Behavioral Measures of Listening Effort in School-Age Children: Examining the Effects of Signal-to-Noise Ratio, Hearing Loss, and Amplification

Ronan McGarrigle,[1] Samantha J. Gustafson,[2] Benjamin W. Y. Hornsby,[3] and Fred H. Bess[3]

Objectives: Increased listening effort in school-age children with hearing loss (CHL) could compromise learning and academic achievement. Identifying a sensitive behavioral measure of listening effort for this group could have both clinical and research value. This study examined the effects of signal-to-noise ratio (SNR), hearing loss, and personal amplification on 2 commonly used behavioral measures of listening effort: dual-task visual response times (visual RTs) and verbal response times (verbal RTs).

Design: A total of 82 children (aged 6–13 years) took part in this study; 37 children with normal hearing (CNH) and 45 CHL. All children performed a dual-task paradigm from which both measures of listening effort (dual-task visual RT and verbal RT) were derived. The primary task was word recognition in multi-talker babble in three individually selected SNR conditions: Easy, Moderate, and Hard. The secondary task was a visual monitoring task. Listening effort during the dual-task was quantified as the change in secondary task RT from baseline (single-task visual RT) to the dual-task condition. Listening effort based on verbal RT was quantified as the time elapsed from the onset of the auditory stimulus to the onset of the verbal response when performing the primary (word recognition) task in isolation. CHL completed the task aided and/or unaided to examine the effect of amplification on listening effort.

Results: Verbal RTs were generally slower in the more challenging SNR conditions. However, there was no effect of SNR on dual-task visual RT. Overall, verbal RTs were significantly slower in CHL versus CNH. No group difference in dual-task visual RTs was found between CNH and CHL. No effect of amplification was found on either dual-task visual RTs or verbal RTs.

Conclusions: This study compared dual-task visual RT and verbal RT measures of listening effort in the child population. Overall, verbal RTs appear more sensitive than dual-task visual RTs to the negative effects of SNR and hearing loss. The current findings extend the literature on listening effort in the pediatric population by demonstrating that, even for speech that is accurately recognized, school-age CHL show a greater processing speed decrement than their normal-hearing counterparts, a decrement that could have a negative impact on learning and academic achievement in the classroom.

Key words: Listening effort, Dual-task paradigm, Response times, Hearing loss, Children, Cognition, Speech recognition.

(Ear & Hearing 2019;40;381–392)

INTRODUCTION

For children with hearing loss (CHL), assessment of auditory function in the clinic typically consists of pure-tone and speech audiometry testing in controlled settings. However, everyday listening occurs in complex (often noisy) acoustic environments. Children develop key cognitive, linguistic, and academic skills in the classroom—an acoustically adverse environment with reported signal-to-noise ratios (SNRs) ranging from +1 to +11 dB (Larsen & Blair, 2008; Sato & Bradley, 2008). Speech understanding in noise relies on cognitive skills such as working-memory and attention, which are still undergoing a process of maturation in school-age children (Gomes et al., 2000; Luna et al., 2004). The increased mental demands of listening in adverse conditions will likely interfere with a child's ability to perform important cognitive functions, like memorizing and comprehending information (Howard et al., 2010; Klatte et al., 2010). Owing to the importance of higher level cognitive abilities for communicative success in CHL, an increasing number of studies have sought to measure the allocation of mental resources during listening (i.e., "listening effort") using behavioral and/or physiological methods (Amlani & Russo, 2016; Grieco-Calub et al., 2017; Gustafson et al., 2014; Hicks & Tharpe, 2002; Howard et al., 2010; Hsu et al., 2017; Hughes & Galvin, 2013; McFadden & Pittman, 2008; McGarrigle et al., 2017; Steel et al., 2015; Stelmachowicz et al., 2007). A sensitive and reliable measure of listening effort could help to provide a more comprehensive assessment of the disability associated with listening in challenging environments for CHL.

One behavioral method commonly used to measure listening effort is the "dual-task" paradigm, which has been used in the field of cognitive psychology to measure attention allocation (Broadbent, 1958; Kahneman, 1973; Styles, 2006). During the dual-task paradigm, participants are asked to perform two tasks (a "primary" task and a "secondary" task), each completed in isolation and concurrently. In the case of measuring listening effort, the primary task typically involves listening (e.g., speech recognition) in different acoustic conditions, while the secondary task varies substantially across studies, but usually involves performance on either a memory or nonauditory (e.g., visual, motor or tactile) response task. Secondary task performance can be quantified in terms of accuracy or response speed and is generally measured in isolation (single-task performance) and in conjunction with the primary task (dual-task performance). Any performance decrement (accuracy or response speed) on the *secondary* task when multi-tasking is interpreted as reflecting the allocation of resources shifted from the secondary task to the primary (listening) task. This decrement is therefore interpreted as the effort required to listen (Gagné et al., 2017).

