Research Article

Effects of Noise on Speech Recognition and Listening Effort in Children With Normal Hearing and Children With Mild Bilateral or Unilateral Hearing Loss

Dawna Lewis,^a Kendra Schmid,^{a,b} Samantha O'Leary,^a Jody Spalding,^a Elizabeth Heinrichs-Graham,^a and Robin High^b

Purpose: This study examined the effects of stimulus type and hearing status on speech recognition and listening effort in children with normal hearing (NH) and children with mild bilateral hearing loss (MBHL) or unilateral hearing loss (UHL).

Method: Children (5–12 years of age) with NH (Experiment 1) and children (8–12 years of age) with MBHL, UHL, or NH (Experiment 2) performed consonant identification and word and sentence recognition in background noise. Percentage correct performance and verbal response time (VRT) were assessed (onset time, total duration).

Results: In general, speech recognition improved as signal-to-noise ratio (SNR) increased both for children with NH

and children with MBHL or UHL. The groups did not differ on measures of VRT. Onset times were longer for incorrect than for correct responses. For correct responses only, there was a general increase in VRT with decreasing SNR.

Conclusions: Findings indicate poorer sentence recognition in children with NH and MBHL or UHL as SNR decreases. VRT results suggest that greater effort was expended when processing stimuli that were incorrectly identified. Increasing VRT with decreasing SNR for correct responses also supports greater effort in poorer acoustic conditions. The absence of significant hearing status differences suggests that VRT was not differentially affected by MBHL, UHL, or NH for children in this study.

In classrooms, the ability to hear and understand verbal information is important to the learning process. When listening to speech in the presence of noise, access to acoustic–phonetic speech cues may be limited. Access to this information may be more negatively affected in individuals with hearing loss than in those with normal hearing (NH), making the learning process more difficult for these children. Even mild or unilateral hearing loss (UHL) may reduce the audibility of acoustic information, potentially affecting academic performance.

Children with mild bilateral hearing loss (MBHL) or UHL may experience difficulties in communication, academics, cognition, and psychosocial interactions (e.g., Bess & Tharpe, 1986; Blair, Peterson, & Viehweg, 1985;

Crandell, 1993; Lieu, Tye-Murray, & Fu, 2012; Lieu, Tye-Murray, Karzon, & Piccirillo, 2010; Oyler, Oyler, & Matkin, 1988; Walker et al., 2013). Children with MBHL or UHL are often considered as a homogeneous group (Bess, Dodd-Murphy, & Parker, 1998; McFadden & Pittman, 2008; Porter, Sladen, Ampah, Rothpletz, & Bess, 2013) even though the underlying mechanisms affecting their performance may be different. MBHL can be defined broadly to include a three-frequency pure-tone average (PTA; 0.5, 1.0, and 2.0 kHz) threshold ≥ 20 and ≤ 45 dB HL or thresholds > 25 dB HL at one or more frequencies above 2 kHz in both ears. UHL includes a three-frequency PTA threshold > 20 dB HL in the poorer ear and ≤ 15 dB HL in the better ear or thresholds > 25 dB HL at one or more frequencies above 2 kHz in the poorer ear. Understanding factors that affect speech perception in children with both MBHL and UHL can lead to the development of intervention strategies that improve communication access for these children.

In quiet conditions, school-age children with MBHL or UHL typically demonstrate speech perception abilities that are similar to those of their peers with NH (Bess,

^aBoys Town National Research Hospital, Omaha, NE

^bUniversity of Nebraska Medical Center, Omaha

Correspondence to Dawna Lewis: dawna.lewis@boystown.org

Editor: Nancy Tye-Murray

Associate Editor: Karen Kirk

Received June 6, 2015

Revision received December 4, 2015

Accepted March 6, 2016

DOI: 10.1044/2016_JSLHR-H-15-0207

Disclosure: Dawna Lewis is a member of the Phonak Pediatric Advisory Board. However, no conflicts with the contents of this article exist. The other authors have declared that no competing interests existed at the time of publication.

Tharpe, & Gibler, 1986; Crandell, 1993). However, most real-world communication, including that in classrooms, takes place in noisy environments. For the most part, reports of the effects of noise on speech perception abilities of children have suggested that children with MBHL or UHL perform more poorly than their peers with NH. However, the extent of those differences varies across studies depending on experimental stimuli or tasks and acoustic conditions (Bess et al., 1986; Crandell, 1993; Hartvig Jensen, Angaard Johansen, & Borre, 1989; Johnson, Stein, Broadway, & Markwalter, 1997; Lewis, Valente, & Spalding, 2015; Ruscetta, Arjmand, & Pratt, 2005). In addition, although both children with MBHL and children with UHL may experience poorer speech understanding compared with peers with NH, the underlying mechanisms that lead to these challenges likely differ depending on whether hearing loss is present in one or both ears. For children with MBHL, reduced audibility can decrease the salience of auditory cues in the speech signal. Although the presence of NH in one ear should provide adequate audibility under many conditions for children with UHL, they lack the full benefits of binaural hearing that are available to children with NH or MBHL, which also may affect speech perception, particularly in adverse listening conditions.

A common means of examining the effects of noise on children's speech perception is to determine their accuracy on speech recognition tasks that use phonemes, words, or sentences. When tested across a range of signal-to-noise ratios (SNRs), performance typically decreases as the SNR decreases. When comparing children with NH and children with MBHL or UHL, those with MBHL or UHL often—but not always—show a greater deficit in noise (e.g., Crandell, 1993; Lewis et al., 2015). Although such measures are useful for determining a child's level of speech recognition, percentage correct scores alone may not reflect the cognitive effort that is required for understanding. The cognitive effort a child must expend to understand what is being said may affect speech perception. The cognitive resources that are required to understand spoken speech have been referred to as *listening effort* (e.g., Downs, 1982; Gosselin & Gagne, 2010; Hicks & Tharpe, 2002). Measures beyond simple percentage correct scores may be required to assess how cognitive resources are being allocated even during simple speech recognition tasks and to provide a more complete picture of the functional impact of noise for this population. The current study addresses both speech recognition and listening effort in children with MBHL or UHL and children with NH.

Listening effort often is associated with the principle of a limited resource capacity model (Broadbent, 1958), which assumes that the brain has a finite capacity for attending to sensory input. As such, conditions that negatively affect the audibility of acoustic or phonetic aspects of the speech signal (e.g., poor acoustics, hearing loss) may result in greater allocation of resources for this bottom-up process, leaving fewer resources for top-down (linguistic and contextual) processes that also are important in speech perception. In noise, children may need to allocate greater

resources to this process than adults (Wightman & Kistler, 2005).

A variety of measures have been used to address children's allocation of cognitive resources when listening to speech in noise (Choi, Lotto, Lewis, Hoover, & Stelmachowicz, 2008; Gustafson, McCreery, Hoover, Kopun, & Stelmachowicz, 2014; Hicks & Tharpe, 2002; Howard, Munro, & Plack, 2010; McFadden & Pittman, 2008; Stelmachowicz, Lewis, Choi, & Hoover, 2007). Listening effort often is assessed using a dual-task paradigm in which the listener is asked to attempt to maintain performance on a primary task (often a speech perception task) while also performing a secondary task (e.g., pushing a button each time a randomly presented light is illuminated). As the primary task becomes more difficult, greater resource allocation is required for that task, and a corresponding decrease in performance on the secondary task occurs (Kahnemann, 1973). This decrease in performance is interpreted as a measure of the listening effort required to perform the primary task.

Although dual-task paradigms have been shown to effectively measure listening effort for adults with NH and with hearing loss (Downs, 1982; Feurstein, 1992; Fraser, Gagne, Alepins, & Dubois, 2010; Gosselin & Gagne, 2010; Hornsby, 2013; Rakerd, Seitz, & Whearty, 1996; Sarampalis, Kalluri, Edwards, & Hafter, 2009), the same may not always be true for children (Choi et al., 2008; Howard et al., 2010; McFadden & Pittman, 2008; Stelmachowicz et al., 2007). For example, in a study of 7- to 14-year-olds with NH, Choi et al. (2008) found that children's ability to allocate attention to the primary task was task dependent, thus complicating the interpretation of results in the dual-task paradigm. In their study, the two tasks were a word recognition task and a digit recall task. Word recognition was assigned as the primary task for half of the children, and digit recall was assigned as the primary task for the other half. Regardless of which task was assigned as primary, dual-task effects were found for digit recall but not for word recognition.

