Effect of Context and Hearing Loss on Time-Gated Word Recognition in Children

Dawna E. Lewis[1], **Judy Kopun**[2], **Ryan McCreery**[2], **Marc Brennan**[2], **Kanae Nishi**[1], **Evan Cordrey**[2], **Pat Stelmachowicz**[2], and **Mary Pat Moeller**[3]

[1]Boys Town National Research Hospital, Center for Hearing Research

[2]Boys Town National Research Hospital, Center for Audiology

[3]Boys Town National Research Hospital, Center for Childhood Deafness

Abstract

Objectives—The purpose of this study was to examine word recognition in children who are hard of hearing (CHH) and children with normal hearing (CNH) in response to time-gated words presented in high- vs. low-predictability sentences (HP, LP), where semantic cues were manipulated. Findings inform our understanding of how CHH combine cognitive-linguistic and acoustic-phonetic cues to support spoken word recognition. It was hypothesized that both groups of children would be able to make use of linguistic cues provided by HP sentences to support word recognition. CHH were expected to require greater acoustic information (more gates) than CNH to correctly identify words in the LP condition. In addition, it was hypothesized that error patterns would differ across groups.

Design—Sixteen CHH with mild-to-moderate hearing loss and 16 age-matched CNH participated (5–12 yrs). Test stimuli included 15 LP and 15 HP age-appropriate sentences. The final word of each sentence was divided into segments and recombined with the sentence frame to create series of sentences in which the final word was progressively longer by the gated increments. Stimuli were presented monaurally through headphones and children were asked to identify the target word at each successive gate. They also were asked to rate their confidence in their word choice using a 5- or 3-point scale. For CHH, the signals were processed through a hearing aid simulator. Standardized language measures were used to assess the contribution of linguistic skills.

Results—Analysis of language measures revealed that the CNH and CHH performed within the average range on language abilities. Both groups correctly recognized a significantly higher percentage of words in the HP condition than in the LP condition. Although CHH performed comparably to CNH in terms of successfully recognizing the majority of words, differences were observed in the amount of acoustic-phonetic information needed to achieve accurate word

recognition. CHH needed more gates than CNH to identify words in the LP condition. CNH were significantly lower in rating their confidence in the LP condition than in the HP condition. CHH, however, were not significantly different in confidence between the conditions. Error patterns for incorrect word responses across gates and predictability varied depending on hearing status.

Conclusions—The results of this study suggest that CHH with age-appropriate language abilities took advantage of context cues in the HP sentences to guide word recognition in a manner similar to CNH. However, in the LP condition, they required more acoustic information (more gates) than CNH for word recognition. Differences in the structure of incorrect word responses and their nomination patterns across gates for CHH compared to their peers with normal hearing suggest variations in how these groups use limited acoustic information to select word candidates.

Keywords

hearing loss; children; time-gating; word recognition; error analysis

INTRODUCTION

Spoken word recognition is a complex and flexible process that relies on interplay between acoustic-phonetic stimulus cues and the listener's linguistic and cognitive processing abilities. When presented with speech, listeners must decode the acoustic-phonetic characteristics of the signal and link them to existing lexical representations in long-term memory. Efficient and accurate word recognition is supported by a coordinated system that relies on various mechanisms, such as auditory perception, working memory, phonological representations in long-term memory, and linguistic knowledge (Munson 2001). For children who are hard of hearing (CHH), subtle weaknesses in any of these mechanisms could interfere with the efficiency and accuracy of word recognition. Hearing loss during development can negatively affect linguistic knowledge (Tomblin et al. 2015) and working memory abilities (Pisoni & Cleary, 2003; but see also Stiles, McGregor & Bentler 2012), both of which are important skills that support word recognition. CHH with better vocabulary and cognitive skills have superior word recognition abilities in quiet and in noise compared to peers with more limited skills in these domains (McCreery et al. 2015). The specific ways in which CHH use the combination of contextual cues and cognitive and linguistic skills to reconstruct the message when information is incomplete or degraded are not well understood. The current study seeks to gain insights into children's word recognition processes in response to partial acoustic-phonetic information presented with and without semantic context cues.

In a review of spoken word recognition theories, Jusczyk and Luce (2002) concluded that most models assume that perception of spoken words involves processes of activation and competition. That is, listeners respond to speech input by bringing up a set of similar-sounding candidates in memory (activation), and the set of candidates is subsequently narrowed through processes of discrimination and competition among them. In adults, word onsets are used to determine the initial set of candidates, with later acoustic-phonetic information being used to narrow the candidate words (Marslen-Wilson & Tyler 1980; Grosjean 1985; Salasoo & Pisoni 1985; Jusczyk & Luce 2002).

Developmental evidence and evolving theories have led some to question whether word-initial input plays the same prominent role for children as it does for adults. Considerable evidence suggests that early speech representations are not organized in regard to phonemes; rather young children process words in a holistic manner. This means that when the recognition system is immature, children give equal weight to acoustic-phonetic information throughout the word, rather than differentially attending to word-initial cues (Walley 1988). As children's vocabularies grow during early and middle childhood, their representations undergo restructuring and/or become increasingly detailed as a result of the need to differentiate among an increasing number of competitor entries in long-term memory (Walley 1988; Charles-Luce & Luce 1995; Nittrouer 1996; Metsala 1997; Garlock et al. 2001). This viewpoint predicts developmental changes in the processing of spoken words including shifts between early childhood and later stages in the ways children divide their perceptual attention.

The relative role of word-onset in triggering lexical searching has been examined in past studies using time-gated words to gain insights about listeners' real-time lexical searching and word-recognition processes (Grosjean 1980, 1985; Walley et al. 1995). A forward, sequential gating paradigm presents listeners with portions of words that are repeated in increasing portions (gates) beginning at the onset of the word and continuing until the entire word is available (Grosjean 1980; Craig 1992). For example, listeners hear partial input at successive gates (de, des, desk), and the task is to guess the target word based on available cues and self-rate confidence that they correctly identified the target word after each incremental gate. The ability to recognize a gated word, especially at early gates, is dependent on the listener's ability to access a lexical item based on this partial information (Munson 2001). It has been suggested that the gating paradigm shares properties with the task of recognizing words in noisy settings, where portions of words are obscured or distorted (Elliott et al. 1990). As such, the gating paradigm has potential for providing clinically-relevant insights about lexical processing and word recognition in CHH.

Another central research concern has been the interplay between listeners' reliance on linguistic knowledge and higher-level cognitive abilities compared to acoustic-phonetic cues in the stimulus. Previous studies have examined the integration of acoustic-phonetic and semantic and syntactic cues by manipulating the contextual information provided (Kalikow et al. 1977; Salasoo & Pisoni 1985; Craig et al. 1993). The presence of semantic and syntactic context was found to accelerate the process of word recognition by allowing listeners to rely on linguistic knowledge and cognitive processes to reduce the initial set of lexical competitors (Grosjean 1980; Tyler & Wessels 1983; Salasoo & Pisoni 1985). Salasoo and Pisoni (1985) conducted error analyses to further understand how listeners combine acoustic-phonetic input and abstract linguistic knowledge to recognize words. They postulated that the structure of incorrect word nominations in a gating task would contain information about the basis of the listeners' guesses. Three possible sources were proposed, including errors that reflected (1) analysis of acoustic-phonetic cues in the signal (guesses shared acoustic-phonetic properties with the target), (2) relying on semantic or syntactic context (guesses predicted from context), or (3) using other sources (e.g., non-words, intrusions from other stimuli presented in the task). They found that adult listeners were heavily reliant on acoustic-phonetic cues in the stimulus in low-predictability contexts and

that the acoustic-phonetic information at the beginnings of words supported word recognition and lexical access. However, in high-predictability contexts, the error responses showed that adult listeners used linguistic knowledge to identify word candidates, even when the answer did not match the acoustic-phonetic properties of the stimulus. This type of error analysis has not been conducted with CHH, and may inform our understanding of lexical processing by this group.

Craig and colleagues (1993) examined listeners' reliance on acoustic-phonetic cues versus higher level cognitive-linguistic processes by presenting gated words in high- and low-predictability contexts to groups of children and adults. High-predictability sentences contain semantic cues to the final word (e.g., *The cow was milked in the **barn***) while low-predictability sentences minimize such cues (e.g., *She talked about the **spice***). Listeners were asked to rate their confidence on a Likert scale after each response, which is a common practice in gating studies. Time-gated words were recognized earlier and with more confidence in the high-predictability condition compared to the low-predictability sentences by both children and adults. Developmental effects were observed with older children (8 – 10 years) showing more effective use of high-predictability contexts than younger children (5–7 years).

Other previous studies using the gating paradigm applied to children found that (1) young children (age 5) required more gates than adults, suggesting that they needed more word-initial input than adults for correct word identification in noise (Walley 1988), (2) young children (ages 5 – 7 years) performed like older adults but more poorly than teenagers in recognizing time-gated words (Elliott et al. 1987), (3) children (preschoolers through second grade) performed best in identifying gated words that were acquired early in development, showing age of vocabulary acquisition effects on word recognition (Garlock et al. 2001), and (4) receptive vocabulary explained individual differences in word recognition in children with phonological disorders or learning disabilities (Elliott et al. 1990; Edwards et al. 2002).

Listeners may demonstrate patterns of response perseveration or change in the gating paradigm. Craig and Kim (1990) observed patterns of perseveration in adult listeners. Once adult listeners made incorrect guesses, they tended to persevere with the same response even after the word offset. In a follow-up study, children were observed to use beginning sounds to formulate initial responses, but if their guesses were not supported by the subsequent acoustic-phonetic information, they had a tendency to change their responses (Craig et al. 1993). Results suggest developmental trends in error response patterns which are not well understood. Further research is needed to explore the mechanisms underlying these differences. Walley and colleagues (1995) noted that successive presentations could contain implicit negative feedback, perhaps indicating to children that they are on the wrong track with their current answers. It is possible that CHH are particularly susceptible to implicit negative feedback which prompts them to change their answers more often than CNH, given that others often repeat words for them when they have not understood. Analysis of various error patterns may provide insight into knowledge sources the respective groups of children rely on as they progress toward word recognition.

As noted by Grosjean (1980), a listener may correctly identify a gated word without being confident in his/her response, continuing to monitor the acoustic-phonetic signal beyond the point of identification before reaching a maximum level of confidence. Even after CHH correctly identify a gated word, reduced audibility could result in a need to attend more closely to the increasing acoustic-phonetic cues in the gated signal relative to CNH before the child is confident about having perceived the word correctly. Confidence in one's recognition may play a role in ongoing speech perception.