Hicks and Tharpe (2002) assessed dual-task performance in a small group of children with normal-hearing (CNH; n = 10) and in children with mild-to-moderate hearing loss (n = 10), all between 5 and 11 years old. All but one of the CHL wore their hearing aids during testing. The primary task was word recognition in multi-talker babble at varying SNRs (+ 20 dB, + 15 dB,

---

[1] Department of Psychology, University of Chester, Parkgate Road, Chester, United Kingdom; [2] Department of Otolaryngology/Head and Neck Surgery, The University of North Carolina at Chapel Hill, Chapel Hill, NC; [3] Department of Hearing and Speech Sciences, Vanderbilt Bill Wilkerson Center, Vanderbilt University School of Medicine, Nashville, TN.

0196/0202/2019/402-381/0 • Ear & Hearing • Copyright © 2018 Wolters Kluwer Health, Inc. All rights reserved • Printed in the U.S.A.

<zdoi; 10.1097/AUD.0000000000000623>

and + 10 dB). The secondary task required participants to press a button every time a light probe was displayed. Overall, secondary task response times (RTs) were slower in the CHL compared with the CNH, suggesting increased listening effort. Given the relatively favorable SNRs used in this study, the authors suggested that these findings might underestimate the difficulty faced by CHL in noisy classrooms. Howard et al (2010) investigated listening effort in CNH using the dual-task paradigm at SNRs more representative of a typical classroom (-4, 0, and +4 dB) and found evidence of a systematic increase in listening effort as SNRs became more challenging. Given that CHL generally show greater difficulty understanding speech in noise than CNH (Leibold et al., 2013; Ruscetta et al., 2005), it is likely they will also experience greater listening effort in more adverse SNRs.

The dual-task method has high ecological validity; individuals are often tasked with performing multiple cognitive operations simultaneously, particularly in a learning environment (e.g., following a teacher's instructions while taking notes). However, interpretation of dual-task results relies on the assumption that specific tasks can be prioritized and/or cognitive resources can be distributed among multiple concurrent tasks (Gagné et al. 2017). While this executive control ability is thought to be intact in most adult populations, previous research indicates that this ability is not yet fully developed in school-age children. For example, Choi et al. (2008) examined top-down control of attention in children (ages 7–14 years) using the dual-task method. In their study, participants were asked to prioritize either the word recognition or digit-recall task. Participants performed consistently well on word-recognition and consistently poor on digit recall in the dual-task condition, despite receiving instructions to prioritize one task over the other. These findings suggest that school-age children have some difficulty allocating their attention in a preferential way during multi-tasking situations.

As an alternative to recording secondary-task RTs in a dual-task paradigm, verbal RTs are also believed to reflect listening effort (Houben et al., 2013). To assess verbal RTs, participants are typically asked to respond to an auditory stimulus, for example, verbally repeating a word/sentence. The interval between the stimulus offset and the verbal response is used as a measure of the effort required to achieve accurate performance. That is, slower verbal RTs are taken to reflect increased effort. Research in the adult population suggests that verbal RTs could provide supplementary information to accuracy measures alone (Houben et al., 2013; Pals et al., 2015). In particular, RTs have been shown to reveal benefit from amplification in listeners with hearing loss (Gatehouse & Gordon, 1990) and increased listening effort in more challenging SNRs (Houben et al., 2013; Pals et al., 2015). The use of the verbal RTs as a measure of listening effort may have some advantages compared with dual-task measures in that, if automated, RT information can be relatively simple to obtain alongside a routine speech recognition assessment (Houben et al., 2013; Pals et al., 2015).

Lewis et al. (2016) examined the effects of stimulus type (consonant, word, and sentence recognition) and SNR (-5, 0, and +5 dB) on verbal RTs in three groups of school-age children (aged 5–12 years): (1) CNH, (2) children with mild bilateral hearing loss (MBHL), and (3) children with unilateral hearing loss (UHL). Both groups of CHL performed the task without the use of a hearing aid. Unlike Hicks and Tharpe (2002), analyses showed a main effect of SNR, but not group (i.e., hearing status). Verbal RTs were generally slower in more challenging SNRs across all groups, but there was no overall difference between CNH, UHL, or MBHL. The lack of group difference is in line with a dual-task study investigating top-down control of attention in a comparable sample of children with UHL and MBHL, which also showed no difference in performance accuracy between groups when using a dot-to-dot game as the secondary task (McFadden & Pittman, 2008). Noting that CHL in Hicks and Tharpe (2002) had poorer hearing sensitivity (mild-to-moderate high-frequency hearing loss), Lewis et al. (2016) suggested that the effects of listening effort in children may be related to hearing loss severity. It remains to be seen whether verbal RTs or dual-task measures are sensitive to differences in listening effort between CNH and school-age children with greater degrees of hearing loss.