Hicks and Tharpe (2002) examined listening effort of children with mild-moderate or high-frequency hearing loss and children with NH using a dual-task paradigm. In their experiment, the primary task was word recognition in quiet and varying levels of background noise. The secondary task was a reaction time measure in which participants were told to push a button as quickly as possible whenever they saw a randomly presented lighted probe. Results revealed that children with hearing loss had longer reaction times for all conditions, even quiet, indicating greater effort for this dual task compared with children with NH. However, average reaction times did not change across conditions for either group, suggesting that neither group expended greater effort as noise levels increased. It is possible that the SNRs used in that study (10, 15, and 20 dB) were not sufficiently difficult to require an increase in listening effort beyond that experienced in quiet for either group.

McFadden and Pittman (2008) questioned whether the lack of an effect in dual-task paradigms for children versus adults could either indicate a difference in the way children perform during simultaneous tasks or be a result

of the tasks used to measure performance. They chose dual tasks that they considered to be common to children's experiences. Children (8–10 years of age) with NH or with minimal hearing loss (bilateral and unilateral) were tested using a word categorization task (primary) and dot-to-dot games (secondary) in quiet and in noise. Although children with hearing loss performed more poorly on the primary task with increased noise levels, there were no significant differences on the secondary task across listening condition or hearing status. Two possible explanations for the results were suggested. According to the authors, it was possible that the children with hearing loss did not realize that they were performing more poorly on the primary task as the acoustic conditions became more difficult, and as a result their effort for that task did not change. Second, the authors stated that it also was possible that their secondary task may have been the easier of the two tasks, making it more likely that the children would focus on that task as conditions became more difficult. In addition to the potential ease of the secondary task, a possibility not suggested by the authors is that children simply were more interested in the dot-to-dot games than they were in the word categorization task.

Another measure that has been used to measure listening effort for speech in both adults and children is verbal response time (VRT; Gatehouse & Gordon, 1990; Gustafson et al., 2014; Houben, van Doorn-Bierman, & Dreschler, 2013; Larsby, Hallgren, Lyxell, & Arlinger, 2005). VRT typically is reported as the time delay between presentation of a speech stimulus and the listener's physical (e.g., selecting items on a keypad) or spoken response (Gustafson et al., 2014; Houben et al., 2013). Studies have shown that response time can increase in poorer listening conditions (Gustafson et al., 2014) as well as under some conditions where effects of increased noise on speech intelligibility scores are small (Tun, Benichov, & Wingfield, 2010).

In the perception of speech, prolonged processing time may limit the amount of new information that can be held in memory at any given time and delay or limit processing of subsequent information. Measures of the time to repeat a speech stimulus may provide information regarding the effort involved in processing that particular signal. It would be expected that shorter, simpler stimuli would require less processing time than longer, more complex stimuli. However, even the former may show differences across increasingly degraded conditions. Assessing listening effort using a variety of materials that commonly are used to assess speech perception in a clinical setting may provide a means of obtaining additional information that is not available when using only percentage correct scores. The goals of the current study were to examine the effects of background noise, stimulus type, age, and hearing status on speech recognition and listening effort in children with NH and children with MBHL or UHL. Speech materials differing in linguistic complexity were used to examine effects of both bottom-up and top-down processing of speech. Identification of consonants in vowel-consonant-vowel nonwords results only from access to the acoustic-phonetic information in the speech signal. These stimuli

also are very short in duration. Single words provide listeners with both acoustic-phonetic and lexical information but could require access to long-term storage of known words in a listener's vocabulary for accurate identification, particularly as noise levels increase. Last, complete sentences provide additional semantic and syntactic information to aid in recognition. However, they also require listeners to hold information in memory for a longer period of time in order to respond correctly and could be influenced by higher level processing skills to a greater extent than the other two stimulus types. Assessing performance at a variety of SNRs provided information across acoustic conditions that can be found in typical classrooms (for a review, see Picard & Bradley, 2001).

Given the potential difficulties that may exist when using traditional dual-task paradigms with children, listening effort was assessed using two measures of VRT. In this study, *onset time* is the delay from the end of the stimulus to the initial vocalization, and *total duration* is onset time plus the time from the initial vocalization to the end of the response (see Figure 1). The measure of onset time is similar to a simple reaction time measure in that it indicates how quickly the listener began to speak once the stimulus ended. However, it is presumed that the listener has processed (or begun processing) the incoming speech signal before beginning to respond. The additional measure of total duration was included on the basis of pilot data suggesting that even when children began their responses quickly, the time required to process the information and complete the response could be longer as acoustics became poorer. In short-term memory tasks, measures of spoken response timing have been used to provide information about the cognitive processes that result in children's correct recall (Cowan, 1992; Cowan et al., 2003). It has been suggested that pauses during verbal responses may reflect both memory search and retrieval.

The first experiment in the current study examined speech recognition and VRT in a group of 45 children with NH. This experiment served as a baseline assessment of acoustic and stimulus effects on verbal processing time across a range of ages in young school-age children with NH. Following findings in previous studies, speech recognition scores were expected to decrease with decreasing SNRs. It was hypothesized that VRT would decrease as SNR increased and that onset time would be longer for incorrect responses, even at the most advantageous SNR. It also was hypothesized that onset time for correct responses would be longer as complexity of the stimuli increased from consonants to words to sentences.

In the second experiment, the effects of hearing loss on performance for the same tasks were examined in a group of children with MBHL or UHL. An equal number of children with NH were tested. The children with NH were selected to provide age-matched counterparts to the children with hearing loss. It was hypothesized that children with MBHL or UHL would demonstrate poorer speech recognition than peers with NH at the poorest SNR but that those with and without hearing loss could show similar

Figure 1. Measurement of onset time and total duration of responses.

performance at the more advantageous SNRs. Across SNRs, it was speculated that children with MBHL or UHL could demonstrate longer VRT than peers with NH but that patterns would be similar across stimuli. Such an outcome would suggest that the children with hearing loss were expending greater listening effort than their peers with NH, even when speech recognition scores were comparable. Within the group of children with hearing loss, several possibilities for outcomes were proposed. First, it was possible that children with UHL would perform better than children with MBHL because the signals were presented at 0° azimuth, and those with UHL would be able to rely on improved audibility at the ear with NH compared with those with MBHL. However, children with UHL have been shown to demonstrate poor speech perception even when speech is presented from the front or is directed to the good ear (Bess et al., 1986; Bovo et al., 1988; Ruscetta et al., 2005), suggesting that factors other than audibility during the task itself may play a role in performance. As such, it was possible that the two groups of children with hearing loss would show similar outcomes.

Experiment 1

Method

Participants

Forty-five children with NH (hearing screened at 15 dB HL from 0.25 to 8.00 kHz) participated in this study. They were divided into five age groups (5, 6, 7–8, 9–10, and 11–12 years) with nine children per group, and there were approximately equal numbers of boys and girls in each group. All of the children were native speakers of English. The Bankson-Bernthal Quick Screen of Phonology (Bankson & Bernthal, 1990) was administered to identify and exclude participants with speech production errors that could influence scoring. The Peabody Picture Vocabulary Test–Fourth Edition (Dunn & Dunn, 2007) was used to assess receptive vocabulary and exclude potential participants who fell below 2 *SD* of age-appropriate norms.

This study was approved by the institutional review board for Boys Town National Research Hospital. Consent and assent were obtained for all children. They were paid \$15/hr for their participation and received a book to take home.