The current investigation explored the effects of semantic context (high vs. low predictability, HP and LP, respectively) on time-gated word recognition in age-matched children, who varied in hearing status (CNH, CHH). The presentation of gated words in these separate contexts provided the opportunity to examine the relative contributions of semantic and acoustic-phonetic cues to children's word recognition. It is possible that CHH might have age-appropriate language abilities, yet still differ from CNH in using acoustic-phonetic cues for word recognition, due to between-group differences in audibility. Speech is fully audible for CNH; in contrast, hearing aids even when optimized for audibility, do not give CHH complete audibility of speech. Under this prediction, CHH would perform comparably to CNH when they can effectively use context to compensate for reduced access to acoustic-phonetic cues (HP condition), but not as well as CNH when such contextual cues are limited and access to robust acoustic-phonetic cues is required (LP condition). Although we did not prospectively match the groups on language ability, both groups scored within the average range on two standardized measures of expressive language. This affords a unique opportunity to examine between-group differences in use of semantic vs. acoustic-phonetic cues to support word recognition, with control of a potentially confounding variable – child language. Knowledge sources used by the groups were further examined through analysis of incorrect word nominations.

This study addresses the following specific research questions:

1. Are words in HP sentences recognized earlier and more successfully than words in LP sentences by children from both groups? Given that the respective groups have comparable language abilities, we predict that both will take advantage of the availability of linguistic (semantic context) cues to achieve higher recognition levels in the HP than in the LP condition.

2. Do CHH require more gates than CNH to accurately recognize time-gated words presented in low-predictability contexts? We predict that CHH will require more acoustic-phonetic information than their peers to achieve correct word identification in the absence of context support (LP condition).

3. Are the error patterns and confidence ratings of CHH distinct from those of CNH? We predict that the CHH will be less able to rely on acoustic-phonetic cues in the signal than CNH in the LP condition leading to different error patterns, and this will also result in CHH providing lower confidence ratings in the LP condition than CNH.

MATERIALS AND METHODS

Participants

Seventeen children ages 5–12 years with mild-to-moderate hearing loss and 17 children ages 5–12 years with normal hearing enrolled in this study. All participants were monolingual speakers of English. The CNH were matched in age to within 6 months of the CHH. One CHH participant demonstrated an error rate for the experimental task that was > 2 SD above the group mean, and therefore was identified as an outlier. This participant's data were removed from analysis, along with those of the age-matched peer with normal hearing, bringing the totals in each group to 16 participants. Table 1 summarizes demographic characteristics of both groups. For CNH, hearing was screened at 15 dB HL for octave frequencies from 250–8000 Hz in both ears. For CHH, behavioral thresholds were obtained at octave frequencies from 250–8000 Hz. In addition, thresholds were obtained at inter-octave frequencies when the difference between thresholds at consecutive octaves was greater than 15 dB. Figure 1 shows the mean and range of hearing thresholds for the CHH and Table 2 summarizes audiological factors for this group. The Speech Intelligibility Index (SII) values in the table are based upon the child's personal hearing aid(s) at use settings as measured with the Audioscan Verifit (Dorchester, Ontario). Fourteen of those participants wore binaural hearing aids, one wore binaural hearing aids at school and a monaural hearing aid at home, and one did not wear hearing aids at the time of testing[1].

This study was approved by the Institutional Review Board of Boys Town National Research Hospital (BTNRH). Written assent for children over 6 years old and parental-consent to participate were obtained for each child and they were compensated monetarily for participation.

Stimuli

Sentence Construction

HP Sentences: Sixty-eight high-predictability (HP) sentences from the Speech Perception in Noise (SPIN) test sentences (Kalikow et al. 1977) were used. In HP SPIN sentences the final (target) word is predictable based on the semantic context offered by the sentence. The 68 HP sentences were selected from the original set of 200 HP SPIN sentences if they did not have embedded clauses, if all the words were within the vocabulary of a first grader (Storkel & Hoover 2010), and if the target word could be easily excised from the preceding word. This excluded sentences where the initial phoneme of the target word was a sonorant, /h/, or where the final phoneme of the preceding word and the initial phoneme of the target word were the same. Sentences in which the target word ended in a vowel (CV-consonant vowel) were eliminated, so that only CVC, CCVC, CVCC, and CCCVC words were included. Some uncommon proper nouns were changed to more common nouns.

In order to avoid sentences where the target word was too easy or too difficult for children to predict, the 68 HP SPIN sentences were printed without the expected last word and 15

---

[1]The participant who did not wear hearing aids had been identified with hearing loss at birth and fit with hearing aids at the age of three years. However, this participant was no longer using the hearing aids when enrolled in this study.

children ages 5, 8 and 12 years old (5 per age category) were asked to fill in the blank. The sentences were read to the 5-year-old participants and their responses were transcribed by the examiner. The 8- and 12-year olds read the sentences themselves and wrote their own responses. The frequency with which the expected target word was selected as the response was calculated for each sentence. Of the 68 HP sentences, the 29 sentences in which the target word was guessed correctly by between 3 and 10 of the 15 participants were selected.

LP Sentences: In low-predictability (LP) SPIN sentences the target word is not predictable based on the semantic context of the sentence, however the sentences are syntactically appropriate. In order to address potential learning effects resulting from use of a sequential gating paradigm, different target words were selected for the LP than those used for the HP sentences. Target words in the LP sentences were chosen to match those in the HP sentences based on duration, syllable structure, consonant manner, and voicing but not on place. These LP target words were placed in modified LP SPIN sentence frames. The original LP sentences contained some words that were not in the vocabulary of first graders, so these words were modified ("considered" was changed to "thought about", "discussed" was changed to "talked about").

Stimulus Recording—The sentences were recorded by a 31-year-old female talker from the Midwest. The talker was seated in a sound booth and was instructed to read the sentences using normal vocal effort. The audio was recorded by a detachable camera shotgun microphone (GY-HM710U JVCPROHD) mounted on a Manfrotto tripod. The recorded file was imported into Final Cut Pro 7 (Apple, Inc., Cupertino, CA) and divided into individual sentence files, which were exported as individual .wav files. Two recordings were made of each sentence. Each of the recordings was evaluated by an experienced listener on the research team and the best quality recording (no noise or interruptions in the signal; clarity of the target word) was selected for the final set of stimuli. Nine HP and nine LP sentences were eliminated because of poor recording quality in either the matching HP or LP sentence, leaving 20 each of HP and LP sentences. Two HP sentences and three LP sentences with non-matching target words were selected for practice. The final set of stimuli included 15 HP and 15 LP sentences (See Supplemental Digital Content 1).

Creation of Gates—The stimuli were gated using PRAAT (Ver. 5.3.51; Boersma & Weenink 2013) software with a script developed by the 5th author. After the beginning of the target word was marked, the target word was excised from the sentence frame at the nearest zero-crossing. Depending on the duration, the target word was cut into 5 to 13 segments. The first segment of each target word was 150 ms in duration and each following segment was 50 ms. Inclusion of a longer first gate followed procedures recommended by Elliott et al. (1987) for use with young children. If the duration of the final gate was > 25 ms, it was left unchanged. If the duration of the final gate was < 25 ms, it was attached to the previous gate. To minimize distortion, each segment was cut at the nearest zero-crossing and ramps were applied to the gates. The gates were then recombined with the sentence frame to create a series of sentences with the final word progressively longer by the gated increments. The mean number of gates for HP sentences was 7.0 (standard deviation [SD] = 2.2) and the mean duration was 481.5 ms (SD = 108 ms). The mean number of gates for LP sentences

was 6.9 (SD = 1.4) and the mean duration was 465.6 ms (SD = 103 ms). The last two presentations of each sentence (final 2 gates) included a complete final word. The additional repetition of the complete target word was used in order to determine if the participant would make any additional changes to their response, such as changing the target word to a plural.

Amplification

A hearing aid simulator program in MATLAB (Mathworks, Natick, MA) (see McCreery et al. 2014, Alexander & Masterson 2015) was used to provide amplification of the stimuli for the CHH in order to avoid inconsistencies in audibility provided by personal amplification. The simulated hearing aid processed the stimuli with a sampling rate of 22.05 kHz using broadband input limiting, filter bank, wide-dynamic range compression, multichannel output compression, and a broadband output limiter. The attack and release times for the wide-dynamic range compressor were 5 and 50 ms, respectively. For a full description of the hearing aid simulator see the description of the extended bandwidth condition in Brennan et al. (2014).

The output of the hearing aid simulator for a 60 dB SPL speech-weighted input (carrot passage, Audioscan Verifit, Dorchester, Ontario) and 90 dB SPL swept pure tone were set to the Desired Sensation Level (DSL) algorithm v. 5.0s (Scollie et al. 2005), pediatric version. Because the DSL algorithm does not provide a target for 8 kHz, the target level at that frequency was set to provide the same sensation level as that prescribed at 6 kHz. Fit-to-target was computed, as described in Brennan et al. (2014), by subtracting the output level of the hearing aid simulator for the carrot passage, 60 dB SPL input level, for each frequency band using 1/3 octave-wide filters (ANSI 2004) from the DSL target level. The resultant values are illustrated in Fig 2. Fit-to-target differences were calculated for the test ear for all participants except for one participant for whom stimuli were inadvertently presented binaurally. For that participant, the output of the hearing aid simulator and hearing aid targets were averaged for the right and left ears prior to inclusion in the group data. For most participants, differences between the simulator output and DSL targets were less than 5 dB.

Procedures

Speech/Language Measures—CHH wore their personal amplification for the speech/language section of the study. In order to ensure that participants' responses were intelligible, an articulation probe was developed that was specific to the stimuli used in this study. Participants were seated in a sound booth and asked to identify pictures of eight objects whose name included phonemes present in the stimuli that might be difficult for participants to say. The targets of the probe were /l/ (be*ll*), /s/ (bu*s*), /ts/ (ca*ts*), /ʃ/ (fi*sh*), /ɹk/ (fo*rk*), /sk/ (ma*sk*), /dʒ/ (oran*ge*), and, /st/ (toa*st*). All identified errors were judged by the examiner as slight to moderate distortions that did not interfere with scoring. Consequently, no participants were excluded because of articulation errors.

