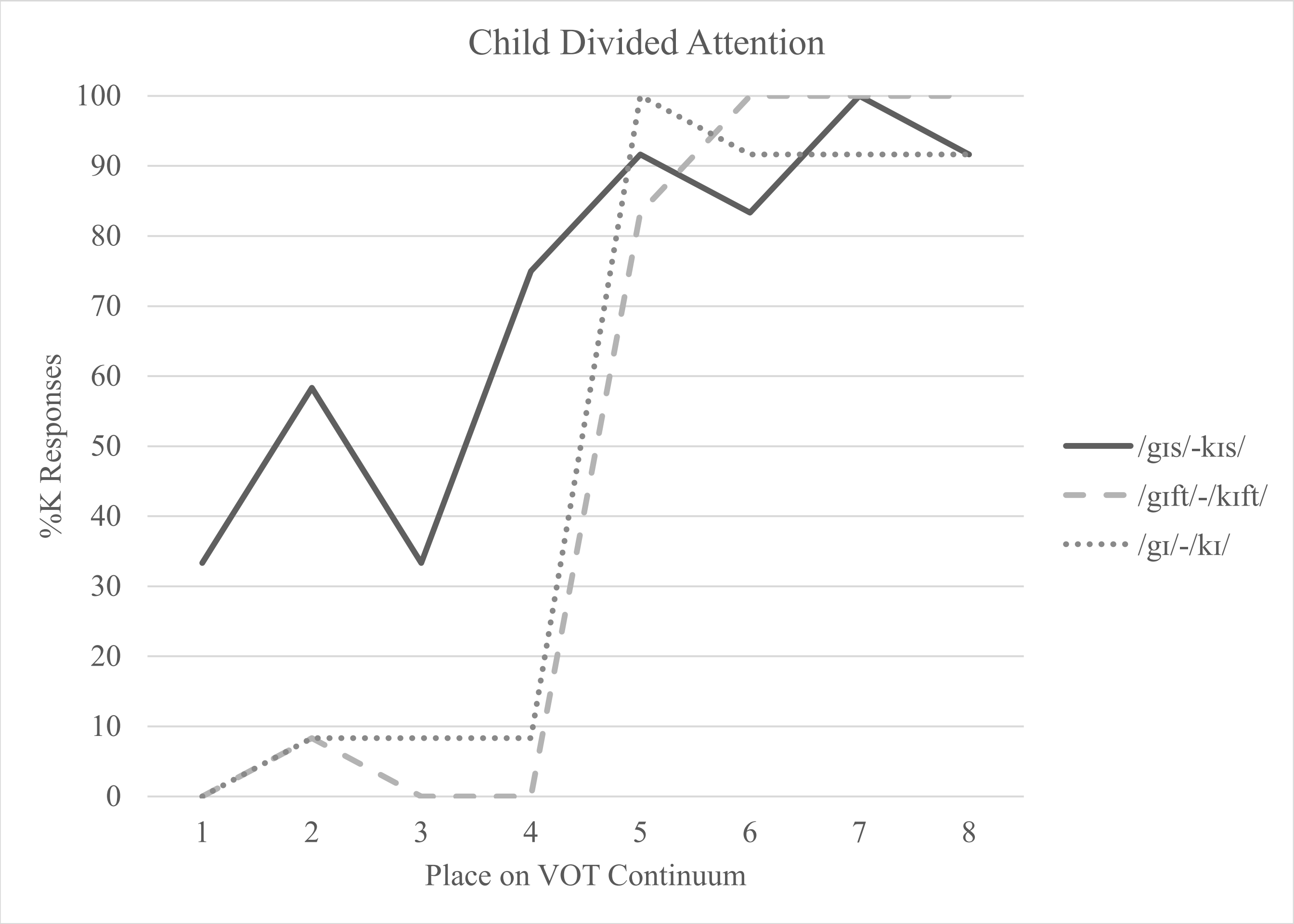

Figure 5. Percent "K" Responses in Child Participants: Divided Attention Conditions for all Continua

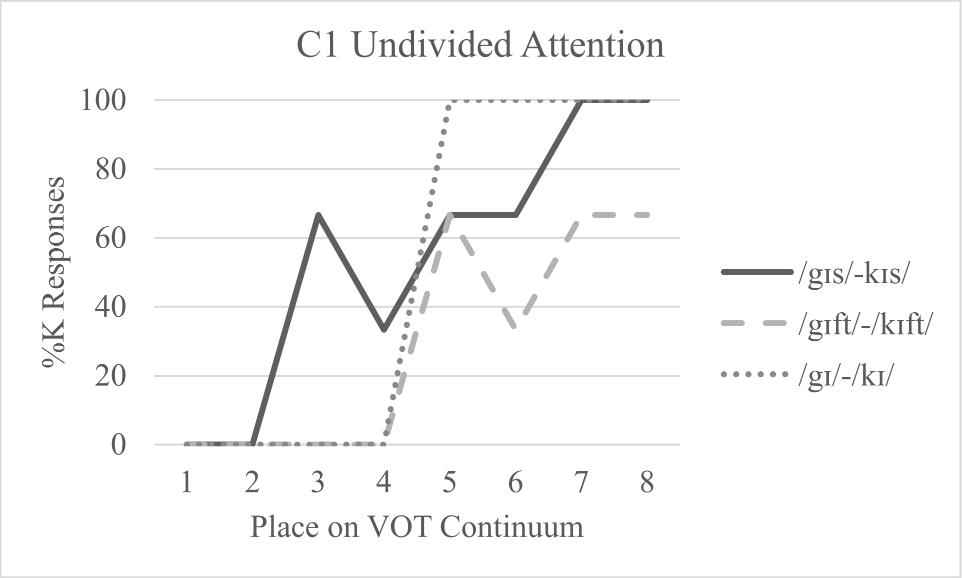

Figure 6. Percent "K" Responses in Participant C1 Under Undivided Attention Conditions for all Continua