Stimuli

Test materials consisted of three sets of (a) a consonant identification task with 15 vowel–consonant–vowels that were constructed using the consonants /p, b, t, d, g, k, r, l, m, n, s, sh, z, f, v/ in an /a/ context (e.g., *apa, aka, asa*), (b) 15 phonetically balanced kindergarten monosyllabic words (Haskins, 1949), and (c) 15 Bamford-Kowal-Bench sentences (Bench, Kowal, & Bamford, 1979) with three target words in each sentence.¹ Stimuli were recorded by a single female talker and digitally mixed with speech-shaped noise using MATLAB (MathWorks, Natick, MA) to create three SNRs (–5, 0, and 5 dB). The presentation order of the stimulus groups was consonants, monosyllabic words, and sentences. Within a stimulus group, presentation order of the SNR conditions was determined using a Latin square design. A Latin square design also was used for the sets of stimuli to vary the order of presentation. The total number of stimulus presentations was 135 per participant (15 consonants, 15 real words, and 15 sentences at each of the three SNRs).

Procedures

All participants were tested in a double-walled sound-treated booth. Speech stimuli were presented via a loudspeaker located 1 m from the child at 0° azimuth. The output of the loudspeaker was calibrated at an average root-mean-square level of 65 dB SPL. Recordings of responses

¹The three sets of 15 words all were taken from the first 45 words of List 1 of the phonetically balanced kindergarten word lists. The word frequency and neighborhood density for each set were determined using the Speech and Hearing Lab Neighborhood Database (Washington University in St. Louis, 2016) and are reported in the Appendix.

were made from a head-worn microphone located at a 45° angle from the child's mouth. Children were instructed to listen and repeat the stimuli exactly as heard. If they were uncertain about a response, they were instructed to make a guess. For all tasks, items were presented using a computer game format with visual feedback (e.g., removal of a puzzle piece to reveal an interesting picture) given immediately after each response. This feedback was not contingent on correct responses and was used only to maintain interest in the task. Testing was completed in a single 1-hr session, and children were given breaks as needed throughout the session. Consonants and real words were scored as either correct or incorrect. Sentences were scored as correct only when all three key words within each sentence were produced accurately. Responses were recorded and a second experimenter labeled the response onset time, length of utterance, and correctness of the response using Praat (<http://www.praat.org>). Any differences in scores were discussed for a final consensus regarding the score for that response.

Statistical Analysis

As is well documented (Gatehouse & Gordon, 1990; Houben et al., 2013; Ratliff, 1979; Whelan, 2008; Zumbo & Coulombe, 1997), there are several methodological concerns in using common statistical methods for response time data analysis. Response time and percentage correct data are known to be skewed, and individual responses are not statistically independent even with carefully controlled, randomized trials. In addition to variance from fixed factors such as experimental tasks, group, age, and so on, the additional subject variance must be taken into account as a random effect. We expected similar variance for the percentage correct and VRT data in the current study. For that reason, as suggested by Baayen and Milin (2010), statistical analysis was conducted using generalized linear mixed models that accounted for both fixed and random effects as well as the nonnormally distributed nature of the data.

Generalized linear mixed models were used with a logit link for the correct and incorrect responses and a gamma link for onset and total duration. SNR, stimulus type, and the interaction between SNR and stimulus type along with age as a continuous variable and interactions of age and stimulus were included in the analysis as fixed effects. The models also included random subject effects to account for correlations due to multiple observations collected from each participant across the combinations of conditions.

Measures of onset time for correct and incorrect responses were included in the model to determine whether there were differences in onset time when participants did or did not respond correctly. Onset time was also analyzed separately for correct responses alone to examine potential differences between stimuli. Total duration was analyzed only for correct responses due to potential duration variability in incorrect responses that could obscure potential interpretation of the results (e.g., large and opposite duration differences between incomplete sentences and sentences with added words). The gamma link in the generalized linear model accounts for skewness of the distributions for onset

and duration times. For that reason, extreme values were not excluded from the analysis.

Simple effects for significant interactions were computed with differences in means for all pairs of levels for one factor at each level of the other factor. For outcomes significantly affected by the interaction of age and stimulus, because age was continuous, it was fixed at 8 and 12 years to investigate differing effects of stimulus at different ages on the outcomes. Multiple comparisons of differences in means were adjusted with the simulation technique—the recommended approach for correlated data models (Westfall, Tobias, & Wolfinger, 2011). These analyses were computed with PROC GLIMMIX from SAS/STAT (Version 9.4) of the SAS System for Windows (SAS Institute, Cary, NC).

Results

Percentage Correct Performance

Figure 2 displays percentage correct performance across SNR for the three types of stimuli. Examination of the figure suggests an improvement in performance for all stimulus types as SNR improved. However, the pattern of performance for stimulus type varied across SNRs. Statistical analysis revealed that the interaction between SNR and stimulus type was significant, confirming that the percentage correct for different stimulus types was not consistent across SNRs, $F(4, 6019) = 27.8, p < .001$. In general, percentage correct increased with increasing SNR (–5, 0, and 5 dB), but the pattern of percentage correct responses for the different stimulus types within an SNR was not consistent across SNRs. When examining the pairwise comparisons within each SNR, all comparisons were significant at $p < .001$ except word versus sentence at an SNR of –5 dB, which was not statistically significant,

Figure 2. Percentage correct for given signal-to-noise ratios (SNR) and stimuli. Boxes represent the 25th to 75th percentiles, and whiskers represent minimum and maximum, excluding extreme values. Within each box, diamonds represent the mean and dashed lines represent the median. Filled circles represent extreme values.

$t(6019) = 0.9, p = .66$. For SNRs of 0 and 5 dB, both sentences and consonants had higher percentage correct than words. There was a significant interaction effect between age and stimulus on percentage correct, $F(2, 6019) = 8.12, p = .0003$, with the odds of a correct response for sentences increasing more with each year of age than the odds of a correct response for consonants and words (odds ratio = 1.11, 1.06, and 1.01, respectively).

Onset Time

Onset time was first analyzed including an indicator variable for correct or incorrect response as a predictor. Here we found a significant interaction of age and response, $F(1, 5605) = 15.6, p < .001$, with shorter onset times occurring with increased age and correct responses. Onset times increased approximately 20 ms for each year increase in age for incorrect responses and decreased approximately 2 ms for each year increase in age for correct responses.

All other two-way interactions were also significant at $p < .001$. As shown in Figure 3a, mean onset times were longer for incorrect responses, but the differences in onset times between correct and incorrect responses varied given the SNR and stimulus type. As seen in the top panel, the differences in mean onset times for incorrect versus correct responses increased with increasing SNR (from 156 ms at 5 dB SNR to 275 ms at 0 dB SNR and 497 ms at -5 dB SNR). The mean difference in onset times between incorrect and correct responses also increased with stimulus complexity (bottom panel). Differences were shortest for consonant identification (152 ms), followed by words (289 ms) and sentences (441 ms).

When both correct and incorrect responses were included, the interaction of SNR and stimulus type (see Figure 3b) revealed that onset times generally were longer for -5 dB SNR, but the pattern of onset times for given stimulus types varied across levels of SNR, similar to that observed with other outcomes. When examining pairwise

Figure 3. (a) Mean onset times (ms) and 95% confidence intervals [error bars] as a function of signal-to-noise ratio (SNR; -5, 0, and 5 dB) and response (Rsp; top panel) and as a function of stimulus (STIM) type (CONS = consonants; RW = real words; SNT = sentences) and response (bottom panel). (b) Mean onset times (ms) and 95% confidence intervals [error bars] depicting the interaction between SNR and stimulus type.

comparisons of levels of stimulus for these interaction effects within each SNR, all pairwise comparisons were significant other than the difference in onset time between words and sentences at 0 dB SNR, $t(5605) = 0.05$, $p = 1.0$, and the difference between consonants and sentences at 5 dB SNR, $t(5605) = 0.1$, $p = .60$.