To assess the contribution of linguistic skills, the Expressive Vocabulary Test, Second Edition (EVT-2) form A (Williams 2007) and the Comprehensive Assessment of Spoken Language (CASL), Sentence Completion subtest (Carrow-Woolfolk 1999) were

administered live voice. The Sentence Completion subtest of the CASL was selected because it examines word retrieval based on the context of the sentence. The examiner reads a sentence that is missing the final word. The participant is asked to provide a word to meaningfully complete the sentence. Parents were also asked to fill out a questionnaire that included information about their child's hearing and amplification history, current amplification use, educational placement, and maternal level of education.

Gating Task—In each of the gating task trials, participants verbally reported their guess for the target word, then reported how confident they were that their response was correct using a 3- or 5-point scale. The confidence scale contained text and a visual analogy that corresponded to a range of confidence ratings. The visual analogy was a series of facial expressions adapted from a confidence scale developed by Roebers (2002) and Roebers et al. (2004). The facial expression varied from a sad face to a happy face by gradually changing the shape of the mouth. The text associated with the faces was as follows: "I don't know" (1), "I'm not sure" (2), "maybe" (3), "I'm pretty sure" (4), and "I'm very sure" (5). Pilot data with normal hearing 5–12 year olds indicated that the youngest children (5–6 years) tended to use the extremes of the 5-point scale, so the number of confidence intervals was reduced to three (1 – 3 – 5) for 5–7 year olds ("I'm not sure" and "I'm pretty sure" were eliminated). Thought bubbles were added above "I don't know" (1) and "I'm very sure" (5) for further clarification. Subsequent 5–6 year-old NH pilot participants were able to use all points on the 3-point confidence scale. The 3-point scale used for the 5–7 year olds represented the same values and descriptions/labels as 1, 3, and 5 for the older children.

Stimuli were presented in a sound-treated booth monaurally through Sennheiser HD 25-1 II headphones. A locally-developed software program presented stimuli, displayed the confidence scale, and recorded confidence ratings using a touchscreen monitor. The presentation level was either 60 dB SPL at the headphone (CNH) or 60 dB SPL to the input of the hearing-aid simulator (CHH). The test ear was randomly selected for the CNH and the ear with the best pure-tone average was selected for the CHH. The exception was the one child who was hard of hearing who inadvertently received binaural presentation of the stimuli.

All children completed training and practice prior to the experimental task (see Supplemental Digital Content 2 for instructions). After the training and practice phase, the gating stimuli were presented sequentially until the participant responded with the correct word and selected the highest confidence rating on two consecutive gates or until all the gates were presented. Once a participant responded with the correct word and selected the highest confidence rating on two consecutive gates, the next sentence series was initiated. The predictability (HP or LP) of the sentence series was alternated for presentation to each participant, with random presentation order within each predictability category. Due to the attention demands of the task, 5–7 year olds completed the 30 sentences over the course of two visits (15 sentences per visit), while 8–12 year olds completed all 30 sentences in a single visit. Breaks were provided as necessary in order to maintain participants' attention. If a second visit was required, an LP practice sentence was presented prior to data collection. Participants' responses were video recorded and all incorrect responses (real words and nonsense words) were transcribed at a later date. In 8 of 4692 trials, the computer program

failed to record the confidence rating or response (correct vs. incorrect) for a single gate. In those cases, the confidence rating was assigned the average of the trial before and after the missing trial. The response from the previous trial (correct or incorrect) was maintained.

For the current study, the isolation point was defined as the stimulus duration (number of gates) at which the child correctly identified the target and did not change the response in subsequent trials. The acceptance point was the duration at which the child accurately identified the target on two successive trials and gave a confidence rating of 5, indicating high confidence. Thus, although the child's response is correct beginning at the isolation point, the acceptance point represents the gate at which the child is confident in that word choice.

Data Analyses

Isolation and acceptance points were used when calculating children's confidence values for the analysis. The following dependent variables were computed and compared across conditions and groups, (1) percent of words for which children reached the isolation point (Grosjean 1980; Walley 1988; Craig et al. 1993), (2) confidence ratings at the isolation point, (3) percent of words on which children reached the acceptance point (Walley 1988; Craig et al. 1993), (4) percent of gates required to reach the isolation and acceptance points, and (5) analysis of incorrect word nominations for the respective groups and conditions. For 24 of 480 full-sentence presentations, there were instances for which individual children did not recognize the target word, even at the longest gate (12 for CHH, 12 for CNH). The procedure described by Elliott et al. (1987) was used to separate the targets recognized at the longest gate from those not recognized at the longest gate. Using this procedure, the duration for the unrecognized items was defined as the stimulus duration plus the duration of one additional gate. The same procedure was used for instances where participants did not reach the acceptance point at the longest gate.

Coding of Incorrect Words and Error Patterns—To examine the nature and distributions of incorrect word candidates across the two groups, error responses were coded using a scheme reported by Salasoo and Pisoni (1985). Their scheme was designed to categorize incorrect responses into three mutually exclusive bins (acoustic-phonetic, grammatical, other). These original categories were intended to capture the hypothesized sources of knowledge underlying young adults' error responses. Preliminary analysis of error patterns in the current study revealed that the children were making vowel errors in early gates, and there was interest in capturing the extent to which these errors occurred for both groups. Therefore, all incorrect word nominations in the first five gates[2] of the LP condition were coded into one of four mutually-exclusive categories by the final author (See Supplemental Digital Content 3 for definitions), (1) acoustic-phonetic analysis of the signal [Acoustic-Phonetic], (2) acoustic-phonetic analysis, but the vowel did not match the target [Vowel], (3) language cues providing context to support the response [Language], and (4) other sources [Other]. The final author coded all error responses, and a trained research

---

[2]The majority of error responses occurred within the first five gates. Few errors occurred after this point, especially for CNH, limiting quantitative comparisons past gate 5.

assistant who was blinded to hearing status of the participants also coded 20% of the error responses for each condition and group across the first five gates to establish inter-judge reliability. Cohen's Kappa was .926, indicating strong agreement.

A secondary error pattern analysis was conducted for the LP condition to determine the degree to which children provided persisting or changing error responses. Error responses occurring at two or more gates were coded into two mutually exclusive categories, as follows.

1. *Persistent*. The child repeated the same word across two or more successive gates, or showed multiple successive repetitions starting at gate 2. For example (target word in brackets), [brick] = bread, bread, bread, bread; [desk] = deck, dust, dust, dust, dust.

2. *Changing*. The child changed the response at each gate, or made multiple changes in nominated words, mixed with some repetitions. For example: [spice] = splash, spin, spider, spike, spite; [flash] = football, flood, flower, flower, flat, flat.

A trained research assistant who was blinded to hearing status of the participants also coded 100% of the response patterns to establish inter-judge reliability. Cohen's Kappa was .936, indicating strong agreement.

RESULTS

Descriptive results from the language measures are summarized in Table 1. Independent sample *t* tests showed no significant differences between groups (p > .05) indicating that the groups were matched on language abilities; both groups performed within the average range relative to the test standardization norms on the vocabulary and sentence completion measures. Mean standard scores were higher than the mean value for the normative sample (100) on both language measures for both groups. Preliminary analyses of gating task performance revealed no effects of age across groups on any analyses (p > .05). For that reason the groups are aggregated across age for reporting of results.

Spoken Word Recognition: High- vs Low-Predictability Conditions

The first research question examined the effects of predictability on time-gated word recognition in CNH and CHH by comparing their performance in response to HP vs LP sentence conditions. Table 3 shows the percentage of words for which children in each group achieved the isolation point and acceptance point as a function of predictability condition. Separate mixed-design analyses of variances (ANOVA) were completed, with predictability (HP vs. LP) and group (CNH, CHH) as factors for each dependent variable. In regard to the isolation point, there was a main effect of predictability [$F(1,30) = 15.348$; $p < 0.001$; $\eta_p^2 = 0.338$], but not of group [$F(1,30) = .113$; $p = .739$; $\eta_p^2 = .004$], and no significant interaction between group and predictability [$F(1,30) = .313$; $p = .580$; $\eta_p^2 = .010$]. Evaluation of the marginal means for predictability revealed significantly higher word recognition for the words in the HP compared to the LP context. The lack of a group difference suggests that, in spite of differences in hearing status, children matched on

language abilities took advantage of linguistic cues available in the HP condition to achieve more accurate word recognition than in the LP condition. In regard to the acceptance point, results indicated effects that are comparable to the isolation point findings: There was a significant effect of predictability [$F(1,24) = 15.53$; $p < 0.001$; $\eta_p^2 = .393$], but no effect of group [$F(1,24) = .024$, $p = .878$, $\eta_p^2 = .001$], and no significant interaction between group and predictability [$F(1,24) = .0006$; $p = .941$, $\eta_p^2 = .000$]. Evaluation of the marginal means for predictability revealed that children reached the acceptance point for a significantly higher percentage of words in the HP condition compared to the LP condition. For both the isolation and acceptance points, the results suggest that regardless of hearing status, the availability of language cues facilitated children's word recognition performance.

Confidence Ratings

Four children (2 NH and 2 HH) in the 5–7 year age range demonstrated that they did not understand the confidence scale by always selecting "Very Sure or "I Don't Know" and their confidence ratings were excluded from that portion of the analysis. Confidence-rating data were non-normally distributed, so nonparametric statistical tests were used in this analysis. In the HP condition, the median confidence rating at the isolation point was 4.8 out of 5 (interquartile range [IQR] = 0.27) for the CNH and 4.7 (IQR = 0.8) for the CHH, indicating high confidence in both groups. In the LP condition, the median confidence rating at the isolation point was 4.2 (IQR = 1.73) for the CNH and 4.4 (IQR = 0.8) for the CHH. To determine whether confidence ratings differed between groups, independent sample Kruskal Wallis tests were conducted and showed that the groups did not differ from each other in their overall confidence ratings in the HP ($p = .282$) or the LP ($p = .898$) conditions. To determine whether confidence ratings differed by context within each group, related-samples Wilcoxon Signed Ranks tests were conducted separately for the CNH and the CHH. CHH gave very similar confidence ratings whether they were recognizing words in the HP or LP condition ($p = .069$). However, for the CNH, confidence ratings were significantly lower with greater range in the LP compared to the HP condition ($p = .002$).