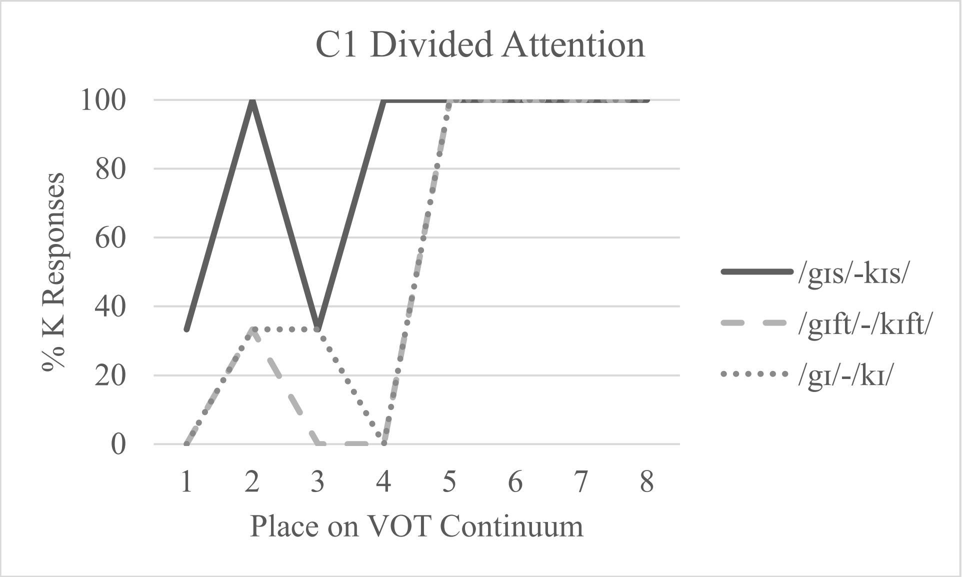

Figure 7. Percent "K" Responses in Participant C1 Under Divided Attention Conditions for all Continua

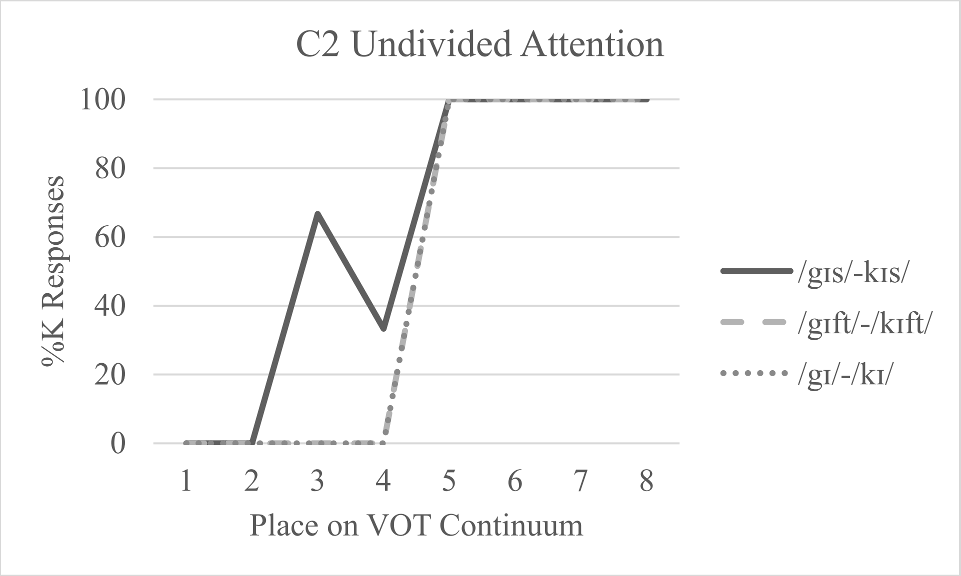

Figure 8. Percent "K" Responses in Participant C2 Under Undivided Attention Conditions for all Continua

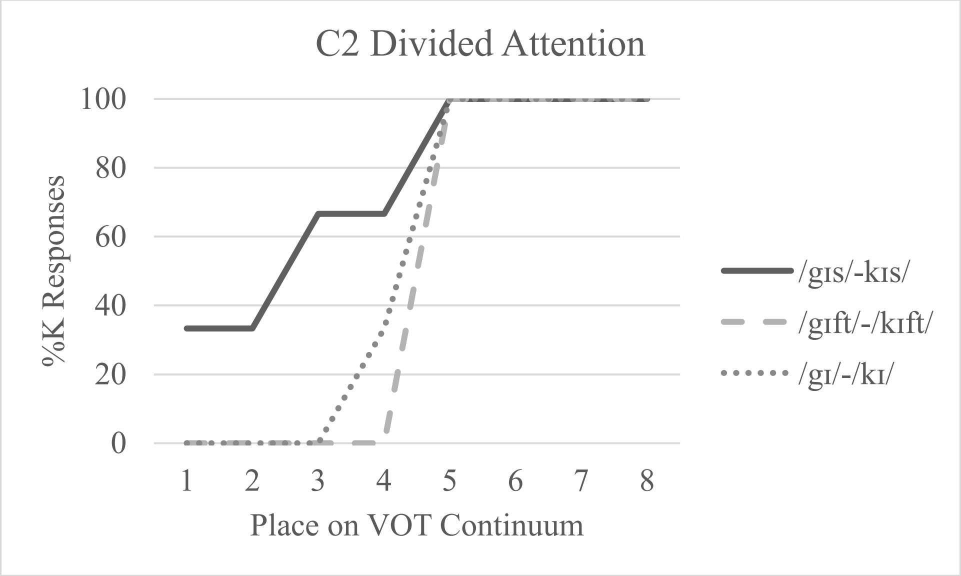

Figure 9. Percent "K" Responses in Participant C2 Under Divided Attention Conditions for all Continua