In analyzing onset time for correct responses only (see Figure 4), a significant interaction between SNR and stimulus was observed, $F(4, 3720) = 8.3$, $p < .001$. In general, the onset times were longest for -5 dB SNR and decreased as SNR increased. However, the pattern of mean onset times for stimulus type within SNR was not consistent. When examining pairwise differences within each SNR, there was not a difference in onset time between words and sentences at -5 dB SNR, $t(3720) = -0.10$, $p = .99$; consonants and words or sentences at 0 dB SNR, $t(3720) = -2.2$, $p = .07$ and $t(3720) = 0.13$, $p = .99$, respectively; or consonants and words at 5 dB SNR, $t(3720) = 1.7$, $p = .20$. All other pairwise comparisons were significantly different. There was no significant effect of age on onset time of correct responses, $F(1, 3720) = 0.4$, $p = .52$. These patterns are in agreement with those seen for the percentage correct scores, indicating that children took longer to process even their correct responses for the speech stimuli for which they demonstrated poorest performance.

Total Duration

When examining the total duration of correct responses, there was neither a significant effect of age nor a significant interaction between SNR and stimulus type: $F(1, 3719) = 2.0$, $p = .15$, and $F(4, 3719) = 2.1$, $p = .08$, respectively. Significant main effects of both SNR and stimulus type were observed: $F(2, 3719) = 129.1$, $p < .001$, and $F(2, 88) = 955.3$, $p < .001$, respectively. Figure 5 shows means and corresponding 95% confidence intervals for duration of correct responses as a function of SNR and stimulus type. For all stimulus types, duration decreased as SNR increased. As expected on the basis of stimulus

Figure 4. Mean onset times (ms) and 95% confidence intervals [error bars] across stimulus type and signal-to-noise ratios (SNR) for correct responses.

Figure 5. Mean total duration (ms) and 95% confidence intervals [error bars] for correct responses across signal-to-noise ratios (SNR; top panel) and stimulus type (bottom panel; CONS = consonants; RW = real words; SNT = sentences).

length, duration of sentences was significantly longer than that of both consonants and real words. However, there was no difference between consonants and real words.

Discussion

As hypothesized, the results of Experiment 1 indicated that speech recognition for children with NH was negatively affected by noise. However, the interaction between speech recognition and stimulus type was more complex. At the poorest SNR (-5 dB), scores for consonant identification were higher than those for real words and sentences. At both 0 and 5 dB SNR, scores for both consonants and sentences were higher than those for words. At the poorest SNR, access to acoustic-phonetic information in the signal would have been the most limited. For consonant identification, only one portion of the stimulus changed from one presentation to the next. As such, less effort would be required to recognize these stimuli than would be required for real words or sentences. Better speech recognition scores were obtained for sentences than for real words at the two higher SNRs. These findings suggest that children may have been able to use both acoustic-phonetic and linguistic or contextual information in the sentences to assist with understanding. As demonstrated at the poorest SNR, however, linguistic or contextual information may not have provided sufficient additional support when acoustic-phonetic information was most limited.

Listening effort was assessed using two measures of VRT: onset time and total duration. As predicted, mean onset times were longer for incorrect than for correct responses for all SNRs and stimulus types, suggesting that greater listening effort was expended during attempts to process those signals. In addition, onset time decreased

as age increased, suggesting greater listening effort for younger children. However, relationships among variables differed across stimulus types and SNRs, indicating a complex relationship where listening effort was differentially affected by stimulus type at different levels of background noise. The differences in onset times between correct and incorrect responses were greatest at the poorest SNR (–5 dB) and for the most complex stimulus (sentences), suggesting increased effort under these conditions.

When evaluating only correct responses, there was a general decrease in both onset time and duration as SNR increased. Thus, even when participants were able to correctly identify the stimuli, listening effort (as measured by VRT) was greatest in the poorer acoustic conditions. Within SNRs, the relationships of onset times across stimulus types were similar to those seen for percentage correct scores (see Figures 2 and 4 for comparison), supporting the complex nature of the relationship. No effect of age was found for either measure, indicating that effort did not differ with age when responses were correct.

The findings in Experiment 1 support the extant literature, showing that speech perception in children with NH decreases as SNR decreases. VRT results suggest a complex relationship between age, linguistic content of the speech materials, and listening effort. However, the general trend of increasing VRT with decreasing SNR held throughout. Differences in VRT across SNR, even for correct responses, support the hypothesis that percentage correct scores alone may not provide a complete picture of the cognitive effort required for children's speech recognition in noise.

Experiment 2

Although there is a body of research showing that children with MBHL or UHL exhibit poorer speech recognition in noise than children with NH (Bess et al., 1986; Crandell, 1993; Hartvig Jensen et al., 1989; Johnson et al., 1997; Ruscetta et al., 2005), efforts to examine listening effort have shown limited effects using dual-task paradigms (e.g., Hicks & Tharpe, 2002; McFadden & Pittman, 2008). To examine how MBHL or UHL may affect both speech recognition and VRT, Experiment 2 compared the same tasks in a group of children with MBHL or UHL with an age-matched group of children with NH. It was hypothesized that children with MBHL or UHL would perform more poorly than their peers with NH on the speech recognition task as SNR decreased and that they would expend greater listening effort on the tasks, manifested as increased VRT. However, it also was possible that listening effort (as measured by VRT) would differ within the group of children with hearing loss depending on whether the hearing loss was in one or both ears.

Method

Participants and Procedures

The children who participated in Experiment 2 were recruited as participants in a series of studies in this lab

examining auditory skills in children with MBHL or UHL (e.g., Lewis et al., 2015). For that series of studies, the lower end of the age range was higher than that for participants who took part in Experiment 1. For Experiment 2, eighteen 8- to 12-year-old children with NH and 18 children with MBHL or UHL participated. Ten of the children with hearing loss presented with UHL (five right ear, five left ear), and eight presented with MBHL. For this experiment, UHL was defined as a three-frequency pure-tone average (PTA) > 20 dB HL in the poorer ear and ≤ 15 dB HL in the better ear or thresholds > 25 dB HL at one or more frequencies above 2 kHz in both ears. Bilateral hearing loss was defined as a three-frequency PTA ≥ 20 and ≤ 45 dB HL or thresholds > 25 dB HL at one or more frequencies above 2 kHz in both ears. For seven of the children with MBHL, the mean better ear PTA was 34.2 dB HL ($SD = 8.3$ dB) and the poorer ear PTA was 36.7 dB HL ($SD = 7.1$ dB). One participant with MBHL had a high-frequency hearing loss. That participant's better ear high-frequency PTA was 45 dB HL (6–8 kHz; left ear) and poorer ear PTA was 50 dB HL (8 kHz). For eight of the children with UHL, the poorer ear PTA was 55 dB HL ($SD = 11.7$ dB). Two participants with UHL presented with high-frequency hearing loss. For one, the high-frequency PTA was 82.5 dB HL (3–8 kHz; right ear), and for the other it was 88.3 dB HL (4–8 kHz; right ear). Per parent report, eight children had congenital hearing loss, three had acquired hearing loss, and seven were unknown.

To recruit sufficient numbers of children with MBHL or UHL in the 8- to 12-year age range, the distribution of children across years of age was not equal. As part of the series of studies examining children with MBHL or UHL, new participants with NH were recruited and were age matched to the children with MBHL or UHL. The age distribution of participants in the group with hearing loss and the group with NH was the same: four 8-year-olds, two 9-year-olds, two 10-year-olds, six 11-year-olds, and four 12-year-olds.

All children scored within 1.25 SD s of the mean on the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999). There was no significant difference between groups on the Peabody Picture Vocabulary Test–Fourth Edition, $t(34) = 0.791$, $p = .435$, and average standard scores were within 1 SD of the mean for both groups (NH: $M = 108.9$, $SD = 16.1$; MBHL or UHL: $M = 104.9$, $SD = 13.7$).