Percent of Gates Required to Reach the Isolation Point

The second research question sought to determine whether there were group differences in the percent of gates required to achieve the isolation point. Figure 3 displays the percent of gates required by each group for the LP and HP conditions. A mixed-design ANOVA was conducted to determine if the percent of gates to reach the isolation point differed by predictability condition (HP vs. LP) or group (CNH, CHH). Results are summarized in Table 4. The significant main effect of predictability [$F(1,30) = 825.61$, $p < .001$, $\eta_p^2 = .965$] indicates that, pooled across groups, the percent of gates required to achieve the isolation point was lower in the HP condition compared to the LP condition. There was also a significant main effect of group [$F(1,30)$, $p < .011$, $\eta_p^2 = .196$], indicating that the CHH required a significantly higher percentage of gates than the CNH to reach the isolation point. However, results also showed a significant interaction between predictability condition and group [$F(1,30) = 12.253$, $p < .001$, $\eta_p^2 = .290$]. Post hoc tests (Fisher's Least Significant Difference) comparing groups in each predictability condition suggested that the isolation point for CHH (20.35%) was not different from the CNH (19.44%) in the HP condition, but CHH required significantly more gates (56.89%) than the CNH (48.04%) in the LP

condition. This suggests that both groups used language cues in a similar manner to support word recognition in the HP condition. However, in the LP condition, CHH required more acoustic information than CNH to reach the isolation point, suggesting they were less effective in using acoustic-phonetic cues.

Percent of Gates Required to Reach the Acceptance Point

Potential group differences in the percent of gates required to achieve the acceptance point for the 28 children who were able to perform the confidence rating task also were examined. A mixed-design ANOVA was conducted to determine if the percent of gates to reach the acceptance point differed by predictability condition (HP vs. LP) or group (CNH, CHH). Results are summarized in Table 5. The significant main effect of predictability [$F(1,26) = 159.03$, $p < .001$, $\eta_p^2 = .856$] indicates that, pooled across groups, the percent of gates required to achieve the acceptance point was lower in the HP condition compared to the LP condition. The main effect of group was not significant ($F[1,26] = 0.203$ $p = .656$, $\eta_p^2 = .008$), indicating that the CNH and CHH did not differ in percentage of gates needed to reach the acceptance point. There also was no interaction between predictability condition and group [$F(1,26) = 0.66$, $p = .799$, $\eta_p^2 = .003$]. The acceptance point incorporates both accuracy and confidence. Although CNH required fewer gates to reach the isolation point (a measure of accuracy) than CHH in the LP condition, they required more gates beyond that point to reach the acceptance point. The absence of a difference in acceptance points for the two groups suggests that both groups needed similar amounts of information from the speech signal to be maximally confident.

Analysis of Error Responses

The next research question examined patterns of incorrect responses provided by the CNH and CHH in the first five gates for both conditions (HP and LP) and is shown in Figure 4. A mixed-design ANOVA was conducted with the four error types (Acoustic-Phonetic, Vowel, Language, and Other) examined as a function of the five gates (1 – 5) and two levels of predictability (HP, LP) for the CNH and the CHH. The degrees of freedom were adjusted using the Greenhouse-Geisser estimates of sphericity, in cases where Mauchly's test indicated that the assumption of sphericity was violated. Results are summarized in Table 6. To analyze the pattern of significant differences across main effects and interactions while controlling for Type I error inflation due to multiple comparisons, post hoc tests using Fisher's Least Significant Difference were conducted with a minimum mean significant difference of 1.18 errors. CHH had more errors than CNH and there were more errors in the LP condition than in the HP condition. The Language error types were higher than other error types in the HP condition, whereas Acoustic-Phonetic and Vowel categories were higher than other error types in the LP condition. The number of errors across all types decreased with each successive gate, but the differences in error types across gates were not significant after controlling for multiple comparisons. The significant four-way interaction indicated that the pattern of vowel errors across gates and predictability depended on whether children had normal hearing or were hard of hearing. In the HP condition, the number of Vowel errors was low and the CNH did not make any Vowel errors in the first three gates, whereas the CHH had Vowel errors at those gates. For the Vowel errors in LP conditions, both groups made more errors than in the HP condition and Vowel errors

decreased across gates. None of the other two- or three-way interactions had significant differences after controlling for multiple comparisons.

In summary, the follow up analysis shows a similar pattern of Language errors that decreased across gates for both groups in the HP condition, but a greater number of Vowel errors for CHH than CNH for the first three gates in the HP condition. A final error analysis examined whether the participants persistently nominated the same response for a target or they kept changing responses in the LP condition. Between group differences are illustrated in Figure 5. A mixed-design ANOVA was conducted with pattern as the within-subjects factor (persist, change) and group (CHH, CNH) as the between-subjects factor. Results revealed no significant main effect for pattern [$F(1,30) = 3.507$, $p = .071$, $\eta_p^2 = .105$]. There was a significant effect for group [$F(1,30) = 4.449$, $p = .043$, $\eta_p^2 = .129$] and a significant interaction between group and pattern [$F(1,30) = 7447$, $p = .011$, $\eta_p^2 = .199$]. CHH changed their responses significantly more often than CNH and persisted with the same response significantly less often.

DISCUSSION

The first purpose of this investigation was to examine the effect of sentence context (HP vs. LP) on time-gated word recognition in age-matched groups of CNH and CHH, who happened to have comparable language abilities. As predicted, children in both groups correctly recognized a significantly higher percentage of words in the HP condition than in the LP condition. This finding is consistent with previous reports suggesting that the availability of semantic context facilitates the perception of words by reducing the initial set of word competitors in listeners of varied ages (Grosjean 1980; Salasoo & Pisoni 1985; McAllister 1988; Craig et al. 1993). Children in both groups in the current study made fewer errors in the HP condition than in the LP condition, and reached the isolation point in the HP condition at less than half the gates required for the LP condition. These results provide further support for the facilitative effects of linguistic context in the efficiency of spoken word processing.

Our findings indicate that the CHH performed comparably to CNH in successfully recognizing most words presented in the time-gated paradigm. Both groups performed near ceiling at the isolation point (> 94% of the words) and at the acceptance point (> 99%) in the HP condition. Similarly robust scores were observed for the isolation point in the LP condition (> 94%). Although acceptance point average scores were lower (84%) in the LP condition compared to the HP condition for both groups, findings reflected fairly strong word recognition and high confidence. Two reasons may account for these high performance levels across groups and conditions. The first is that language skills (vocabulary, sentence completion) were similar between the groups, consistent with their ability to use the available semantic cues in the HP sentences in the service of word recognition. Secondly, the words employed in the current study were controlled to be within the vocabularies of typical first graders (Storkel & Hoover 2010), suggesting that the target words should have been familiar to the majority of participants in the study. However, previous research employing the gating paradigm with children (Garlock et al. 2001) found that later-acquired words require more gates. The gating paradigm in the current study might have been insensitive to

group differences associated with word acquisition, because words were selected to be in the vocabularies of young children. It may be premature to draw such a conclusion. The fact that the groups are comparable in language abilities appears to account for their equivalent success in overall word recognition, and this impression is supported by the evidence that language skills are predictive of children's speech recognition abilities (Elliott et al. 1990; Edwards et al. 2002; McCreery et al. 2015). However, it is possible that sensitivity of the gating paradigm could be increased by including later-developing or low frequency words. Further research is needed to confirm this hypothesis. Although both groups reached the isolation point for the majority of words, there were differences in the percent of gates needed by the CNH and CHH to achieve that point. In the HP condition, both groups required only 20% of the gates on average to reach the isolation point, suggesting that they were able to rely on knowledge/semantic cues to support the limited acoustic-phonetic cues present in the signal. However, a different pattern was seen in the LP condition. The CHH needed more than half of the gates (57%) on average to reach the isolation point, whereas the CNH required less than half of the gates (48%) to reach this point. Although the effect size is modest (.196), this finding may reflect inherent differences in audibility between the groups, resulting in the CHH needing more gates in order to have equivalent acoustic-phonetic information to the CNH. Another contributing factor to differences in performance on LP sentences could be related to differences in working memory abilities between CNH and CHH. Some studies have shown minimal differences in working memory abilities between CNH and CHH (Stiles, McGregor & Bentler 2012), whereas other studies have indicated that working memory may be an area of developmental risk for CHH (e.g. Pisoni & Cleary 2003; Conway et al. 2009). Working memory abilities were not documented for children in this study, but CHH with stronger working memory abilities have been shown to have better word recognition abilities in quiet and in noise (McCreery et al. 2015). Audibility and working memory are two factors that might help to explain the differences in performance between groups, despite the similarities in language abilities.

Because the CHH required mores gates to achieve the isolation point in the LP but not the HP condition, we explored possible sources of differences by examining the structure of incorrect word responses and their patterns across gates, adapting methods of Salasoo and Pisoni (1985). Through analysis of error words, Salasoo and Pisoni found that adults integrated acoustic-phonetic and linguistic context cues (when available) in the process of word recognition. We reasoned that the patterns of errors in the two groups may be distinct, and this analysis could explain why CHH required more acoustic information than CNH to reach the isolation point. In the HP condition, incorrect word nominations that reflected reliance on language knowledge dominated the early gates for both groups, and errors declined systematically at subsequent gates. However, the CNH made no vowel errors at early gates. The few vowel errors made by CNH did not appear until gates 4 or 5. In contrast, vowel errors were observed at each gate in the CHH. In general, both groups appeared to be combining context cues and acoustic-phonetic cues from the signal for word recognition in the HP condition. In the LP condition, acoustic-phonetic errors were the prominent category followed by vowel errors for the CNH, and errors systematically declined at each successive gate in each error category except Language, which was near 0 from the beginning. Similar error patterns were observed in the CHH. Although vowel errors

accounted for more errors at gate 1 than acoustic-phonetic errors for the CHH, this difference was not significant.

The error analysis led us to question why children would focus on the initial consonant but suggest word candidates that allowed the vowel to vary. Studies of adult lexical selection processes suggest that vowel information is often regarded as more mutable than consonants in auditory word recognition. Listeners appear to allow more potential word candidates based on the vowel than they do for consonants (van Ooijen 1996; Cutler et al. 2000). This seems counterintuitive because vowels are more acoustically salient than consonants and are misperceived less often than consonants by children with hearing loss (Boothroyd 1984). However, certain regions in the vowel space have tokens from multiple categories that overlap one another (Hillenbrand et al. 1995) making precise vowel identity challenging at times. Van Ooijen points out that "salience is not synonymous with unambiguous" (page 574). She proposed that listeners may find the vowel information less reliable than consonants, which could lead listeners to postpone decisions about exactly which vowel was heard until they have more information. Results reported by Cutler and colleagues (2000) suggest that listeners treated consonants differently than vowels when there was uncertainty, including relaxing the constraints on vowel identity. This line of reasoning leads us to consider the possibility that the source of the vowel errors was not an issue of low level misperception, but rather that children, like adults, may postpone considering the precise vowel identity as they nominate word candidates until they have more information. It also would appear from our results that the CNH needed less acoustic-phonetic information than the CHH to narrow in on word candidates with vowels that matched the target. One limitation was that the CNH were listening to the signals through normal-hearing ears while CHH were listening to an amplified signal through a hearing aid simulator, which would have been different from listening to that same signal through their own hearing aids. Possible influences of these differences cannot be ruled out.