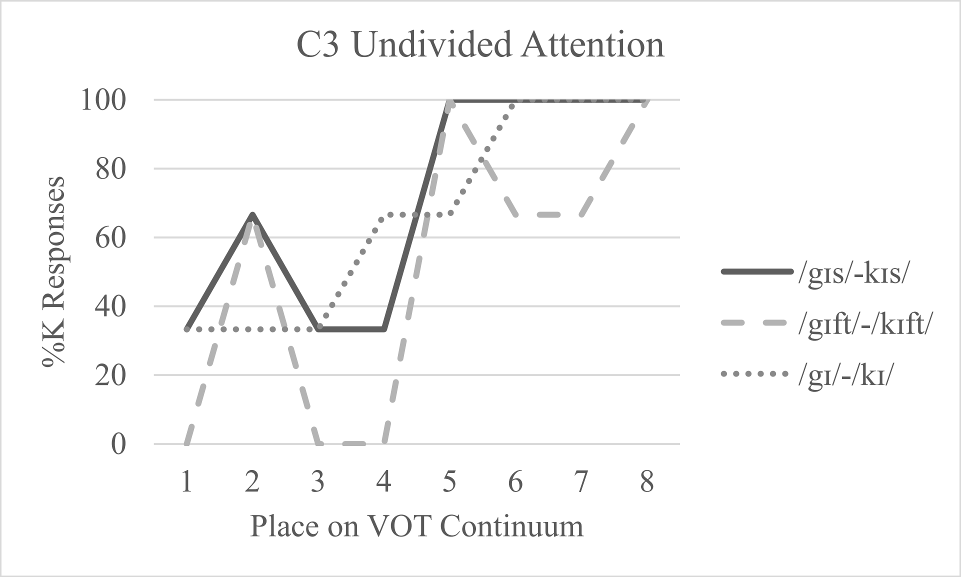

Figure 10. Percent “K” Responses in Participant C3 Under Undivided Attention Conditions for all Continua

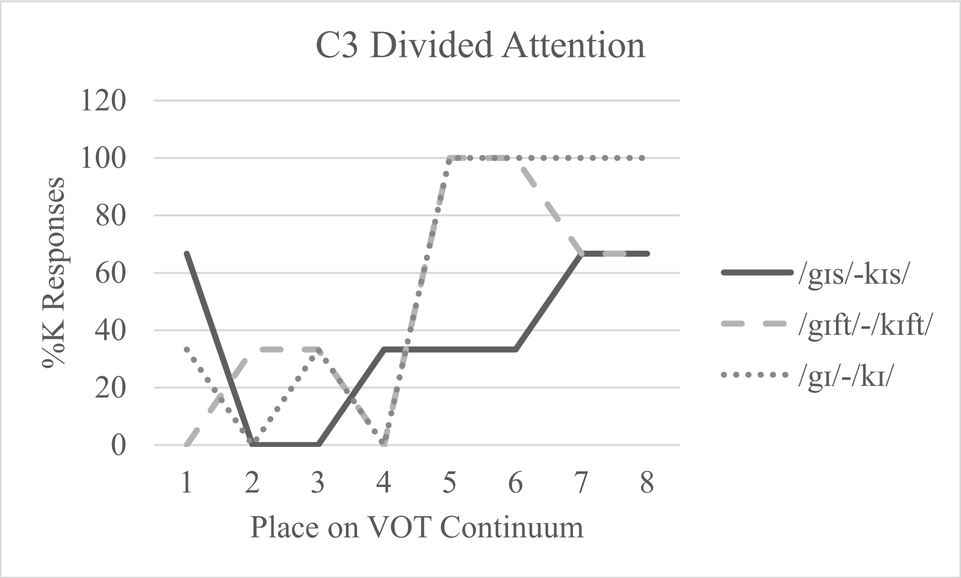

Figure 11. Percent “K” Responses in Participant C3 Under Divided Attention Conditions for all Continua

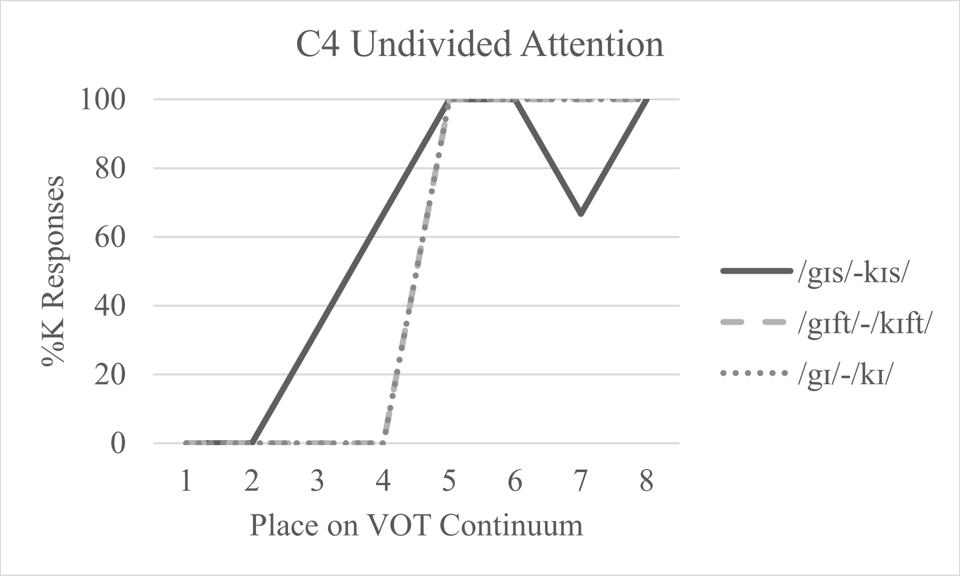

Figure 12. Percent "K" Responses in Participant C4 Under Undivided Attention Conditions for all Continua