With regard to ameliorating the increased listening effort experienced by CHL, previous studies have investigated the effects of digital noise reduction technology (Gustafson et al., 2014) and auditory stimulus bandwidth (Stelmachowicz et al., 2007). However, the more general effect of personal amplification (unaided versus aided) on listening effort has not yet been examined in school-age CHL. In Gustafson et al. (2014), CNH listened to speech processed through a hearing aid with and without digital noise reduction and were presented with consonant-vowel-consonant nonwords in broadband background noise. Verbal RTs were slower when digital noise reduction was switched off versus when it was switched on, suggesting a reduction in listening effort associated with the use of digital noise reduction technology. Stelmachowicz et al. (2007) investigated the effect of stimulus bandwidth (5 versus 10 kHz) on dual-task performance in a group of CNH and CHL (aged 7–14 years). Word recognition and digit recall were used as primary and secondary tasks, respectively. No difference in listening effort was found between CNH and CHL or between stimulus bandwidth conditions. Clearly, further research is needed to better understand the effects of personal amplification on dual-task and verbal RT measures of listening effort in CHL.

The current study examined the effects of hearing status, SNR, and amplification on listening effort in a group of school-age children. Specifically, we were interested in comparing the sensitivity of two commonly used behavioral measures of listening effort: (1) dual-task visual RTs, and (2) verbal RTs. We hypothesized:

1. slower dual-task visual RTs and verbal RTs (i.e., increased listening effort) in more challenging SNRs across both CNH and CHL, consistent with Howard et al. (2010) and Lewis et al. (2016);
2. slower dual-task visual RTs and verbal RTs (i.e., increased listening effort) in CHL versus CNH, based on the findings from Hicks and Tharpe (2002) and the prediction of Lewis et al. (2016);
3. slower dual-task visual RTs and verbal RTs (i.e., increased listening effort) in unaided versus aided conditions, supporting previous research showing reduced listening effort in school-age CNH using digital noise reduction technology (Gustafson et al., 2014).

MATERIALS AND METHODS

This study was a part of a broader research program designed to examine the effects of listening effort and fatigue on school-age CHL—see Bess et al. (2014) for an overview of that program.

Participants

A total of 82 children between the ages of 6.0 and 12.9 years were included in this study. Within this sample, 37 children had normal hearing sensitivity, and 45 children had hearing loss. Data for 11 additional CHL and two additional CNH were not included in the analyses for the following reasons: (1) word recognition performance was less than 20% in all test conditions (n = 7), (2) they could not master the dual-task paradigm (n = 2), (3) they elected to stop participation prior to completing the task (n = 3), or (4) a computer error resulted in a completely missing data set (n = 1). Children had no diagnosis of learning disability or cognitive impairment as reported by parents. The parent of one child only (in the CHL-aided group) reported that their child had vision problems not fully corrected by glasses. For the purposes of the larger study, children were also excluded based on factors known to affect fatigue. This criterion resulted in the exclusion of (a) children who were bilingual or whose primary language in the home was not listening and spoken language, (b) children with Autism Spectrum Disorder, (c) children with linear metabolic or endocrine disorders (e.g., diabetes or hypothyroidism), (d) children with a chronic medical condition, and (e) children who utilized stimulant medications. Children with cochlear implants and children with UHL were not included in the study.

Participants were recruited from Vanderbilt's pediatric audiology clinics, school systems throughout the middle Tennessee area, advertisements in a local parenting magazine, and through the Vanderbilt Kennedy Center's Study Finder website. This study was reviewed and approved by the Institutional Review Board of Vanderbilt University. All children provided their assent, and parents/caregivers provided written informed consent prior to the initiation of any research procedures. Children were compensated for their participation.

Upon entry into the larger study, each child received an audiological assessment and a series of standardized tests to assess language and nonverbal intelligence. Language ability was measured using the core language index of the Clinical Evaluation of Language Fundamentals – Fourth Edition (Wiig et al., 2004), providing a reliable, norm-referenced measure of language performance by age. Children also received the Test of Nonverbal Intelligence - Fourth Edition (Brown et al., 2010). Demographic, nonverbal intelligence, and language information obtained at study entry is shown in Table 1 for each group.