It may also be, however, that something intrinsic to the gating paradigm introduced challenges for CHH in perceiving vowels or vowel formant transitions. When vowels are truncated as they are in a gating paradigm, CHH may be more challenged than CNH to apprehend vowel transition cues. More research is needed to determine if children are performing like adults in allowing greater flexibility with vowels versus consonants or if the gating paradigm differentially affects recognition of vowels in CHH compared to CNH.

The second aspect of error analysis examined the sequences of incorrect word nominations and classified them as persistent or changing. The CNH used persistent responses and changing response patterns to similar degrees (52% and 48% respectively). In contrast, the CHH used a change pattern significantly more often than the CNH (67% of the time). It is possible that CHH are accustomed to having words repeated to them when they have not understood someone. Thus, it may be that they are more prone than CNH to interpret the repetition of the target at successive gates as a form of negative feedback (i.e., "my answer must be incorrect"). This tendency to change their answers in the gating paradigm may provide objective insight into their confidence in their responses. It is possible that the CNH persist with the same response on words for which they are confident in the portion they heard so far. What has been called response perseveration (answers that persist) could

possibly serve as a placeholder for children until they have sufficient acoustic-phonetic cues to disconfirm the established portion and narrow in on the correct target. In daily settings, CHH may alternate between guessing to fill in missing details or they may simply lack awareness of what they missed. Both experiences could impact self-appraisals of confidence in CHH. In contrast to our prediction, CHH were equally confident at the isolation point for both the HP and LP conditions. The CNH were less confident in the LP condition compared to the HP condition, which is consistent with other studies involving children (Craig et al. 1993). This difference between conditions within the CNH group may be related to the fact that they had experienced fewer gates/less acoustic detail at the isolation point in the LP condition compared to the CHH (who required more gates to succeed). Interestingly, the mean confidence rating for the CNH at the isolation point was 3.9. At this point, almost 95% of CHH were incorrect in their identification of the target word. Despite the low accuracy for CHH, their mean confidence rating at the same percentage of gates as the mean isolation point for CNH was 4.1 (+/− .5 SD). Given their high confidence even when their guesses were incorrect, it is not surprising that the CHH required a comparable percentage of gates to reach the acceptance point when compared to CNH. Thus, the overall results lead us to question whether the CHH were able to accurately self-rate their confidence in this task. This impression is supported by their frequency of changing their nominated word candidates at successive gates. We suggest that a pattern of frequent changing of responses may index the child's confidence better than subjective ratings.

Implications for Research

The gating paradigm presented in different predictability contexts has utility as a research tool to inform our understanding regarding children's approach to integrating linguistic, cognitive and acoustic-phonetic cues, which may reflect their routine experiences of coping with missing information (Elliott et al. 1990). In the current study, both groups of children performed well overall in time-gated word recognition, yet the paradigm revealed some subtle differences in their respective approaches. Future studies should examine the impact of predictability on word recognition using stimuli that are allowed to vary to a greater extent in familiarity and complexity than was done for this study. Such studies may further inform developmental effects in word recognition and rehabilitative practices for preparing CHH to cope with missing information.

Future studies should also examine time-gated word recognition in children who have greater than mild-moderate degrees of hearing loss. One participant who was excluded from the current study had moderately-severe hearing loss and demonstrated a much higher error rate (> 2 SDs) across gates than the other CHH. This child also had a pronounced tendency to nominate words that differed in vowels and to change his responses frequently across gates. Although the error rate and types may have been idiosyncratic to this particular child, it leaves open the possibility that the gating paradigm in the current study may yield different results depending on degree of hearing loss. Such findings would further support the notion that differences in audibility between the CNH and CHH influenced their performance in using acoustic-phonetic cues in the service of word recognition. It is possible that varying levels of audibility within CHH would contribute to gating performance outcomes in a larger, more diverse group.

Summary

In summary, our results indicate that age-matched groups of CNH and CHH with typical language abilities were able to identify words faster and with greater accuracy in HP condition compared to LP condition. This is consistent with previous studies demonstrating the facilitative effects of context on spoken word processing (Grosjean 1980; Tyler & Wessels 1983; Salasoo & Pisoni 1985; Craig at al. 1993). Although error patterns were similar across groups, some differences were observed in vowel errors. CHH required more gates than CNH to reach the isolation point reflecting accurate word recognition. This may be explained by differences in audibility between CNH and CHH. CHH were more inclined to change their responses at subsequent gates than the CNH. Although more research is needed, we suggest that this pattern of changing responses could provide more insight into their confidence in their responses than self-ratings of confidence. The gating paradigm provides insights about word recognition processes in CHH.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Alexander JM, Masterson K. Effects of WDRC Release Time and Number of Channels on Output SNR and Speech Recognition. Ear Hear. 2015; 36:e35–e349. [PubMed: 25470368]

Boersma, P., Weenink, D. Praat: doing phonetics by computer [Computer program]. 2013. Version 5.3.51, retrieved 8 November 2013 from http://www.praat.org/

Boothroyd A. Auditory perception of speech contrasts by subjects with sensorineural hearing loss. J Speech Hear Res. 1984; 27:134–144. [PubMed: 6716999]

Brennan MA, McCreery R, Kopun J, et al. Paired Comparisons of Nonlinear Frequency Compression, Extended Bandwidth, and Restricted Bandwidth Hearing Aid Processing for Children and Adults with Hearing Loss. J Am Acad Audiol. 2014; 25:983–998. [PubMed: 25514451]

Carrow-Woolfolk, E. Comprehensive Assessment of Spoken Language (CASL). Circle Pines, MN: American Guidance Service; 1999.

Charles-Luce J, Luce PA. An examination of similarity neighbourhoods in young children's receptive vocabularies. J Child Lang. 1995; 22:727–735. [PubMed: 8789521]

Conway CM, Pisoni DB, Kronenberger WG. The importance of sound for cognitive sequencing abilities the auditory scaffolding hypothesis. Current Dir Psychol Sci. 2009; 18(5):275–279.

Craig CH. Effects of aging on time-gated isolated word-recognition performance. J Speech Lang Hear Res. 1992; 35:234–238.

Craig CH, Kim BW. Effects of time gating and word length on isolated word-recognition performance. J Speech Lang Hear Res. 1990; 33:808–815.

Craig CH, Kim BW, Rhyner PMP, et al. Effects of word predictability, child development, and aging on time-gated speech recognition performance. J Speech Lang Hear Res. 1993; 36:832–841.

Cutler A, Sebastián-Gallés N, Soler-Vilageliu O, et al. Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. Mem Cog. 2000; 28:746–755.

Edwards J, Fox RA, Rogers CL. Final Consonant Discrimination in Children: Effects of Phonological Disorder, Vocabulary Size, and Articulatory Accuracy. J Speech Lang Hear Res. 2002; 45:231–242. [PubMed: 12003507]

Elliott LL, Hammer MA, Evan KE. Perception of gated, highly familiar spoken monosyllabic nouns by children, teenagers, and older adults. Percept Psychophys. 1987; 42:150–157. [PubMed: 3627935]

Elliott LL, Scholl ME, Grant JO, et al. Perception of gated, highly familiar spoken monosyllabic nouns by children with and without learning disabilities. J Learn Disabil. 1990; 23:248–252. [PubMed: 2324637]

Garlock VM, Walley AC, Metsala JL. Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. J Mem Lang. 2001; 45:468–492.

Grosjean F. Spoken word recognition processes and the gating paradigm. Percept Psychophys. 1980; 28:267–283. [PubMed: 7465310]

Grosjean F. The recognition of words after their acoustic offset: Evidence and implications. Percept Psychophys. 1985; 38:299–310. [PubMed: 3831907]

Hillenbrand J, Getty L, Clark M, Wheeler K. Acoustic characteristics of American English vowels. J Acoust Soc Am. 1995; 97:3099–3111. [PubMed: 7759650]

Jusczyk PW, Luce PA. Speech perception and spoken word recognition: Past and present. Ear Hear. 2002; 23:2–40. [PubMed: 11881915]

Kalikow DN, Stevens KN, Elliott LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J Acoust Soc Am. 1977; 61:1337–1351. [PubMed: 881487]

McAllister JM. The use of context in auditory word recognition. Percept Psychophys. 1988; 44:94–97. [PubMed: 3405735]

McCreery RW, Alexander J, Brennan M, et al. The influence of audibility on speech recognition with nonlinear frequency compression for children and adults with hearing loss. Ear Hear. 2014; 35:440–447. [PubMed: 24535558]

McCreery R, Walker E, Spratford M, et al. Speech recognition and parent ratings from auditory development questionnaires in children who are hard of hearing. Ear Hear. 2015; 36(Suppl 1):60S–75S. [PubMed: 26731160]

Metsala JL. An examination of word frequency and neighborhood density in the development of spoken-word recognition. Mem Cog. 1997; 25:47–56.

Munson B. Relationships between vocabulary size and spoken word recognition in children aged 3 to 7. Contemp Iss Commun Sci Disord. 2001; 28:20–29.

Nittrouer S. Discriminability and perceptual weighting of some acoustic cues to speech perception by 3-year-olds. J Speech Lang Hear Res. 1996; 39:278–297.

Pisoni DB, Cleary M. Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. Ear Hear. 2003; 24(1 Suppl):106S. [PubMed: 12612485]

Roebers CM. Confidence Judgements in Children's and Adults' Event Recall and Suggestibility. Develop Psychol. 2002; 38:1052–1067.