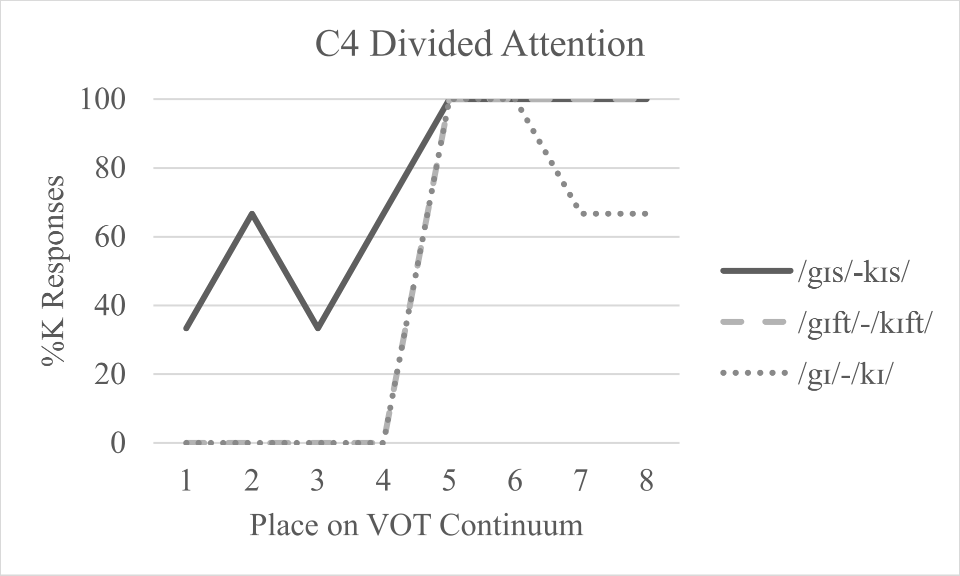

Figure 13. Percent "K" Responses in Participant C4 Under Divided Attention Conditions for all Continua

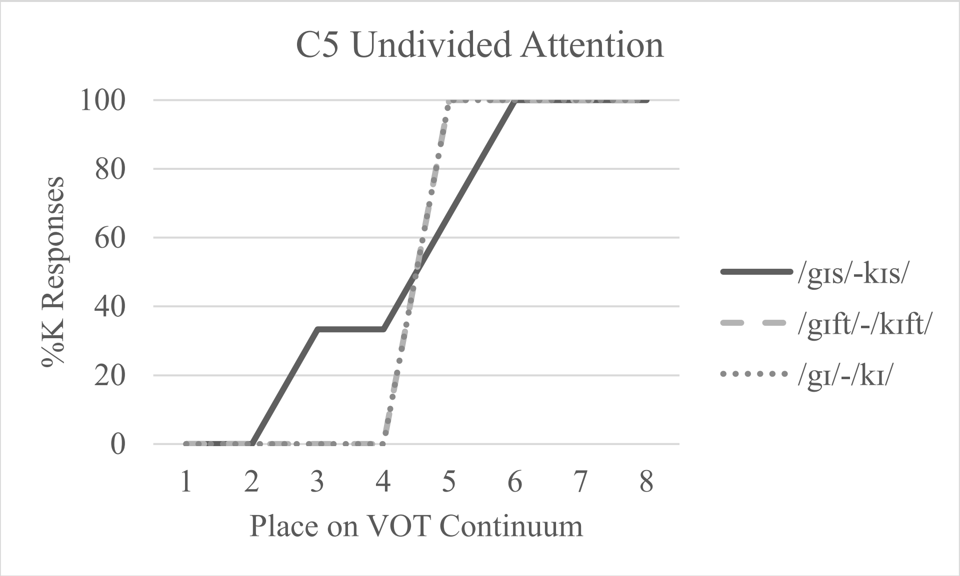

Figure 14. Percent "K" Responses in Participant C5 Under Undivided Attention Conditions for all Continua

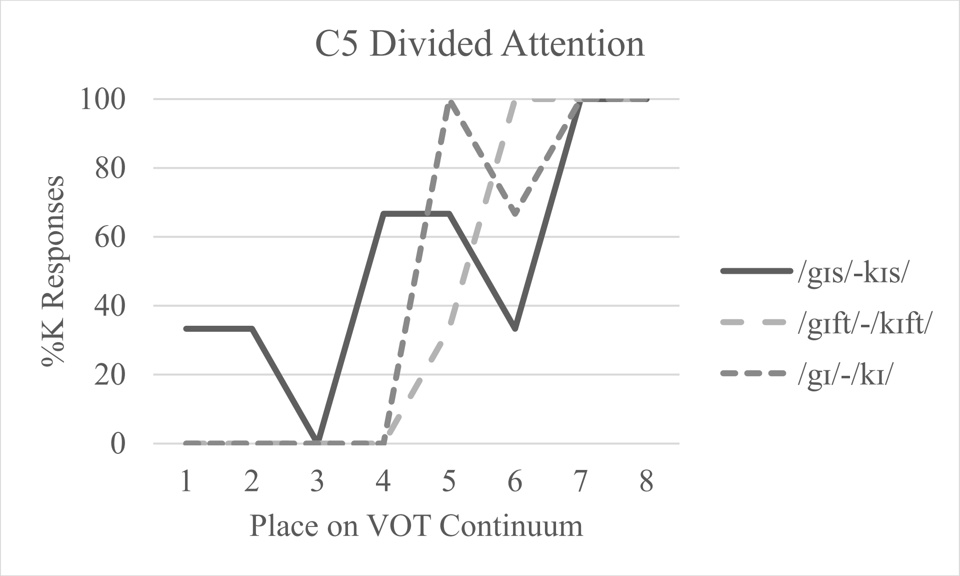

Figure 15. Percent "K" Responses in Participant C5 Under Divided Attention Conditions for all Continua