All participants had normal middle ear function verified by tympanometry as well as unremarkable otoscopic examinations. CNH received a standard hearing screening at 15 dB HL for octave frequencies ranging from 0.25 to 8.0 kHz, bilaterally. CHL received an audiological examination including bilateral air and bone conduction threshold testing. CHL had permanent, bilateral, sensorineural hearing loss. Losses were in the mild-to-moderate range in at least the better-hearing ear. We defined mild hearing loss using the following criteria as measured in the better hearing ear: a pure tone average (PTA; thresholds at 0.5, 1.0, and 2.0 kHz) between 20 and 40 dB or thresholds greater than 25 dB HL at two or more frequencies above 2.0 kHz. Moderate hearing loss was defined as a PTA of > 40–70 dB HL in the better ear. During aided testing, all CHL wore two hearing aids. Figure 1 shows a composite audiogram for the CHL included in this data set.

TABLE 1. Participant Characteristics

	CNH	CHL	Significance (*p*)
Participants; males/females	37; 24/13	45; 24/21	
Age (y)	8.92 (2.22)	10.09 (1.87)	0.011
Language*	110.19 (10.15)	97.02 (20.97)	0.001
Nonverbal intelligence quotient†	110.11 (9.15)	104.44 (12.58)	0.025

Mean (and SD) age, language, nonverbal intelligence quotient, and number of children with mild versus moderate hearing losses for both groups of children enrolled in the study. Mild/Moderate HL 0/0 29/16.

**Standard score on the core language index of the Comprehensive Evaluation of Language Fundamentals-Fourth Edition.*

†Standard score on Test of Nonverbal Intelligence-Fourth Edition.

Materials

Twenty isophonemic lists, each containing 10 consonant-vowel-consonant words, were used as test stimuli (Mackersie et al., 2001). Of the 20 word lists, the same two lists (lists 13 and 20) were used for screening, practice, and for secondary-task alone trials. The remaining 18 word lists were split into two groups for testing in the primary or dual-task conditions (Primary task: lists 1–6, 14–16; Dual-task: lists 7–12, 17–19). There were two versions of these groups (Word List A and Word List B), in which the words were presented in a different order. In general, Word List A was used for unaided testing and Word List B for aided testing. Thus, there was no replication of word lists during a single unaided or aided visit. For those CHL tested in both unaided and aided conditions, all word lists were replicated on the second visit, although the order of words within lists varied between visits. The background noise consisted of uncorrelated segments of multi-talker babble sampled from the "Connected Speech Test" (Cox et al., 1987) and presented from four loudspeakers located around the listener (3.5 meters from the listener at 45°, 135°, 225°, and 315° azimuths). The level of each background-noise loudspeaker was equated and

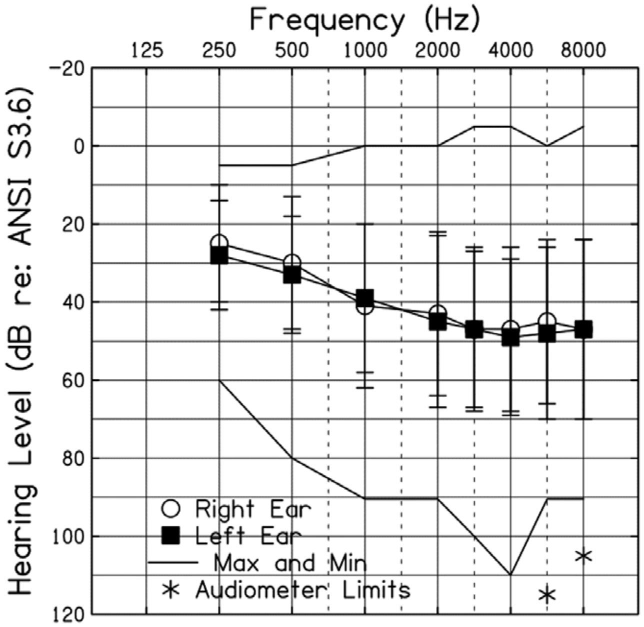

Fig. 1. Composite audiogram for CHL enrolled in the study.

adjusted to produce an overall noise level of 56 dB(A) when all four loudspeakers were played simultaneously. The background noise was presented continuously throughout testing. We opted to present our noise at a fixed level of 56 dB (A), rather than fixing the speech level, to reduce the likelihood of children becoming uncomfortable while listening to the continuous noise in a reverberant setting. The speech level was adjusted by the examiner during testing to create three SNRs. These levels were chosen to systematically vary task difficulty while limiting floor and ceiling effects. For the CNH, this goal could be achieved using three, fixed, SNRs (-4, 0, +4 dB). However, research suggests that CHL may require more favorable SNRs than CNH to achieve comparable performance (Leibold et al., 2013; Ruscetta et al., 2005). Therefore, to limit floor and ceiling effects, we used three potential SNR combinations (-4, 0, +4 dB; 0, +4, +8 dB; and +4, +8, +12 dB) for the CHL. Screening procedures used to determine each child's SNR combination are described below. To illustrate the systematic variation in task difficulty in each SNR combination, we will refer to SNRs within each combination as Hard, Moderate, and Easy.