Roebers CM, Gelhaar T, Schneider W. "It's magic!" The effects of presentation modality on children's event memory, suggestibility, and confidence judgements. J. Exper Child Psychol. 2004; 87:320–335. [PubMed: 15050457]

Salasoo A, Pisoni DB. Interaction of knowledge sources in spoken word identification. J Mem Lang. 1985; 24:210–231. [PubMed: 23226691]

Scollie S, Seewald R, Comelisse L, et al. The Desired Sensation Level multistage input/output algorithm. Trends Amplif. 2005; 9:159–197. [PubMed: 16424945]

Stiles DJ, McGregor KK, Bentler RA. Vocabulary and working memory in children fit with hearing aids. J Speech Lang Hear Res. 2012; 55(1):154–167. [PubMed: 22199188]

Stiles DJ, McGregor KK, Bentler RA. Wordlikeness and word learning in children with hearing loss. Int J Lang Commun Disord. 2013; 48:200–206. [PubMed: 23472959]

Storkel HL, Hoover JR. An on-line calculator to compute phonotactic probability and neighborhood density based on child corpora of spoken American English. Behav Res Methods. 2010; 42:497–506. [PubMed: 20479181]

Tomblin JB, Harrison M, Ambrose SE, Walker EA, Oleson JJ, Moeller MP. Language outcomes in young children with mild to severe hearing loss. Ear Hear. 2015; 36(Suppl 1):76S–91S. [PubMed: 26731161]

Tyler LK, Wessels J. Quantifying contextual contributions to word-recognition processes. Percept Psychophys. 1983; 34:409–420. [PubMed: 6657445]

Van Ooijen B. Vowel mutability and lexical selection in English: Evidence from a word reconstruction task. Mem & Cog. 1996; 24:573–583.

Walley AC. Spoken word recognition by young children and adults. Cog Develop. 1988; 3:137–165.

Walley AC, Michela VL, Wood DR. The gating paradigm: Effects of presentation format on spoken word recognition by children and adults. Percept Psychophys. 1995; 57(3):343–351. [PubMed: 7770325]