Although children with MBHL or UHL may be fitted with hearing aids, this is not universal and those who have hearing aids may not wear them consistently (American Academy of Audiology, 2013; Davis, Reeve, Hind, & Bamford, 2002; Fitzpatrick, Durieux-Smith, & Whittingham, 2010; Fitzpatrick, Whittingham, & Durieux-Smith, 2014; Walker et al., 2013). To obtain baseline performance from these children with potentially common but nonoptimal audibility, the children with MBHL or UHL did not wear amplification during testing.

This experiment was approved by the institutional review board for Boys Town National Research Hospital. Consent and assent were obtained for all children.

Children were paid \$15/hr for their participation and received a book to take home.

The stimuli and procedures were identical to those utilized in Experiment 1 with the exception that the Bankson-Bernthal Quick Screen of Phonology was not used to assess articulation errors for these older children.

Statistical Analysis

Experiment 2 was similar to Experiment 1, with a different group of children, including both those with MBHL or UHL and age-matched children, being tested. The statistical analysis of Experiment 2 follows that of Experiment 1 with the addition of a categorical predictor variable indicating whether a participant had hearing loss or NH.

Results

For the children with hearing loss, it was possible that there would be differences between children with MBHL and those with UHL across measures as a function of type of hearing loss. To examine this possibility, separate analyses of the hearing loss group were conducted with type of hearing loss as the between-subjects factor. Results revealed no significant differences across type of loss for percentage correct performance, $F(1, 2333) = 2.55, p = .11$; onset time for all responses, $F(1, 2192) = 0.99, p = .32$; or onset time for correct responses, $F(1, 1379) = 3.66, p = .06$. No significant interactions between type of hearing loss were found with either SNR or stimulus for any of the aforementioned outcomes. For total duration, there was a significant interaction between hearing loss group and stimulus, $F(2, 1469) = 6.65, p = .001$. For the tasks examined in the current study, children with UHL and MBHL were combined for all analyses, with additional separate reporting of the interaction between hearing loss group and stimulus for total duration.

Percentage Correct

Figure 6 displays percentage correct performance across SNR and hearing status group for the three types of stimuli. The analysis of the percentage of correct responses revealed significant interaction effects between hearing status group and stimulus type, $F(2, 68) = 5.4, p = .006$, and between SNR and stimulus type, $F(4, 4666) = 14.0, p < .001$. There was neither a significant interaction between hearing status group and SNR, $F(2, 68) = 0.6, p = .56$, nor a significant effect of age, $F(1, 4666) = 1.6, p = .20$. The children with NH performed better than the children with MBHL or UHL on consonants, which relied on acoustic-phonetic information, $t(68) = 2.9, p = .004$, and on sentences, which provided contextual information but also required greater processing, $t(68) = 4.4, p < .001$. However, the groups did not differ for words, $t(68) = 1.31, p = .20$. The interaction effect between SNR and stimulus was similar to that observed in Experiment 1 in that percentage correct increased with increasing SNR ($-5, 0$, and 5 dB), but the pattern of percentage correct responses for a

given stimulus type was not consistent within a given SNR. As would be expected, for all stimuli the largest difference in percentage correct was between -5 and 5 dB. However, the difference between -5 and 0 dB was much larger for sentences (33.8% – 86.0%) than for the other two stimuli (33.0% – 61.6% for words; 48.2% – 80.0% for consonants), whereas the difference in percentage correct between 0 and 5 dB was similar for all (9% – 20% difference). This effect did not depend on hearing status, because the three-way interaction was not significant.

Onset Time

Onset time, when including all responses (correct and incorrect), was not significantly different between the group with MBHL or UHL and the group with NH, $F(1, 4616) = 0.9, p = .37$. The analysis of onset times for all responses yielded a significant interaction of age and stimulus, $F(2, 4616) = 7.71, p < .001$. For the youngest children, consonants had the shortest mean onset times, followed by words and then sentences ($535, 585$, and 674 ms, respectively). For the oldest children, the longest onset times were for words (consonants = 524 ms; words = 656 ms; sentences = 605 ms). The interactions between SNR and stimulus, $F(4, 4616) = 2.8, p = .03$, and between stimulus and response, $F(1, 616) = 37.4, p < .001$, were also significant (see Figure 7). For the interaction between stimulus and response (see Figure 7a, bottom panel), all pairwise comparisons were significant, but the differences in onset time between correct and incorrect responses was larger for sentences than for the other stimuli, resulting in the significant interaction. For the interaction between SNR and stimulus (see Figure 7b), all pairwise comparisons were significant other than the difference between real words and sentences at SNRs of 0 and 5 dB— 0 dB SNR: $t(4616) = 1.1, p = .60$; 5 dB SNR: $t(4616) = 1.6, p = .30$.

The analysis of onset times for only correct responses revealed no significant interactions between hearing status group, SNR, and stimulus type. There was a significant interaction between age and stimulus, $F(2, 3166) = 4.99, p = .007$. For consonants, onset time decreased approximately 10 ms for each 2-year increase in age, whereas the decrease was approximately 17 ms for sentences. Onset time for words actually increased with age (approximately 33 ms per 2-year increase in age). As observed in Experiment 1, there was a significant main effect of SNR, $F(2, 68) = 43.6, p < .001$, with mean onset times decreasing as SNR increased (see Figure 8). There was no significant effect of hearing status on mean onset time, $F(1, 34) = 1.3, p = .30$. Thus, for both children with MBHL or UHL and children with NH, acoustics affected processing time for all stimuli, even when they were correctly identified.

Total Duration

When examining the total duration of correct responses, results were similar to those of Experiment 1. Significant interactions were observed for hearing status by stimulus type, $F(4, 3150) = 5.5, p < .001$; SNR by stimulus type, $F(4, 3150) = 3.7, p = .005$; and age by stimulus,

Figure 6. Percentage correct across stimulus type and signal-to-noise ratios (SNR) for participants with mild bilateral hearing loss (MBHL) or unilateral hearing loss (UHL; gray) and participants with normal hearing (NH; white). Boxes represent the 25th to 75th percentiles, and whiskers represent minimum and maximum, excluding extreme values. Within each box, diamonds represent the mean and dashed lines represent the median. Filled circles represent extreme values.

Figure 7. (a) Mean onset times (ms) and corresponding 95% confidence intervals [error bars] for hearing status (top panel) and for the interaction between stimulus (STIM) type and response (Rsp, bottom panel; CONS = consonants; RW = real words; SNT = sentences). (b) Mean onset times (ms) and corresponding 95% confidence intervals [error bars] for the interaction between signal-to-noise ratio (SNR) and stimulus.

Figure 8. Mean onset times and corresponding 95% confidence intervals [error bars] for correct responses across signal-to-noise ratios (SNR).

$F(2, 3150) = 23.1, p < .001$ (see Figure 9). Within hearing status groups, all pairwise comparisons of stimuli were significant at $p < .001$ other than the difference between consonants and words for children with NH, $t(3150) = -1.17, p = .48$, and children with UHL, $t(3150) = -1.47, p = .31$. The pattern of duration varied by stimuli and type of hearing loss. For consonants, duration for children with UHL was much lower than for the other two groups (889 vs. 1,038 ms for children with MBHL and 1,027 ms for children with NH), but for sentences, children with MBHL had the longest total duration (2,210 vs. 1,952 ms for children with UHL and 2,080 ms for children with NH). As in Experiment 1, for all levels of SNR, duration of sentences was significantly longer than both consonants and words. However, differences between consonants and words were significant for -5 dB SNR and 5 dB SNR (both $p = .02$), with words averaging 83 ms longer at -5 dB SNR and 43 ms longer at 5 dB SNR. In examining the interaction between age and stimulus, the youngest children had longer durations than the oldest children for sentences (2,256 vs. 1,914 ms), but the duration for other stimuli was similar (consonants: 1,013 vs. 952 ms; words: 1,010 vs. 1,061 ms). In the youngest children, there was no significant difference in duration between consonants and words, but all other pairwise comparisons were significant with $p < .001$.