Procedures

Hearing Aid Measurement • During aided testing, CHL wore personal hearing aids using their "as-worn" settings (i.e., no programming adjustments were made by research staff prior to testing). Real ear–aided responses were measured by a trained research assistant for each hearing aid using probe-microphone methods on the Audioscan Verifit Real Ear System (Audioscan, Dorchester, Ontario) prior to the start of a child's aided appointment. Hearing aid output in the child's ear, at octave frequencies, was compared to Desired Sensation Level (DSL; Scollie et al., 2005) v5.0 targets using the Verifit's "standard speech signal" (the carrot passage) presented at 65 dB SPL. Figure 2 shows DSL v5.0 targets and measured real ear–aided response values, at octave frequencies from 250 to 4000 Hz, averaged across participants who completed the listening effort task in the aided condition (n = 41). To assess whether a poor match to DSL targets affected listening effort, we used these data to calculate a better-ear average root-mean-square (RMS) error for each fitting. We also quantified audibility for each participant using the speech intelligibility index (SII; ANSI S3.5). SII's were calculated using the Verifit in unaided and aided conditions for average-level speech in quiet (65 dB SPL).

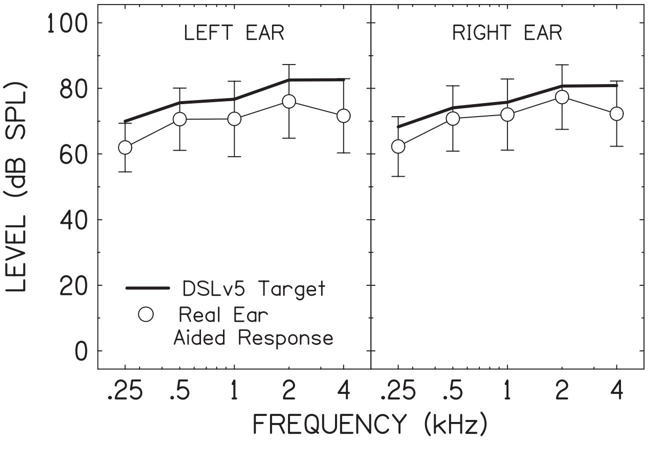

Fig. 2. DSL v5.0 targets and measured real ear aided response values at octave frequencies from 250 to 4000 Hz averaged across participants included in the aided testing condition

Experimental Testing • Listening effort testing was conducted in a 5.9 × 5.1 × 2.5 m reverberation chamber modified with acoustic blankets to create a moderately reverberant condition (average reverberation time; RT60 ~ 450 ms). Children were seated in the center of the chamber at a desk that had colored handprints on the surface showing where to place their hands. They were instructed to rest their hands on the handprints during all tasks unless they were pressing the button during the secondary or dual-task conditions. A trained examiner was seated in the test room during testing to provide verbal instruction and ensure that children remained on-task.

Screening Test • The SNR combination used for each child was determined using a screening test completed upon entry to the larger study. For CHL, this screening was completed first without hearing aids. Children were asked to repeat two 10-word lists (lists 13 and 20) at two of the following three SNRs: 0, -4, or +8 dB SNR. The goal of this screening was to determine a range of SNRs that would minimize floor and ceiling effects for CHL. To this end, we chose the lowest (hardest) SNR at which each child could score at least 20% correct in the "hard" condition. Thus, children were tested first at 0 dB SNR. If they scored better than 20% correct at 0 dB SNR, they were retested at -4 dB SNR. However, if they scored worse than 20% correct at 0 dB SNR, they were retested at +8 dB SNR. Children who scored at least 20% correct in the -4 dB SNR condition were tested with the -4, 0, +4 dB SNR combination. Children who scored poorer than 20% at a -4 dB SNR but scored at least 20% correct at 0 dB SNR were tested with the 0, +4, +8 dB SNR combination. Children who scored poorer than 20% in the 0 dB SNR but scored at least 20% correct at +8 dB SNR were tested using the +4, +8, +12 dB SNR combination. Table 2 shows the number of children tested in each SNR combination.

For CHL who completed aided and unaided testing, the same SNR combination was used in both conditions. If a child scored less than 20% in the +8 dB SNR condition, he/she did not complete listening effort testing in the unaided condition. In this case, the child wore their hearing aids and repeated the screening test to determine the lowest (hardest) SNR (from -4 dB to +8 dB) at which they could score at least 20% correct. During this screening, the same two word lists used in the unaided testing were used for screening in the aided condition. A total of 18 children were unable to complete the testing unaided. Of these children, 12 had a "moderate" hearing loss (PTA > 40 dB), while six had a "mild" hearing loss (PTA < 40 dB).