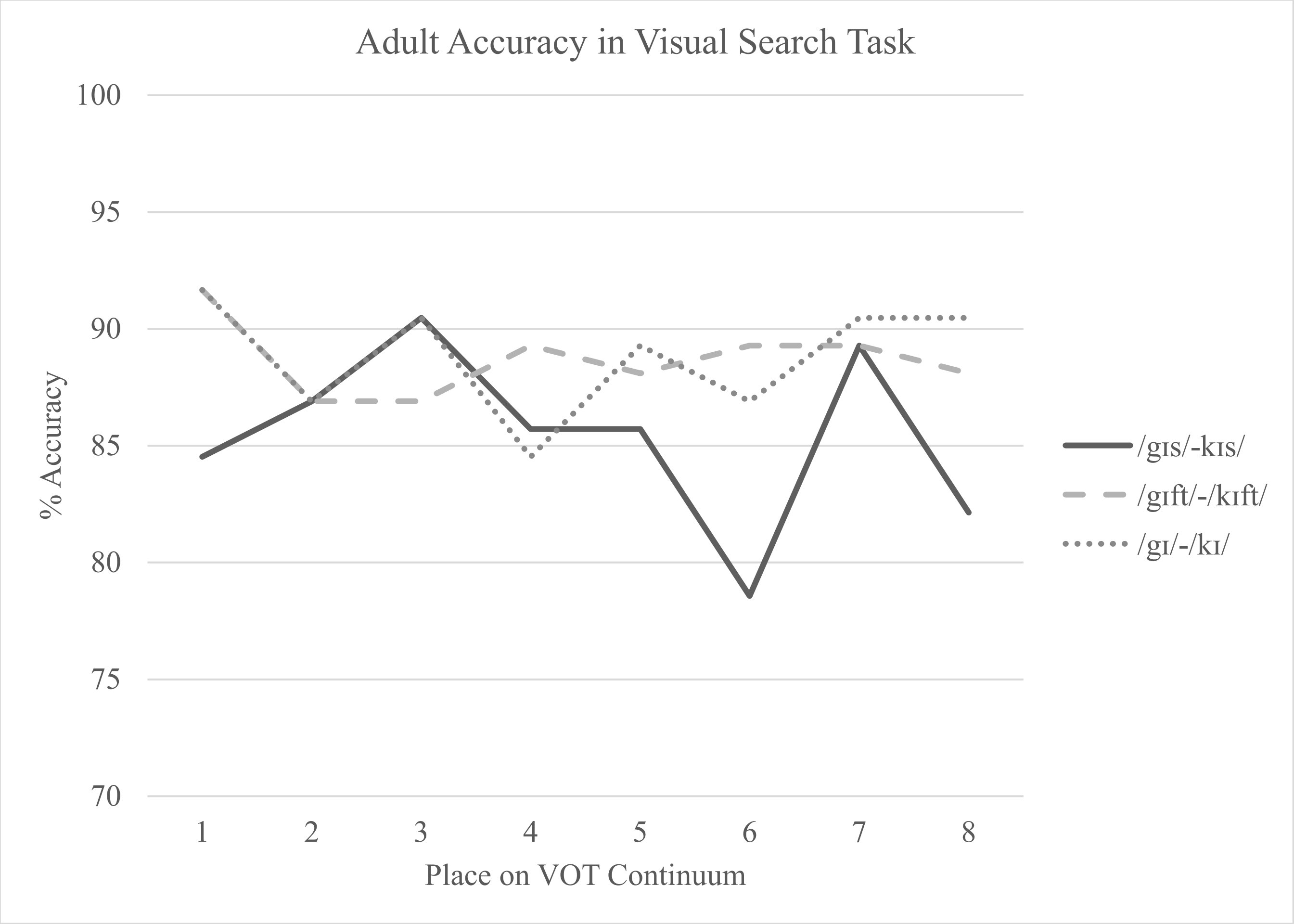

Figure 16. Adult Accuracy in Identifying Presence or Absence of Red Square in Visual Search Task Across all Continua