Discussion

Experiment 2 examined speech recognition and listening effort in 8- to 12-year-old children with MBHL or UHL and their peers with NH. Speech recognition improved similarly for both groups as SNR increased. As in Experiment 1, the pattern of responses across stimulus types within SNRs varied. Children with MBHL or UHL performed more poorly than their peers for both consonants and sentences but not for real words. Because the children with NH would have had a history of consistently better access to acoustic-phonetic information in speech signals than children with MBHL or UHL, they may have been

Figure 9. Mean total durations and corresponding 95% confidence intervals [error bars] across hearing status and stimulus type (top panel) and signal-to-noise ratio (SNR) and stimulus type (bottom panel) for correct responses in participants with mild bilateral hearing loss (MBHL) or unilateral hearing loss (UHL) and participants with normal hearing (NH). CONS = consonants; RW = real words; SNT = sentences.

better able to perceive the single target sounds in the consonants. For the sentences, it is possible that children with NH also were better able to use the available acoustic-phonetic information, supplemented by linguistic and contextual information to assist in speech recognition. For single real words, linguistic and contextual information is more limited

than for sentences, and the acoustic–phonetic content is greater than in consonants. The limited amount of both types of information may not have been sufficient to bolster their scores relative to peers with MBHL or UHL for these stimuli.

As in Experiment 1, onset times were longer for incorrect than for correct responses, and these differences were greater for sentences than for real words or consonants. Such results suggest that greater effort was expended when attempting to process speech inputs that were not correctly identified and that the time required for that attempt was longer for the more complex, sentence-length material. The fact that both onset time and total duration increased with decreasing SNR, even when stimuli were correctly identified, also suggests that additional cognitive processing was required for poorer acoustic conditions. These efforts did not differ across age.

Although there were differences in speech recognition across the hearing status groups, the groups did not differ for either of the VRT measures. These results can be compared to those of studies examining listening effort in children with hearing loss using dual-task paradigms. Hicks and Tharpe (2002) showed differences in listening effort between children with NH and children with greater degrees of (mild–moderate) or high-frequency hearing loss when using a dual-task paradigm. In that study, children with hearing loss exhibited longer reaction times for the secondary task than did children with NH. Although they also showed poorer speech recognition scores on the primary task, the means for both groups were $\geq 85\%$. McFadden and Pittman (2008) compared performance for children with NH and children with MBHL or UHL. In their study, the children with hearing loss exhibited poorer word categorization than peers with NH as noise increased. However, the dual-task paradigm resulted in similar effects on performance for the secondary task for both groups. The absence of significant hearing status differences in the current study suggests that listening effort (as measured by VRT) was not differentially affected by MBHL or UHL for children from 8 to 12 years of age when the groups were matched for age and did not differ in terms of IQ or receptive vocabulary. Taken together, findings show effects of degree of hearing loss on measures of listening effort, with children who have greater degrees of hearing loss exhibiting effects not seen in children with mild or unilateral loss. Further research is needed to address these potential differences.

General Discussion

This study examined speech recognition and VRT in noise in children with NH and children with MBHL or UHL. VRT was used as a means of examining listening effort in these children. It was hypothesized that speech recognition would decrease and VRT would increase as SNR decreased and that younger children would perform more poorly than older children. It also was hypothesized that children with MBHL or UHL would demonstrate

poorer speech recognition and longer VRT than children with NH.

There were differences in age effects across the two experiments. These differing findings may be related to the smaller number of participants in Experiment 2 relative to Experiment 1 as well as the inclusion of younger children in the first experiment. Additional studies with larger numbers of participants with MBHL or UHL across a range of ages will be necessary to address these differences and clarify potential age effects.

In both experiments, onset times were longer for incorrect than for correct responses. These findings indicate that children with NH and those with MBHL or UHL spent more time attempting to process stimuli that they were unable to correctly identify and that processing time increased both as the acoustics deteriorated and as the stimulus became more complex. In addition, when examining correct responses only, there was a general increase in onset time and total duration for decreasing SNRs. Thus, even when these children were able to correctly identify the speech stimuli, results suggest that greater effort was expended as the acoustic conditions deteriorated. In classrooms, where acoustic conditions are often less than ideal (Crandell & Smaldino, 1994; Knecht, Nelson, Whitelaw, & Feth, 2002; Sato & Bradley, 2008), the need to exert greater listening effort may help account for reports of increased fatigue, decreased attention, and poorer academic performance (Dockrell & Shield, 2006; Hornsby, Werfel, Camarata, & Bess, 2013; Shield & Dockrell, 2008).

Two measures of VRT—onset time and total duration—were used in the current study. Onset time measures provided information about processing of both incorrect and correct responses, whereas total duration could be used only to assess correct responses. These differences in response time may provide useful information about the effort exerted during attempts to process speech that is difficult to understand. In addition, duration measures were of limited value in comparing performance across stimuli. As such, onset time appears to be the more useful measure of the two when assessing listening effort using these stimuli.

The absence of a hearing status effect in Experiment 2 could indicate that the stimuli and acoustic conditions used in this study were not sufficiently difficult to differentially affect VRT in the children with MBHL or UHL. Measures of VRT were chosen to obtain additional information beyond that provided by percentage correct scores alone using stimuli that are commonly used to assess speech recognition in clinical settings. However, these speech recognition tasks that used short-duration stimuli (i.e., words) or high-context sentences may not have resulted in a cognitive load that was significantly greater for the children with MBHL or UHL. Tasks that are higher in the listening hierarchy (Erber, 1982) and that tax more of the cognitive resources available for speech understanding may be necessary to differentiate listening effort for this population. Studies that include children with greater degrees of hearing loss will be needed to determine whether these materials can show differences for those populations.

The current study used speech noise as a masker. Research has shown that speech perception in the presence of competing speech is a better predictor of functional hearing than performance in steady noise (Brungart, 2001; Hall, Grose, Buss, & Dev, 2002; Hillock-Dunn, Taylor, Buss, & Leibold, 2014). In addition, children with hearing loss have shown a greater effect of competing talkers on speech perception than have children with NH (Hillock-Dunn et al., 2014; Leibold, Hillock-Dunn, Duncan, Roush, & Buss, 2013). Future tasks that utilize maskers that more closely resemble those that children will encounter in real-world listening conditions may present a more accurate picture of their performance and, in turn, may be better at differentiating children with MBHL or UHL from their peers with NH.

In summary, VRT (as measured by onset time) has potential as a means of examining listening effort in children using the types of speech recognition tasks common in clinical practice. However, further studies using materials with increased cognitive load are needed to assess potential differences in listening effort in children with MBHL or UHL and children with NH.

Acknowledgments

This work was supported by National Institutes of Health Grants R03 DC009675 (awarded to Dawna Lewis), T35 DC08757 (awarded to Samantha O'Leary; Michael Gorga, PI), P20 GM109023 (awarded to Dawna Lewis; Walt Jesteadt, PI), and P30 DC004662 (Michael Gorga, PI). The content of this article is the responsibility and opinions of the authors and does not necessarily represent the views of the National Institutes of Health. We appreciate the contributions of Kanae Nishi in development of Praat scripts and procedures for coding verbal response times.