Dual-Task Paradigm • The dual-task paradigm consisted of three conditions: primary task alone, secondary task alone, and dual-task. The primary task required children to listen to and repeat monosyllabic words presented in noise. Speech stimuli were presented using a custom MATLAB program and routed to an audiometer (GSI 61) to adjust speech levels. The amplified speech was presented from a single loudspeaker at 0° azimuth, at a distance of 1.5 meters from the child.

Primary Task Conditions • Nine word lists (three, 10-word lists × 3 SNRs; 90 words total) were administered in the primary task conditions. Word-list order was held constant and SNR condition was counterbalanced across participants to avoid order effects. Prior to the presentation of the auditory stimuli, a green circle with a cartoon ear was shown on the computer screen to prompt the child to listen for the upcoming word. Children were

TABLE 2. Total Number of Participants Who Completed Testing in Each of the Three SNR Combinations

Test SNRs (in dB)	CNH	CHL Unaided Only	CHL Aided Only	CHL Unaided and Aided	Total
–4, 0, +4	37	16	7	13	73
0, +4, +8	0	3	3	3	9
+4, +8, +12	0	8	8	7	23

encouraged to guess if they were unsure of the word presented. Children's verbal responses were recorded using a head-worn microphone and saved for later analysis.

Secondary Task Conditions • The secondary task was a measure of visual response time. Children were instructed to monitor a computer screen located directly below the front loudspeaker for the presence of a brief (125 ms), randomly presented, visual target (a 24 × 18 cm white rectangle against a gray background), and to press a button as quickly as possible when it appeared. Prior research has shown that auditory background noise can influence response times on visual tasks and that this influence can vary depending on the characteristics of the background noise (Bendixen et al., 2010; Fischer et al., 2010; Stoffels & van der Molen, 1988). Therefore, to obtain a reliable estimate of the baseline cognitive resources required for secondary task performance in our test conditions, primary task auditory stimuli were also presented during the secondary task assessment. Secondary task testing was conducted at the same SNRs used during the primary and dual-tasks, but the child was instructed to ignore the words and respond as quickly as possible to the visual target. Six, 10-word lists (two lists × three SNRs) were used to obtain baseline performance on the secondary task. Eight trials in each 10-word list, contained one (four trials) or two (four trials) visual target(s). Two trials did not contain a visual target resulting in a total of 12 visual targets in each 10-word block. When two visual stimuli were presented during a single word trial they were separated in time by 1–2 seconds. The eight trials containing visual flashes, as well as the number of flashes (one or two) during a trial, were randomly selected. When present, a visual target could occur before, during, or after the presentation of the word—with no more than two targets presented within a single trial. Thus, a total of 72 visual targets (24 in each SNR) were randomly presented during the six, 10-word lists for the secondary task. Responses to these stimuli were used to determine baseline secondary task RTs (single-task visual RT). RTs to visual stimuli were recorded using MATLAB. Due to a programming error, we were unable to obtain reliable data regarding response accuracy to visual targets.

Dual-Task Conditions • For the dual-task, children were asked to repeat back the words in noise (primary task) while remaining vigilant for the visual target (secondary task). Nine additional 10-word lists and 108 visual targets (36 in each SNR) were presented in the dual-task condition. Children were not asked to prioritize one task over the other, given that this strategy has been shown to be ineffective for this particular age-group (Choi et al., 2008).

Children were given at least one practice session consisting of a 10-word list presented at the Easy SNR in each of the three conditions (primary, secondary, and dual-task). This practice list was not repeated during experimental testing. During the secondary and dual-task practice conditions, children were required to correctly respond to at least 80% of the secondary task trials. For the primary and dual-task practice, children were required to attempt to repeat at least 8 of the 10 words.

Word Recognition • Responses from the primary and dual-task conditions were recorded using a head-worn microphone and were transcribed off-line by trained graduate students. Responses were scored as correct or incorrect at the word level. Responses were deemed correct if all phonemes were correctly recalled, and no additional phonemes uttered (e.g., a response of "cats" would be incorrect if the stimulus was "cat"). Two independent reviewers transcribed responses, with discrepant scores resolved by a third trained reviewer.

Measures of Listening Effort

Dual-Task Visual RTs • The time elapsed from the visual target flash to the child's button press in the secondary task was calculated automatically using MATLAB. Based on recommendations in Gagné et al. (2017), a proportional dual-task visual RT was calculated using the following formula: baseline RT—dual RT/baseline RT × 100. "Baseline RT" refers to secondary task RT when the visual monitoring task was performed in isolation. "Dual RT" refers to secondary task RT in the dual-task condition. Dual-task visual RTs therefore represent the proportional RT change in secondary task performance. This method of quantifying listening effort helps to account for the influence of potential baseline differences in secondary task performance across participants. For ease of interpretation, scores were subtracted from zero (i.e., inverted) so that higher scores indicate increased listening effort.