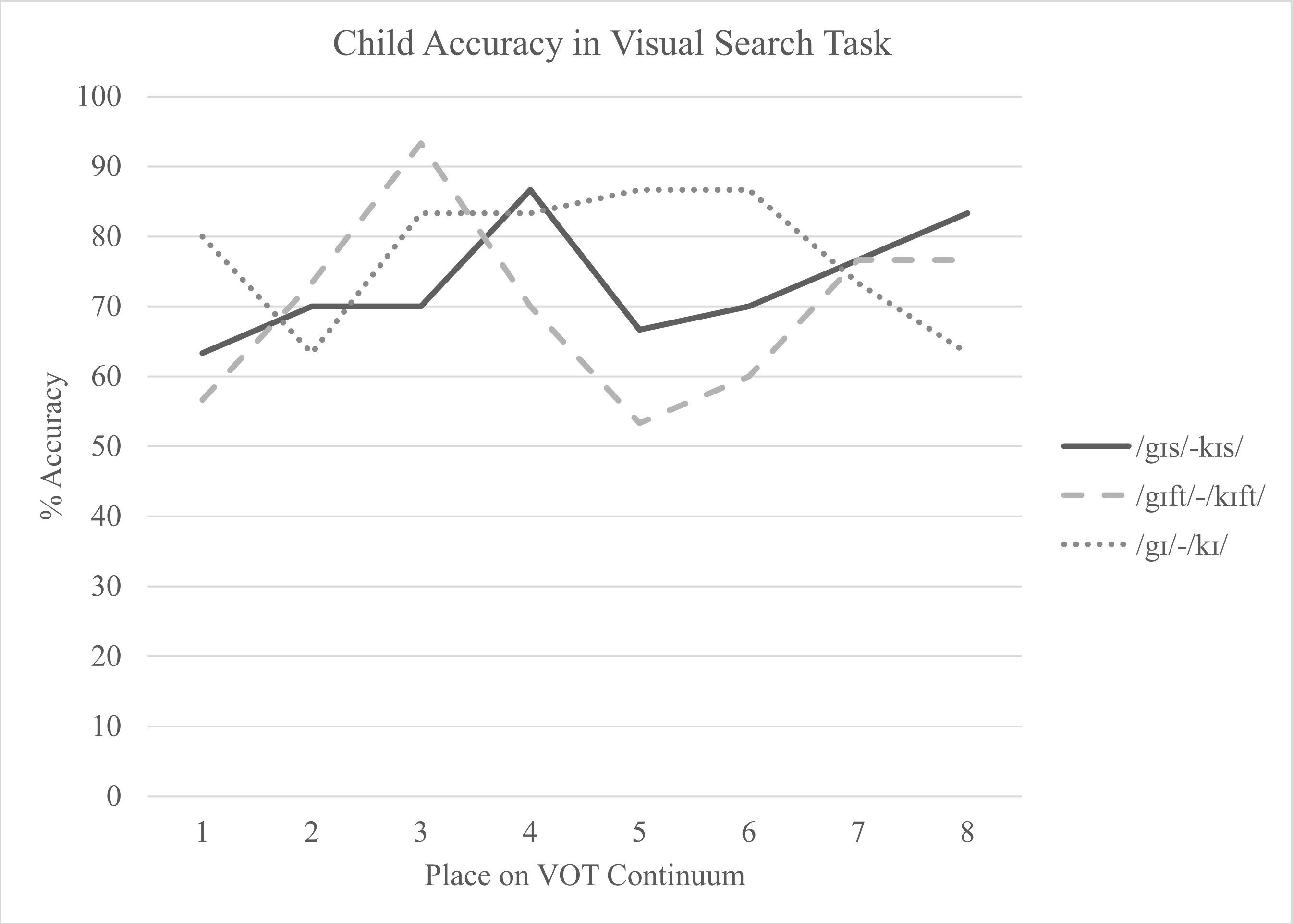

Figure 17. Child Accuracy in Identifying Presence or Absence of Red Square in Visual Search Task Across all Continua

5.0 Discussion

5.1 **Ganong Effect**

The difference in frequency of “K” responses by both children and adults for the three continua in the undivided condition support the presence of the Ganong effect. With the /kɪs/-/gɪs/ continuum, participants identified “K” as the initial sound in ambiguous presentations significantly more than with the other two continua. Therefore, it can be concluded that participants favored the creation of the word, “kiss”, over the creation of the nonword, “giss”. Similarly, participants perceived the initial sound in the /kɪft/-/gɪft/ continuum as “G” with more frequency than in the other two continua. In this situation, participants appeared to favor the creation of the word “gift” over the creation of the nonword “kift”. However, the /kɪft/-/gɪft/ continuum was not substantive different from the nonsense /kɪ/-/gɪ/ continuum. This lack of difference calls into question the lexical influences observed with the /kɪft/-/gɪft/ continuum.

Both area under the curve and VOT boundary were significantly biased towards the end of continua endings resulting in real words. This pattern was consistent with that found by Ganong (1980) and Mattys and Widget (2011), but significant difference between the adult and child groups was not observed. However, the presence of the Ganong effect in children has not been previously reported. It can be assumed that any speech perception, phonemic awareness and lexical immaturity in the child group was not sufficient to affect the overall results. Yet, dismissing any age/maturation effect should be done cautiously given the morphology of the child identification functions and the low number of participants.

5.2 **Divided Attention**

It can be assumed that participants were engaged in the divided attention task given the accuracy with which both adult and child participants identified the presence of the red square in the visual matrix. The divided attention conditions did not appear to strengthen the Ganong effect, especially in the adults. The presence of visual stimuli increased the frequency of “K” responses for adults and children, but the changes were not significant. The frequency of “K” responses, for the /kɪft/-/gɪft/ continuum did not differ from the undivided condition and remained similar to the nonsense CV syllable continuum. As indicated above, the lack of difference between these two continua weakens the argument for lexical influences on perception in this study and may be stimulus specific.

The inconsistencies and lack of differences in this study might a reflection of individual participant variability. For example, only 6 out of 14 adult participants exhibited enhanced lexical drift in the divided attention condition, but only for the /kɪs/-/gɪs/ continuum. These participants, based on VOT boundary alone, exhibited the lexical drift associated with divided attention that was hypothesized. Yet, the participants who exhibited this effect represented less than half of adult participants, and in most cases also exhibited a VOT boundary shift in the same direction in the kɪft/-/gɪft/ continuum, but to a limited extent.

The divided attention results contradict those published by Mattys and Widget (2011), who found that visual divided attention tasks significantly increased the strength of the Ganong effect. It is unclear from the current study why no added lexical drift was found in the divided attention task. It is possible that the use of natural speech stimuli may have affected the results. Mattys and Widget used synthetic speech, which may have heightened the lexical effect in the dual-task condition because of increased perceptual difficulty. The natural speech used in the current study may have included stronger end-point anchors and there also may have been conflicting cues at the mid-points because only VOT and corresponding vowel durations were altered. Other voiceless markers, such as burst amplitude and frequency, and aspiration also may have influenced the perceptions.

In total, the results were somewhat inconclusive on the lexical mediation of ambiguous speech signals during perception, although divided attention did not appear to impact the results. The results also were inconclusive with regards to interactive and autonomous approaches to processing, although the study was not developed to test those types of models. Age and maturity differences could not be documented but increasing the number of children may produce differences.

5.3 **Individual Child Participants**

The child results (individual and group) showed much more variability in their responses than did the adults. In the undivided attention conditions, the majority of children correctly identified endpoints, although participants C3 and C2 were exceptions for specific continua. However, in the divided attention condition, no individual child correctly identified all endpoints of all continua. Furthermore, the functions of many individual children did not continuously increase in percent “K” responses across the continua but fluctuated with increasing VOT. In at least one condition, all of the children experienced both rising and falling percent “K” responses over the course of a single continuum. This variability suggested that children may not have the cognitive resources to attend to the task as consistently as adults.