Williams, K. Expressive Vocabulary Test, Second Edition (EVT-2). Minneapolis, MN: NCS Pearson; 2007.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript
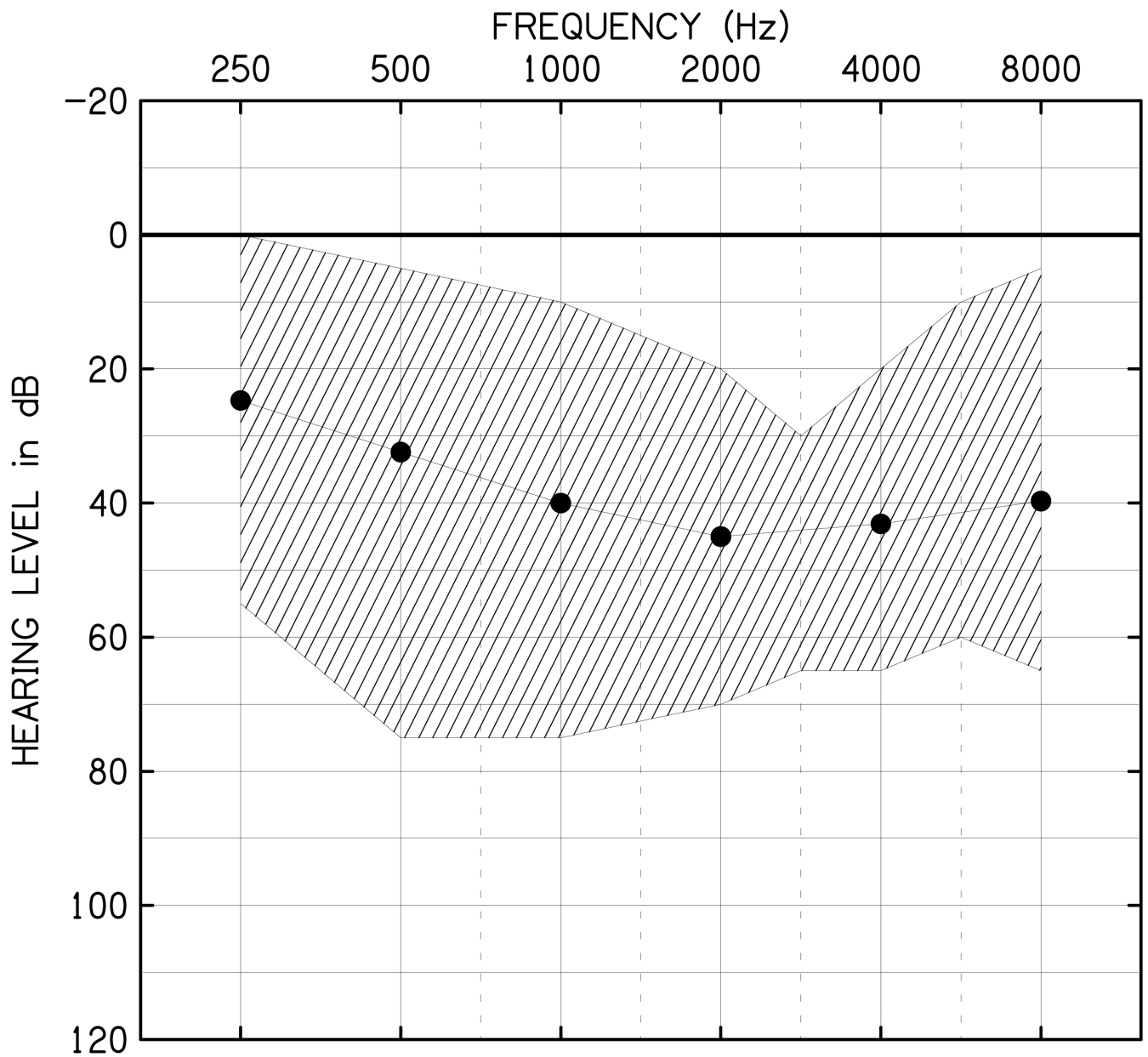
Figure 1.
Hearing thresholds in dB HL as a function of frequency in Hz for the participants with hearing loss (n = 16). Solid dots represent the mean thresholds for the test ear at each frequency. The cross-hatched area represents the range of thresholds in the sample at each frequency.
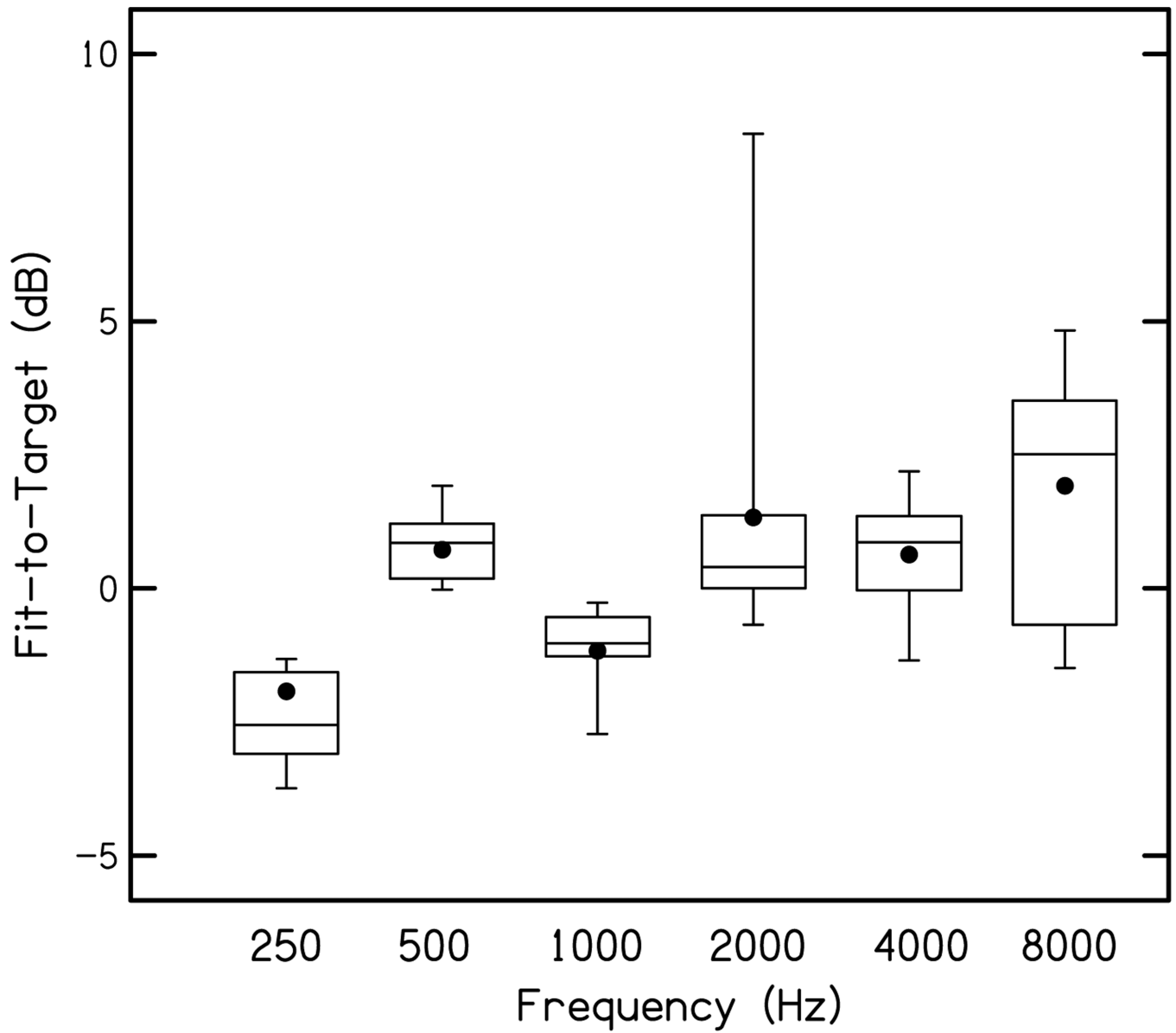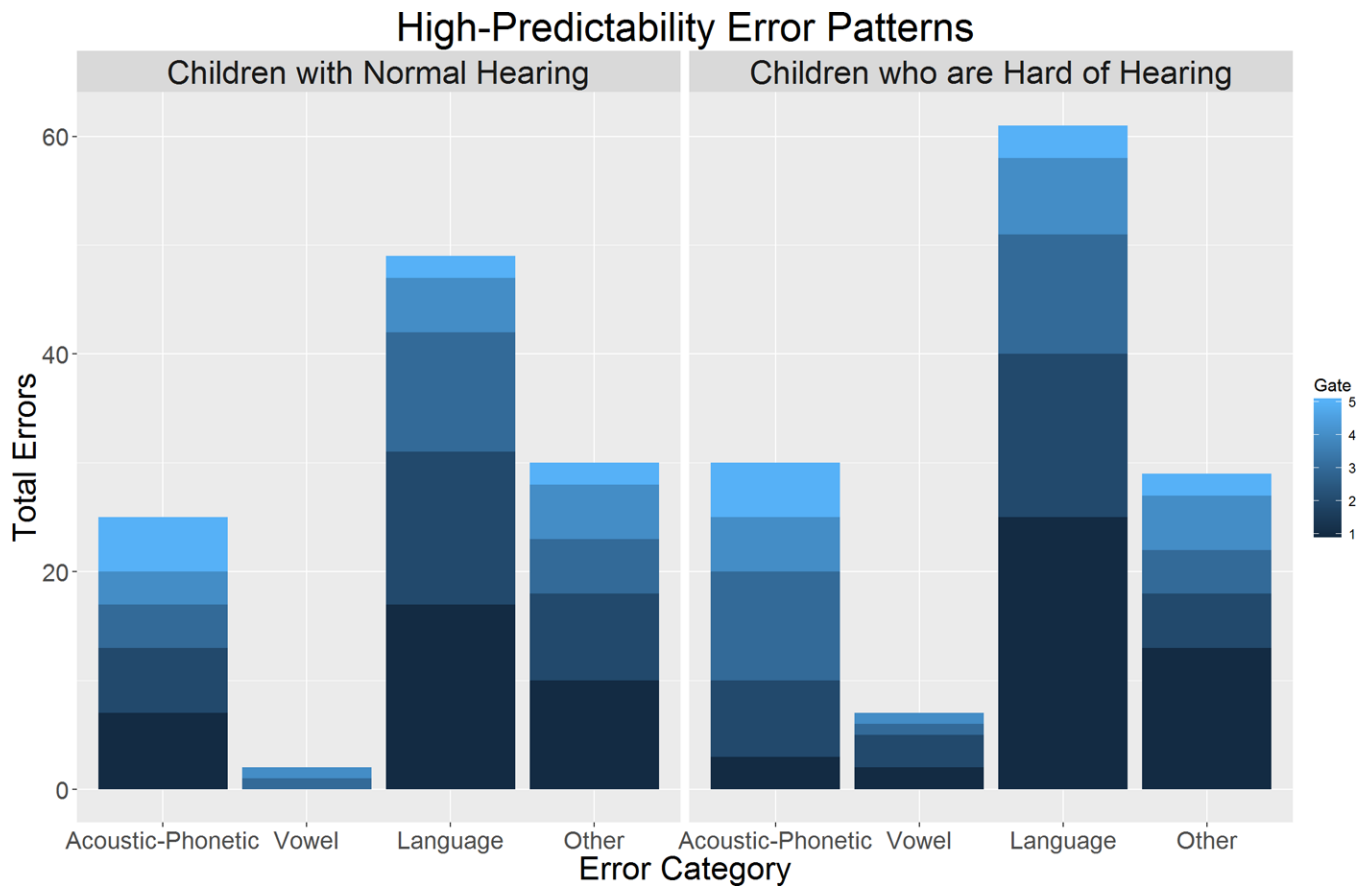
Figure 2.
Fit-to-target differences between the output of the hearing aid simulator and DSL targets across frequency. The boxes represent the 25th to 75th percentiles. The whiskers represent the 5th and 95h percentiles. Within each box, the line represents the median, and the filled circle represents the mean.
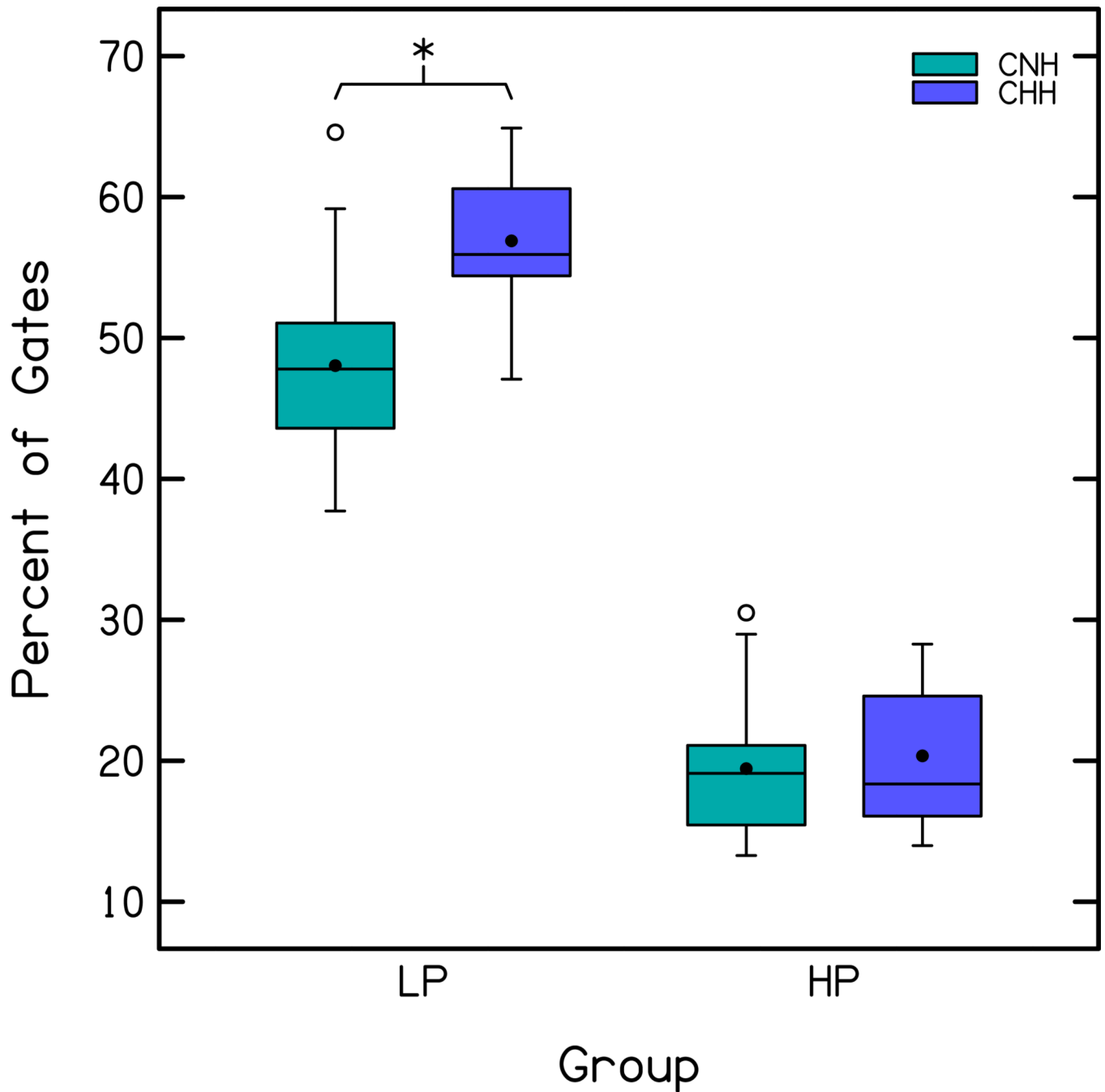
Figure 3.
Percent of gates required to reach the isolation point in the high (HP) and low (LP) predictability sentence conditions for the CNH ($n = 16$) and the CHH ($n = 16$). The boxes represent the 25th to the 75th percentiles (interquartile range) and the whiskers represent the 5th to 95th percentiles. Solid lines are the median and solid circles represent the mean. Open circles represent responses greater than the 95th percentile. *$p < .001$; CNH = children with normal hearing; CHH = children who are hard of hearing
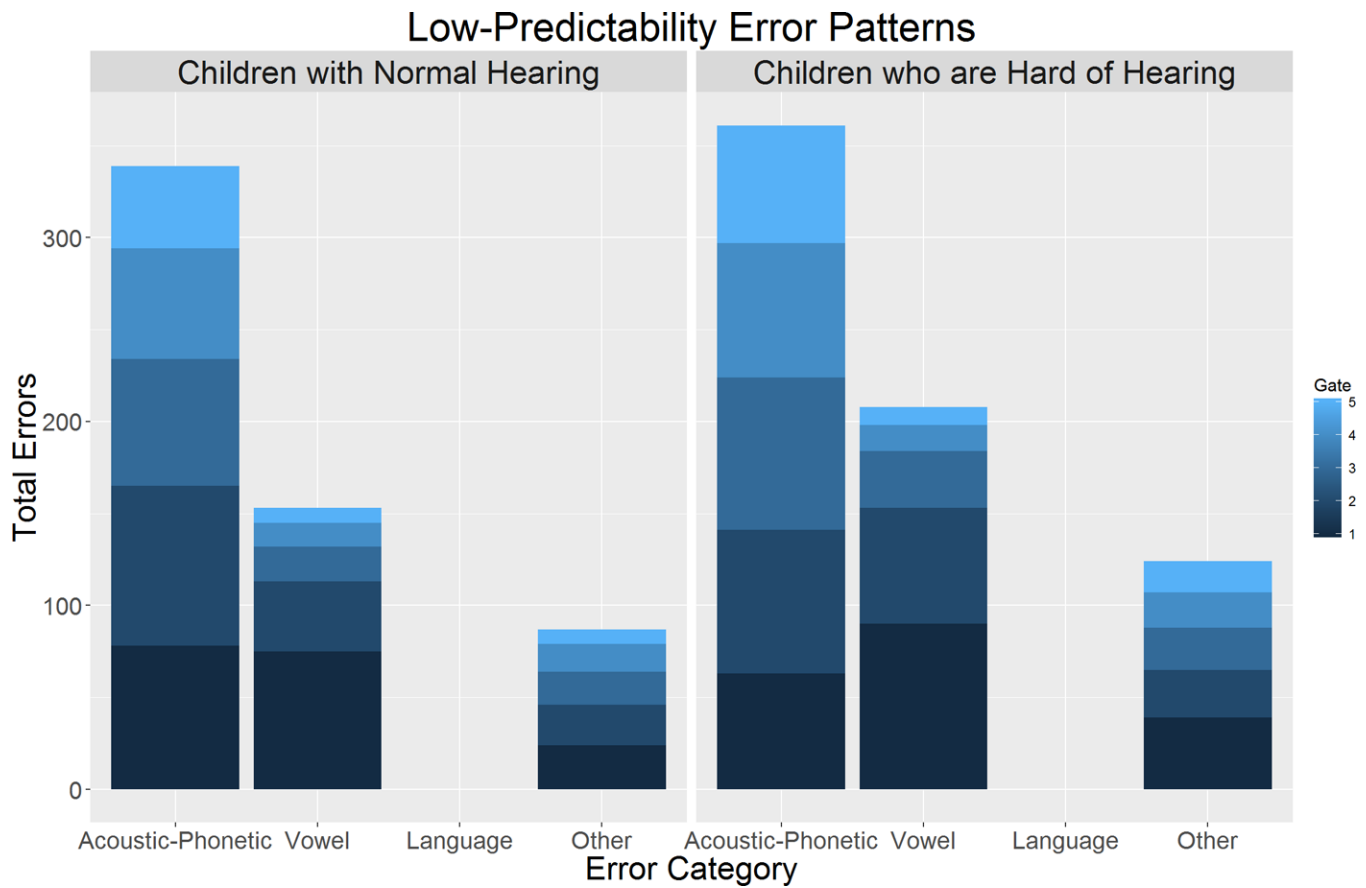
High-Predictability Error Patterns