Verbal RTs • Using audio recordings of the children's responses in the primary task condition, we calculated the time elapsed from the onset of the stimulus to the onset of the response. Trained research assistants used a custom-made MATLAB program to manually select the onset of each child's response in each trial via visual examination of the time waveform and auditory confirmation of the recording. Incidental sounds (mouth noises, room noise, and so on) unrelated to speech production were not included as part of the participant's response when determining the response onset time. In situations where the participant's response was halted or self-corrected, the response onset time analyzed was the onset of the self-corrected, second utterance. Coders had the opportunity to identify trials in which they felt further review from a second coder was needed. In the event that the second coder's selected time showed a discrepancy of > 25 ms from the initial coder's value, the two coders met in-person for consensus coding. Of the 247 flagged trials, only 4.86% required consensus coding (i.e., a discrepancy of > 25 ms was present between coders). Periodic reliability checks were conducted to compare different coders' analyses of a subset of trials. For every condition analyzed by one coder, a separate coder randomly selected and analyzed 10% of responses from that same condition. Reliability between coders was calculated monthly to monitor for coder drift. Interrater reliability remained high ($r > 0.95$) for the duration of the data collection process.

Statistical Analyses

Results of analyses conducted over the entire data set are reported. Some previous listening effort studies report listening effort measures derived using data from correct word recognition responses only, as well as the entire data set (e.g., Lewis et al. 2016). Analysis of reaction times based on only correct response data revealed the same main effects and/or interaction effects as analysis over the entire data set in the current study. Thus, for simplicity, we report only data from the full data set (i.e., results based on both correct and incorrect responses). Listening effort (dual-task visual RTs and verbal RTs) data were analyzed using a Linear Mixed-effects Model (LMM). Analyses were conducted with SNR (Easy, Moderate, Hard) as a within-subjects factor and hearing status (CNH versus CHL Unaided; CNH versus CHL Aided) as a between-subjects factor. A separate within-subjects analysis was then conducted to assess the effect of amplification on the subset of CHL ($n = 23$) who completed the experiment in both aided and unaided conditions. Because our groups differed significantly in age, language, and nonverbal intelligence (Table 1), all three were added as covariates for any between-subjects analyses. Random effects were included to account for individual variance in intercepts and slopes for any within-subject factors (SNR and amplification). All analyses were conducted in R (R Development Core Team, 2011). The "afex" package was used for all mixed-effects modeling analyses (Singmann et al., 2015).

It is well documented that raw RT and binary response data are often skewed and therefore not normally distributed, rendering parametric statistical analyses inappropriate (Lo & Andrews, 2015; Ratcliff, 1979; Whelan, 2008). Generalized Linear Mixed-effects Models (GLMMs) were therefore used for the statistical analysis of verbal RTs and primary task word recognition data. A GLMM is a form of LMM that allows the researcher to specify the response distribution of the dependent variable. For verbal RTs analysis, a "gamma" distribution was specified in the model, as this distribution more closely approximates the surface characteristics of raw RT data; a unimodal skew with all responses greater than 0 (Lo & Andrews, 2015). For primary task performance data, a "binomial" distribution was specified to capture the binary nature of the response data (i.e., "0" for incorrect, and "1" for correct verbal responses). This approach to analyzing non-normal data is advocated in Baayen et al. (2008) and was also used in the analysis of verbal RTs and speech recognition performance data in Lewis et al. (2016). *p* Values for both LMM and GLMM analyses were obtained using likelihood ratio tests, as recommended in Barr et al. (2013) for data with a large number of observations and subjects. Post hoc pairwise comparisons were conducted using least-squares means testing from the "lsmeans" package in R (Lenth & Hervé, 2014). The "Tukey HSD" method for *p* value adjustment of multiple-comparison testing was applied.

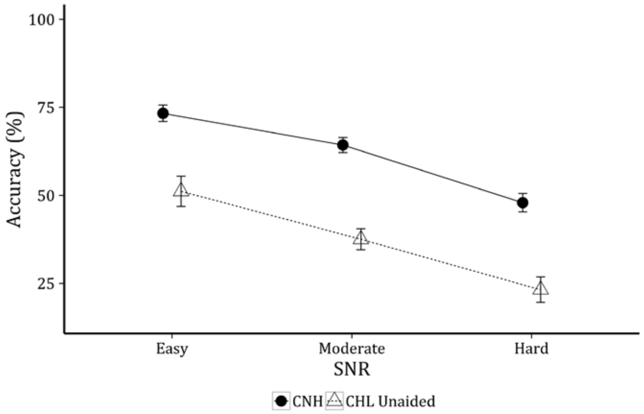

Fig. 3. Across-group mean performance accuracy (%) scores (± 95% CIs) as a function of hearing status (CNH, filled black circles; CHL Unaided, open triangles) and SNR (Easy, Moderate, Hard).