The children also varied on the screening measures. Participant C2 exhibited notably lower than average Peabody Picture Vocabulary Test scores, suggesting that his language was not as mature as the other children in the study. However, his individual results on the Ganong task did not differ substantively from those of the other child participants, suggesting that high-level vocabulary skills are not necessary to complete the task and show the Ganong effect. The words “gift” and “kiss” likely were sufficiently known by all of the children in the study. However, if more mature lexical knowledge was necessary children with lower receptive vocabulary scores may have performed differently than those with higher vocabulary scores.

Participant C3 showed the most variability, as evidenced by increasing and decreasing functions in both the undivided and divided attention conditions. This participant exhibited the highest Peabody Picture Vocabulary test percentile (96 percentile) but scored the lowest out of the child participants on the training task (83%), and was the oldest of the children (8 years). Perhaps this child experienced difficulties with the task itself, which would account for the low training score and variable undivided and divided attention functions. A simpler task developed for young children may provide insight into whether the task itself impacted the results of this particular child.

5.4 **Further Research**

Because many previous studies using the Ganong procedure used synthetic speech stimuli, a study exploring the within subject differences of the Ganong effect strength between natural and synthetic speech stimuli is warranted. Research directly comparing adult and child perception of ambiguous natural and synthetic stimuli could provide more insight into the differences observed between this study and previous research. Furthermore, isolating the cues in natural speech that prevent lexical drift could provide insight into what elements of natural speech should be included and manipulated in synthetic speech.

Additionally, the limited size of the pediatric group posed an issue in comparing the strength of the Ganong effect between children and adults. A greater sample size in both groups could provide more insight into potential differences between adults and children. Also, more trials of each speech token might result in less variability within and across participants, allowing for more clear distinction between groups and conditions.

Increased reaction time during a Ganong-type tasks has been well documented in other adult studies (McQueen, 1991; Pitt & Samuel, 1993). Adding reaction time as a dependent measure might provide insight into the differences observed between the current study and Mattys and Widget (2011). Mattys and Widget hypothesized that the lexical drift observed in their study might have been due to the latency in response times found in the divided attention condition, and not the divided attention task itself. Although reaction times in the current study was not measured, it is possible that participants did not exhibit response delays in the divided attention condition. The lack of delay would explain the lack of lexical drift observed during the divided attention condition in this study and confirm Mattys and Widgets' suspicions that lexical drift might not be attributed to a lack of cognitive resources.

Administering a phonological awareness screening measure could provide insight into the relationship between phonological awareness and the strength of the Ganong Effect. Because of the recruitment procedures used, it is quite possible that the children in the current study had higher than average language and phonological awareness skills, and therefore were not representative of the general pediatric population. Most of the children had high percentile rankings on the Peabody

Picture Vocabulary Test, so it is possible that they had sufficient cognitive resources and language skills given the task demands.

Finally, developing the Ganong task for use with younger children and children with perceptual, linguistic and cognitive problems may advance our understanding of how more immature and atypical children process speech and how and when they use language to mediate their perceptions. This type of research could advance our understanding of optimal stimulus conditions for auditory learning and training.

Bibliography

Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America*, *113*, 544-552.

Badian, N. A. (1993). Phonemic awareness, naming, visual symbol processing, and reading. *Reading and Writing*, *5*, 87-100.

Blomert, L., & Mitterer, H. (2004). The fragile nature of the speech-perception deficit in dyslexia: Natural vs. synthetic speech. *Brain and Language*, *89*, 21-26.

Brady, S. A., & Shankweiler, D. P. (2013). Phonological awareness: Implications for prereading and early reading instruction. In *Phonological Processes in Literacy* (pp. 57-64). Routledge.

Burton, M. W., Baum, S. R., & Blumstein, S. E. (1989). Lexical effects on the phonetic categorization of speech: The role of acoustic structure. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 567.

Burton, M. W., & Blumstein, S. E. (1995). Lexical effects on phonetic categorization: The role of stimulus naturalness and stimulus quality. Journal of Experimental Psychology: *Human Perception and Performance*, *21*, 1230-1235.

Cardoso-Martins, C., Mesquita, T. C. L., & Ehri, L. (2011). Letter names and phonological awareness help children to learn letter–sound relations. *Journal of experimental child psychology*, *109*, 25-38.

Coady, J. A., Evans, J. L., Mainela-Arnold, E., & Kluender, K. R. (2007). Children with specific language impairments perceive speech most categorically when tokens are natural and meaningful. *Journal of Speech, Language, and Hearing Research*, *50*, 41-57.

Corbin, N. E., Bonino, A. Y., Buss, E., & Leibold, L. J. (2016). Development of open-set word recognition in children: Speech-shaped noise and two-talker speech maskers. *Ear and Hearing*, *37*, 55.

Courage, M. L., Bakhtiar, A., Fitzpatrick, C., Kenny, S., & Brandeau, K. (2015). Growing up multitasking: The costs and benefits for cognitive development. *Developmental Review*, *35*, 5-41.

Crandell, C. C., & Smaldino, J. J. (2000). Classroom acoustics for children with normal hearing and with hearing impairment. *Language, Speech, and Hearing Services in Schools*, *31*, 362-370.

Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody Picture Vocabulary Test*. Circle Pines MN: American Guidance Service.

Evans, J. L., Viele, K., Kass, R. E., & Tang, F. (2002). Grammatical morphology and perception of synthetic and natural speech in children with specific language impairments. *Journal of Speech, Language, and Hearing Research*, *45*, 494-504.

Francis, A. L., & Nusbaum, H. C. (2009). Effects of intelligibility on working memory demand for speech perception. *Attention, Perception, & Psychophysics*, *71*, 1360-1374.