Figure 4.
Incorrect word nominations categorized by hypothesized knowledge sources for the children
with normal hearing and children who are hard of hearing in the HP (upper panel) and LP
(lower panel) conditions. Shading in each bar represents the successive gates (1 – 5). For
both groups, the number of errors in the upper panel (HP) is fewer than in the lower panel
(LP).
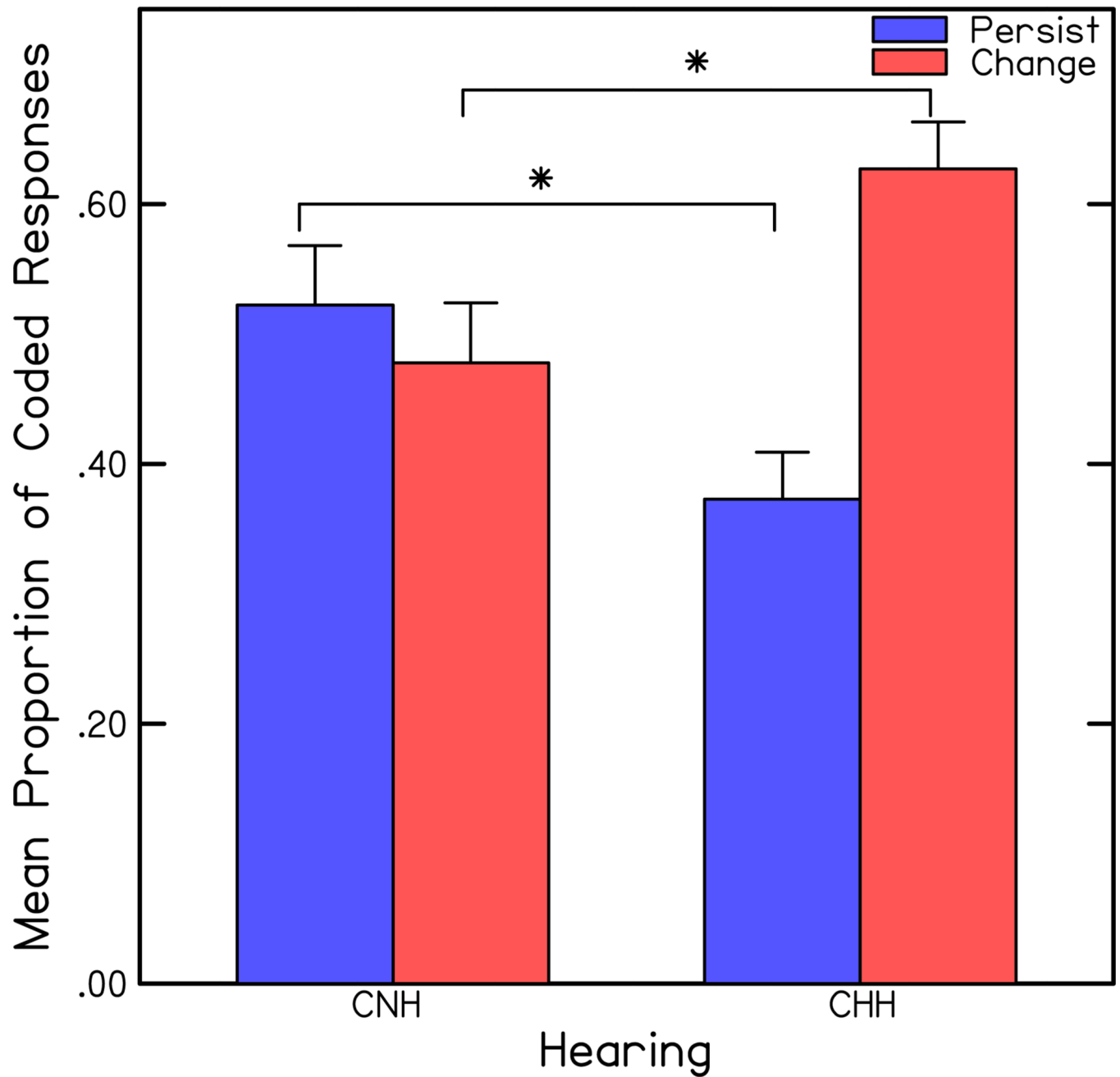
Figure 5.
Average proportion of error patterns categorized as *persist* or *change* for the children with normal hearing (CNH) and children who are hard of hearing (CHH) in the low predictability condition. Asterisks represent $p = .02$ for between group comparisons for both categories (persist and change).

Table 1

Demographic characteristics for the participants.

	CNH (*n* = 16)	CHH (*n* = 16)	*p*
Age (years)			
Mean	8.75	8.85	.914
SD	2.65	2.43	
Range	5.2–12.5	5.5–12.9	
Maternal Ed			
High school or less	1 (6.3%)	5 (31.3%)	.133[a]
Some college	5 (31.3%)	5 (31.3%)	
College graduate	3 (18.8%)	1 (6.3%)	
Graduate education	7 (43.7%)	5 (31.3%)	
Sex			
Male	5 (31.3%)	7 (43.7%)	.465[b]
Female	11 (68.7%)	9 (56.3%)	
EVT Standard Score (SS)			
Mean	108.19	113.56	.273
SD	12.32	14.80	
Range	88–135	93–148	
CASL Sentence Completion SS			
Mean	109.44	110.00	.905
SD	9.48	16.14	
Range	94–133	82–148	

Note. CNH = children with normal hearing; CHH = children who are hard of hearing; Maternal Ed = maternal education levels; EVT = Expressive Vocabulary Test; CASL = Comprehensive Assessment of Spoken Language; SS = Standard Score;

[a] Kruskal Wallis Test;

[b] Pearson Chi Square

Table 2

Audiological data for the children who are hard of hearing.

Variable	n	M	SD	Range
Age at Identification (months)	16	22.31	20.23	0 – 48
Age at HA fit (months)	15	39.09	33.26	2 – 96
HA use at school (hours per day)	15	6.81	0.75	5–8
HA use weekends (hours per day)	15	8.19	5.82	0–14
PTA-right (dB HL)	16	39.56	17.52	20–85
PTA- left (dB HL)	16	38.25	10.72	18–52
PTA-test ear (dB HL)	15	37.10	11.30	18–52
Aided SII – right ear	15	80.20	15.72	33 – 94
Aided SII – left ear	15	83.26	7.99	65–93
Test ear aided SII	15	83.23	8.44	68–94

Note. HA=hearing aid, PTA = pure tone average, SII = speech intelligibility index.

Table 3

Mean percentage of words for which children achieved the isolation point (IP: n= 16 per group) and acceptance point (AP: n = 13 per group) by predictability condition. Comparisons are from the mixed-design analyses of variance.

| Dependent Variable | CNH | | CHH | | Comparisons | | | | |
| | HP | LP | HP | LP | Between Condition | | Between Group | |
	M (SD)	M (SD)	M (SD)	M (SD)	p	η^2_p	p	η^2_p
IP	99.17 (3.33)	94.17 (10.29)	99.17 (2.28)	95.42 (5.29)	.001	.34	.74	.01
AP	97.44 (9.24)	84.10 (19.72)	98.46 (2.92)	84.62 (21.49)	.001	.39	.88	.001

Note: CNH = children with normal hearing; CHH = children who are hard of hearing; HP = high predictability condition; LP = low predictability condition

Table 4

Percent of gates to reach the isolation point for children who are normal hearing (CNH: n = 16) and children who are hard of hearing (CHH: n = 16) by condition Comparisons are from the mixed-design analysis of variance.

| | CNH | | CHH | | Comparisons | | | |
| | HP | LP | HP | LP | Between Condition | | Between Group | |
	M (SD)	M (SD)	M (SD)	M (SD)	p	η^2_p	p	η^2_p
Percent of Gates	19.43 (5.30)	48.04 (7.44)	20.35 (5.43)	56.89 (5.69)	.001	.97	.011	.196

Note: HP = high predictability condition; LP = low predictability condition.

Table 5

Percent of gates to reach the acceptance point for children who are normal hearing (CNH: n = 14) and children who are hard of hearing (CHH: n = 14) by condition Comparisons are from the mixed-design analysis of variance.

| | CNH | | CHH | | Comparisons | | | |
| | | | | | Between Condition | | Between Group | |
	HP	LP	HP	LP	p	η^2_p	p	η^2_p
	M (SD)	M (SD)	M (SD)	M (SD)				
Percent of Gates	39.08 (17.86)	78.80 (16.78)	42.10 (15.02)	80.24 (11.31)	<.001	.86	.656	.008

Note: HP = high predictability condition; LP = low predictability condition.

Table 6

Results of Mixed-design ANOVA analyzing the structure of incorrect word candidates for the first 5 gates.

Effect	F	p	η_p^2
Error Type (E)[*]	96.03	<0.001	0.762
Gates (G)	152.59	<0.001	0.836
Predictability (P)	756.57	<0.001	0.962
Hearing (H)	7.80	0.009	0.206
E x G[*]	15.66	<0.001	0.343
E x P[*]	128.28	<0.001	0.810
E x H[*]	.52	0.672	0.017
G x P	67.15	<0.001	0.691
P x H	756.57	0.020	0.168
E x G x H[*]	3.94	<0.001	0.116
E x P x H[*]	.93	0.429	0.030
E x G x P[*]	30.77	<0.001	0.506
G x P x H	1.34	0.260	0.043
E x G x P x H[*]	2.39	0.027	0.074

[*] Greenhouse-Geisser correction used due to violation of sphericity.