References

- American Academy of Audiology.** (2013). *Clinical practice guidelines. Pediatric amplification*. Retrieved from http://audiology-web.s3.amazonaws.com/migrated/PediatricAmplificationGuidelines.pdf_539975b3e7e9f1.74471798.pdf
- Baayen, R. H., & Milin, P.** (2010). Analyzing reaction times. *International Journal of Psychological Research*, 3, 12–28.
- Bankson, N., & Bernthal, J.** (1990). *Bankson-Bernthal Test of Phonology*. San Antonio, TX: Special Press.
- Bench, J., Kowal, A., & Bamford, J.** (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology*, 13, 108–112.
- Bess, F., Dodd-Murphy, J., & Parker, R.** (1998). Children with minimal sensorineural hearing loss: Prevalence, educational performance, and functional status. *Ear and Hearing*, 19, 339–354.
- Bess, F., & Tharpe, A.** (1986). Case history data on unilaterally hearing-impaired children. *Ear and Hearing*, 7, 14–19.
- Bess, F., Tharpe, A. M., & Gibler, A. M.** (1986). Auditory performance of children with unilateral hearing loss. *Ear and Hearing*, 7, 20–26.
- Blair, J., Peterson, M., & Viehweg, S.** (1985). The effects of mild sensorineural hearing loss on academic performance of young school-age children. *The Volta Review*, 87, 87–93.
- Bovo, R., Martini, A., Agnoletto, M., Beghi, A., Carmignoto, D., Milani, M., & Zangaglia, A. M.** (1988). Auditory and academic performance of children with unilateral hearing loss. *Scandinavian Audiology Supplement*, 30, 71–74.
- Broadbent, D.** (1958). *Perception and communication*. New York, NY: Pergamon Press.
- Brungart, D.** (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, 109, 1101–1109.
- Choi, S., Lotto, A., Lewis, D., Hoover, B., & Stelmachowicz, P.** (2008). Attentional modulation of word recognition by children in a dual-task paradigm. *Journal of Speech, Language, and Hearing Research*, 51, 1042–1054.
- Cowan, N.** (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory and Language*, 31, 668–684.
- Cowan, N., Towse, J., Hamilton, Z., Saults, J. S., Elliott, E., Lacey, J., ... Hitch, G.** (2003). Children's working-memory processes: A response-timing analysis. *Journal of Experimental Psychology: General*, 132, 113–132.
- Crandell, C.** (1993). Speech recognition in noise by children with minimal degrees of sensorineural hearing loss. *Ear and Hearing*, 14, 210–216.
- Crandell, C., & Smaldino, J.** (1994). An update of classroom acoustics for children with hearing impairment. *The Volta Review*, 96, 291–306.
- Davis, A., Reeve, K., Hind, S., & Bamford, J.** (2002). Children with mild and unilateral hearing impairment. In R. Seewald & J. Gravel (Eds.), *A sound foundation through early amplification 2001: Proceedings of the Second International Conference* (pp. 179–186). Stäfa, Switzerland: Phonak Communications AG.
- Dockrell, J., & Shield, B.** (2006). Acoustical barriers in classrooms: The impact of noise on performance in the classroom. *British Educational Research Journal*, 32, 509–525.
- Downs, D.** (1982). Effects of hearing aid use on speech discrimination and listening effort. *Journal of Speech and Hearing Disorders*, 47, 189–193.
- Dunn, L. M., & Dunn, D. M.** (2007). *Peabody Picture Vocabulary Test—Fourth Edition*. Bloomington, MN: Pearson.
- Erber, N.** (1982). *Auditory training*. Washington DC: Alexander Graham Bell Association.
- Fitzpatrick, E., Durieux-Smith, A., & Whittingham, J.** (2010). Clinical practice for children with mild bilateral and unilateral hearing loss. *Ear and Hearing*, 31, 392–400.
- Fitzpatrick, E., Whittingham, J., & Durieux-Smith, A.** (2014). Mild bilateral and unilateral hearing loss in childhood: A 20-year view of hearing characteristics, and audiological practices before and after newborn hearing screening. *Ear and Hearing*, 35, 10–18.
- Feurstein, J.** (1992). Monaural versus binaural hearing: Ease of listening, word recognition, and attentional effort. *Ear and Hearing*, 30, 80–86.
- Fraser, S., Gagne, J. P., Alepins, M., & Dubois, P.** (2010). Evaluating the effort expended to understand speech in noise using a dual-task paradigm: The effects of providing visual speech cues. *Journal of Speech, Language, and Hearing Research*, 53, 18–33.
- Gatehouse, S., & Gordon, J.** (1990). Response times to speech stimuli as measures of benefit from amplification. *British Journal of Audiology*, 24, 63–68.
- Gosselin, A. P., & Gagne, J. P.** (2010). Use of a dual-task paradigm to measure listening effort. *Canadian Journal of Speech Language Pathology and Audiology*, 34, 63–68.
- Gustafson, S., McCreery, R., Hoover, B., Kopun, J., & Stelmachowicz, P.** (2014). Listening effort and perceived clarity for normal-hearing children with the use of digital noise reduction. *Ear and Hearing*, 35, 183–194.
- Hall, J., Grose, J., Buss, E., & Dev, M.** (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing*, 23, 159–165.

- Hartvig Jensen, J., Angaard Johansen, P., & Borre, S. (1989). Unilateral sensorineural hearing loss in children and auditory performance with respect to right/left ear differences. *British Journal of Audiology*, 23, 207–213.
- Haskins, H. (1949). *A phonetically balanced test of speech discrimination for children* (Unpublished master's thesis). Northwestern University, Evanston, IL.
- Hicks, C., & Tharpe, A. M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language, and Hearing Research*, 45, 573–584.
- Hillock-Dunn, A., Taylor, C., Buss, E., & Leibold, L. (2014). Assessing speech perception in children with hearing loss: What conventional clinical tools may miss. *Ear and Hearing*, 36, e57–e60.
- Hornsby, B. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear and Hearing*, 34, 523–534.
- Hornsby, B., Werfel, K., Camarata, S., & Bess, F. (2013). Subjective fatigue in children with hearing loss: Some preliminary findings. *American Journal of Audiology*, 23, 129–134.
- Houben, R., van Doorn-Bierman, M., & Dreschler, W. (2013). Using response time to speech as a measure of listening effort. *International Journal of Audiology*, 52, 753–762.
- Howard, C. S., Munro, K., & Plack, C. J. (2010). Listening effort at signal-to-noise ratios that are typical of the school classroom. *International Journal of Audiology*, 49, 928–932.
- Johnson, C. E., Stein, R., Broadway, A., & Markwalter, T. (1997). “Minimal” high-frequency hearing loss and school-age children: Speech recognition in a classroom. *Language, Speech, and Hearing Services in Schools*, 28, 77–85.
- Kahnemann, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Knecht, H. A., Nelson, P. B., Whitelaw, G. M., & Feth, L. L. (2002). Background noise levels and reverberation times in unoccupied classrooms: Predictions and measurements. *American Journal of Audiology*, 11, 65–71.
- Larsby, B., Hallgren, M., Lyxell, B., & Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: Effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *International Journal of Audiology*, 44, 131–143.
- Leibold, L., Hillock-Dunn, A., Duncan, N., Roush, P., & Buss, E. (2013). Influence of hearing loss on children's identification of spondee words in a speech-shaped noise or a two-talker masker. *Ear and Hearing*, 34, 575–584.
- Lewis, D., Valente, D. L., & Spalding, J. (2015). Effect of minimal/mild hearing loss on children's speech understanding in a simulated classroom. *Ear and Hearing*, 36, 136–144.
- Lieu, J., Tye-Murray, N., & Fu, Q. (2012). Longitudinal study of children with unilateral hearing loss. *Laryngoscope*, 122, 2088–2095.
- Lieu, J., Tye-Murray, N., Karzon, R., & Piccirillo, J. (2010). Unilateral hearing loss is associated with worse speech-language scores in children. *Pediatrics*, 125, 2009–2448.
- McFadden, B., & Pittman, A. (2008). Effect of minimal hearing loss on children's ability to multitask in quiet and in noise. *Language, Speech, and Hearing Services in Schools*, 39, 342–351.
- Oyler, R., Oyler, A., & Matkin, N. (1988). Unilateral hearing loss: Demographics and educational impact. *Language, Speech, and Hearing Services in Schools*, 19, 201–210.
- Picard, M., & Bradley, J. (2001). Revisiting speech interference in classrooms. *Audiology*, 40, 221–244.
- Porter, H., Sladen, D., Ampah, S., Rothpletz, A., & Bess, F. (2013). Developmental outcomes in early school-age children with minimal hearing loss. *American Journal of Audiology*, 22, 263–270.
- Rakerd, B., Seitz, P., & Whearty, M. (1996). Assessing the cognitive demands of speech listening for people with hearing losses. *Ear and Hearing*, 17, 97–106.
- Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, 86, 446–461.
- Ruscetta, M. N., Arjmand, E. M., & Pratt, R., Sr. (2005). Speech recognition abilities in noise for children with severe-to-profound unilateral hearing impairment. *International Journal of Pediatric Otorhinolaryngology*, 69, 771–779.
- Sarampalis, A., Kalluri, S., Edwards, B., & Hafter, E. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. *Journal of Speech, Language, and Hearing Research*, 52, 1230–1240.
- Sato, H., & Bradley, J. (2008). Evaluation of acoustical conditions for speech communication in working elementary schools. *The Journal of the Acoustical Society of America*, 123, 2064–2077.
- Shield, B., & Dockrell, J. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. *The Journal of the Acoustical Society of America*, 123, 133–144.
- Stelmachowicz, P., Lewis, D., Choi, S., & Hoover, B. (2007). Effect of stimulus bandwidth on auditory skills in normal-hearing and hearing-impaired children. *Ear and Hearing*, 26, 483–494.
- Tun, P., Benichov, J., & Wingfield, A. (2010). Response latencies in auditory sentence comprehension: Effects of linguistic versus perceptual challenge. *Psychology & Aging*, 25, 730–735.
- Walker, E., Spratford, M., Moeller, M., Oleson, J., Ou, H., Roush, P., & Jacobs, S. (2013). Predictors of hearing aid use time in children with mild-to-severe hearing loss. *Language, Speech, and Hearing Services in Schools*, 44, 73–88.
- Washington University in St. Louis. (2016). *Speech and Hearing Lab neighborhood database* [Online database]. Retrieved from <http://neighborhoodsearch.wustl.edu/Neighborhood/Home.asp>
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: Pearson.
- Westfall, P., Tobias, R. D., & Wolfinger, R. D. (2011). *Multiple comparisons and multiple tests using SAS* (2nd ed.). Cary, NC: SAS Institute.
- Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record*, 58, 475–482.
- Wightman, F., & Kistler, D. (2005). Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *The Journal of the Acoustical Society of America*, 118, 3164–3176.
- Zumbo, B., & Coulombe, D. (1997). Investigation of the robust rank-order test for non-normal populations with unequal variances: The case of reaction time. *Canadian Journal of Experimental Psychology*, 51, 139–149.