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of experimental psychology: Human Perception and Performance*, *6*, 110-125.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language*, *45*, 468-492.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, *28*, 267-283.

Hall III, J. W., Grose, J. H., Buss, E., & Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing*, *23*, 159-165.

Hembrooke, H., & Gay, G. (2003). The laptop and the lecture: The effects of multitasking in learning environments. *Journal of Computing in Higher Education*, *15*, 46-64.

Jamieson, D. G., Kranjc, G., Yu, K., & Hodgetts, W. E. (2004). Speech intelligibility of young school-aged children in the presence of real-life classroom noise. *Journal of the American Academy of Audiology*, *15*, 508-517.

Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122-149.

Just, M.A., Carpenter, P.A. Keller, T.A., Emery, L., Zajac, H., & Thulborn, K.R. (2001). Interdependence of Nonoverlapping Cortical Systems in Dual Cognitive Tasks. *NeuroImage*, *14*, 417–426

Kingston, J., Levy, J., Rysling, A., & Staub, A. (2016). Eye movement evidence for an immediate Ganong effect. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1968-1998.

Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech, Language, and Hearing Research*, *18*, 686-706.

Koul, R. (2003). Synthetic speech perception in individuals with and without disabilities. *Augmentative and Alternative Communication*, *19*, 49-58.

Leibold, L. J., & Buss, E. (2013). Children's identification of consonants in a speech-shaped noise or a two-talker masker. *Journal of Speech, Language, and Hearing Research*, *56*, 1144-1155.

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and speech*, *10*, 1-28.

Lowenstein, J. H., & Nittrouer, S. (2008). Patterns of acquisition of native voice onset time in English-learning children. *The Journal of the Acoustical Society of America*, *124*(2), 1180-1191.

Lundberg, I., Olofsson, Å., & Wall, S. (1980). Reading and spelling skills in the first school years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of psychology*, *21*, 159-173.

Macken, M. A., & Barton, D. (1980). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*, *7*, 41-74.

Mattys, S. L., & Palmer, S. D. (2015). Divided attention disrupts perceptual encoding during speech recognition. *The Journal of the Acoustical Society of America*, *137*, 1464-1472.

Mattys, S. L., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, *65*, 145-160.

Mattys, S. L., Barden, K., & Samuel, A. G. (2014). Extrinsic cognitive load impairs low-level speech perception. *Psychonomic Bulletin & Review*, *21*, 748-754.

McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception?. *Trends in Cognitive Sciences*, *10*, 363-369.

McNeil, M.R., Pratt, S.R., Szuminsky, N., Sung, J.E., Fossett, T.R.D., Fassbinder, W., & Lim, K.Y. (2015). Reliability and validity of the Computerized Revised Token Test: Comparison of reading and listening versions in persons with and without aphasia. *Journal of Speech, Language, and Hearing Research*, *58*, 1-14.

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, *41*, 329.

Miller, J. L., Grosjean, F., & Lomanto, C. (1984). Articulation rate and its variability in spontaneous speech: A reanalysis and some implications. *Phonetica*, *41*, 215-225.

Morris, R. J., McCrea, C. R., & Herring, K. D. (2008). Voice onset time differences between adult males and females: Isolated syllables. *Journal of Phonetics*, *36*, 308-317.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English?. *Reading Research Quarterly*, *19*, 304-330.

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, *86*, 214-255.

Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *The Journal of the Acoustical Society of America*, *87*, 2705-2715.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*, 204-238.

Pashler, H. (1984). Processing Stages in Overlapping Tasks: Evidence for a Central Bottleneck. Journal of Experimental Psychology: *Human Perception and Performance*, *10*, 358-377.

Pitt, M. A., & McQueen, J. M. (1998). Is compensation for coarticulation mediated by the lexicon?. *Journal of Memory and Language*, *39*, 347-370.

Pitt, M. A., & Samuel, A. G. (1993). An empirical and meta-analytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 699-725.

Reynolds, M., & Jefferson, L. (1999). Natural and synthetic speech comprehension: Comparison of children from two age groups. *Augmentative and Alternative Communication*, *15*, 174-182.

Roring, R. W., Hines, F. G., & Charness, N. (2007). Age differences in identifying words in synthetic speech. *The Journal of the Human Factors and Ergonomics Society*, *49*, 25-31.

Rosner, J., & Simon, D. P. (1971). The auditory analysis test: An initial report. *Journal of Learning Disabilities*, *4*, 384-392.

Utz, T. (2009). *Age and Phonetic Context Effects in Children vs. Adults* (BPhil Honors Thesis, University of Pittsburgh).

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: a meta-analysis. *Psychology and Aging*, *18*, 443-460.

Walley, A. C., & Flege, J. E. (1999). Effect of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics*, *27*, 307-332.

Warren, R. M., & Obusek, C. J. (1971). Speech perception and phonemic restorations. *Perception & Psychophysics*, *9*, 358-362.

Wickens, C. D. (2008). Multiple resources and mental workload. *The Journal of the Human Factors and Ergonomics Society*, *50*, 449-455.

Wilson, R. H., Carnell, C. S., & Cleghorn, A. L. (2007). The Words-in-Noise (WIN) test with multitalker babble and speech-spectrum noise maskers. *Journal of the American Academy of Audiology*, *18*, 522-529.

Windmann, S. (2007). Sentence context induces lexical bias in audiovisual speech perception. *Review of Psychology*, *14*, 77-91.

Yap, M. J., Sibley, D. E., Balota, D. A., Ratcliff, R., & Rueckl, J. (2015). Responding to nonwords in the lexical decision task: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 597-613.

Yopp, H. K., & Yopp, R. H. (2000). Supporting phonemic awareness development in the classroom. *The Reading Teacher*, *54*, 130-143.

Zhang, X., & Samuel, A. G. (2014). Perceptual learning of speech under optimal and adverse conditions. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 200-217.