Appendix

Lexical Frequency and Density for the Phonetically Balanced Kindergarten Words on the Basis of Calculations From the Speech and Hearing Lab Neighborhood Database (Washington University in St. Louis, 2016)

Orthography	Frequency	Log Frequency	Familiarity	Density A	Mean Frequency A	Mean Log Frequency A	Density B	Mean Frequency B	Mean Log Frequency B
Set 1									
<i>please</i>	62.00	2.79	7.00	3.00	2.33	1.23	6.00	7.17	1.52
<i>great</i>	668.00	3.82	7.00	18.00	10.44	1.55	21.00	24.86	1.75
<i>sled</i>	1.00	1.00	7.00	8.00	11.38	1.78	10.00	231.30	2.20
<i>pants</i>	9.00	1.95	7.00	8.00	32.38	1.76	11.00	33.18	1.84
<i>rat</i>	6.00	1.78	7.00	31.00	399.35	2.12	37.00	480.27	2.08
<i>bad</i>	143.00	3.16	7.00	27.00	241.74	2.01	30.00	222.67	2.04
<i>pinch</i>	6.00	1.78	6.75	6.00	2.17	1.25	10.00	9.20	1.56
<i>such</i>	1303.00	4.11	6.83	11.00	279.55	2.48	12.00	256.33	2.35
<i>bus</i>	35.00	2.54	7.00	17.00	288.29	2.01	20.00	279.10	2.06
<i>need</i>	361.00	3.56	6.92	16.00	26.94	1.84	21.00	30.53	1.92
<i>way</i>	913.00	3.96	6.92	20.00	638.80	2.80	30.00	436.67	2.48
<i>five</i>	286.00	3.46	7.00	11.00	50.64	1.91	12.00	46.50	1.84
<i>mouth</i>	103.00	3.01	7.00	6.00	48.67	1.97	7.00	41.86	1.83
<i>rag</i>	10.00	2.00	7.00	24.00	12.33	1.56	28.00	11.25	1.53
<i>put</i>	437.00	3.64	7.00	14.00	20.00	1.93	14.00	20.00	1.93
<i>M</i>	289.53	2.84	6.96	14.67	137.67	1.88	17.93	142.06	1.93
Set 2									
<i>fed</i>	42.00	2.62	6.17	18.00	200.89	2.44	20.00	182.25	2.36
<i>fold</i>	7.00	1.85	7.00	10.00	121.70	2.73	12.00	156.58	2.70
<i>hunt</i>	10.00	2.00	7.00	8.00	6.38	1.53	10.00	6.60	1.57
<i>no</i>	2884.00	4.46	7.00	19.00	369.26	2.49	27.00	285.63	2.44
<i>box</i>	70.00	2.85	7.00	3.00	5.00	1.37	5.00	4.20	1.36
<i>are</i>	4459.00	4.65	6.92	6.00	991.00	3.05	19.00	375.58	2.62
<i>teach</i>	41.00	2.61	7.00	11.00	41.73	2.00	13.00	107.77	2.21
<i>slice</i>	13.00	2.11	7.00	5.00	16.20	1.79	8.00	11.13	1.62
<i>is</i>	10099.00	5.00	7.00	10.00	3967.30	3.12	13.00	3590.15	2.92
<i>tree</i>	59.00	2.77	7.00	11.00	115.73	2.12	16.00	86.69	2.10
<i>smile</i>	58.00	2.76	6.75	4.00	168.75	2.56	5.00	144.60	2.59
<i>bath</i>	26.00	2.41	7.00	17.00	122.18	2.26	17.00	122.18	2.26
<i>slip</i>	19.00	2.27	7.00	14.00	11.64	1.72	17.00	10.94	1.71
<i>ride</i>	49.00	2.69	7.00	24.00	90.08	2.22	30.00	81.10	2.22
<i>end</i>	410.00	3.61	7.00	3.00	10066.00	3.53	13.00	2338.46	2.19
<i>M</i>	1216.40	2.98	6.92	10.87	1086.25	2.33	15.00	500.26	2.19
Set 3									
<i>pink</i>	48.00	2.68	7.00	10.00	50.00	1.83	16.00	35.38	1.68
<i>thank</i>	36.00	2.56	7.00	10.00	59.70	2.03	11.00	54.36	1.94
<i>take</i>	611.00	3.79	7.00	23.00	72.13	2.04	25.00	67.72	2.04
<i>cart</i>	5.00	1.70	7.00	10.00	97.80	2.17	14.00	104.43	2.18
<i>scab</i>	1.00	1.00	7.00	3.00	5.67	1.71	4.00	7.25	1.80
<i>lay</i>	139.00	3.14	7.00	20.00	409.10	2.81	35.00	252.97	2.44
<i>class</i>	207.00	3.32	6.92	10.00	35.50	1.70	12.00	29.83	1.61
<i>me</i>	1181.00	4.07	7.00	25.00	1033.12	2.73	32.00	823.91	2.64
<i>dish</i>	16.00	2.20	7.00	12.00	110.50	2.22	12.00	110.50	2.22
<i>neck</i>	81.00	2.91	7.00	13.00	15.92	1.74	13.00	15.92	1.74
<i>beef</i>	32.00	2.51	7.00	13.00	24.77	1.92	15.00	460.00	2.18
<i>few</i>	601.00	3.78	6.92	6.00	37.67	1.69	12.00	294.83	1.80
<i>use</i>	589.00	3.77	7.00	9.00	61.22	2.15	13.00	296.08	2.19
<i>did</i>	1044.00	4.02	7.00	20.00	23.50	1.96	21.00	22.43	1.92
<i>hit</i>	115.00	3.06	7.00	29.00	391.69	2.35	33.00	609.94	2.35
<i>M</i>	313.73	2.97	6.99	14.20	161.89	2.07	17.87	212.37	2.05