RESULTS

CNH Versus CHL Unaided

Word Recognition • Figure 3 displays word recognition performance as a function of SNR and hearing status. Analysis revealed a significant main effect of hearing status (χ^2 [1] = 36.56, $p < 0.001$), with word recognition scores significantly higher for the CNH ($M = 62\%$, $SD = 7\%$) compared with CHL Unaided ($M = 37\%$, $SD = 9\%$). The analysis also revealed a main effect of SNR (χ^2 [2] = 110.87, $p < 0.001$) but no interaction (χ^2 [2] = 2.73, $p = 0.26$). Post hoc tests comparing SNR conditions collapsed across hearing status revealed a systematic increase in word recognition as the SNR improved (Easy > Moderate, $\beta = 0.51$, $z = 6.87$, $p < 0.001$; Moderate > Hard, $\beta = 0.74$, $z = 9.94$, $p < 0.001$). Mean word recognition performance improved in Easy versus Moderate by 11%, and an additional 15% in Moderate versus Hard. Overall, word recognition was significantly better in the single ($M = 52\%$, $SD = 9\%$) versus dual-task ($M = 47\%$, $SD = 9\%$) condition (χ^2 [1] = 6.33, $p = 0.01$). However, there was no interaction between task (single versus dual) and hearing status (χ^2 [1] = 1.54, $p = 0.22$) or between task and SNR (χ^2 [2] = 2.10, $p = 0.35$). There was also no significant three-way interaction between task, hearing status, and SNR (χ^2 [2] = 1.27, $p = 0.53$; in situations where there is a significant change in primary task performance between single- and dual-task conditions, Gagné et al. (2017) suggest analyzing the proportional change in primary task performance from single- to dual-conditions also as a potential marker of listening effort. Analysis on proportional change data yielded similar findings to the dual-task visual RT data (all *p*s > 0.05).

Dual-Task Visual RTs • Figure 4 (left panel) displays dual-task visual RTs (i.e., the percentage change in dual RT versus baseline RT performance) as a function of SNR and hearing status (CNH versus CHL Unaided). Analysis of dual-task visual RTs revealed no main effect of hearing status (χ^2 [1] = 0.67, $p = 0.41$) or SNR (χ^2 [2] = 2.35, $p = 0.31$) and no interaction (χ^2 [2] = 1.27, $p = 0.53$). Analyses on baseline RTs revealed no significant main effects of hearing status or SNR, and no interaction (all *p*s > 0.05).

Verbal RTs • Figure 4 (right panel) displays verbal RTs as a function of SNR and hearing status (CNH versus CHL Unaided). Analysis revealed a significant main effect of hearing status (χ^2 [1] = 16.58, $p < 0.001$), with verbal RTs significantly slower in CHL Unaided compared with CNH. The analysis also revealed a main effect of SNR (χ^2 [2] = 26.22, $p < 0.001$) but no interaction (χ^2 [2] = 0.10, $p = 0.95$). Post hoc tests comparing SNR conditions collapsed across hearing status revealed that verbal RTs became faster as the SNR improved. However, while the difference between Hard and Moderate was significant

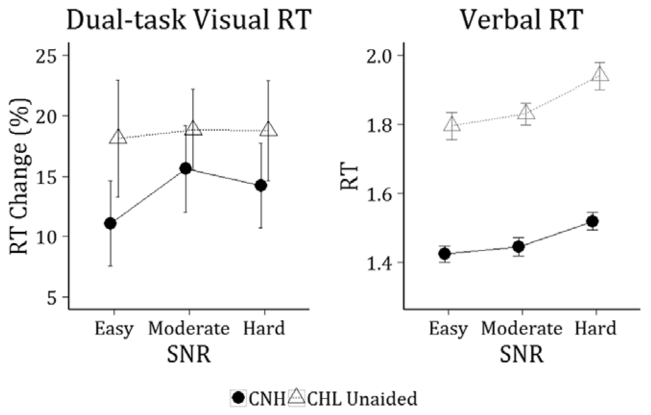

Fig. 4. Across-group means (± 95% CIs) for dual-task visual RTs (left panel) and verbal RTs (right panel; in seconds) as a function of hearing status (CNH, filled black circles; CHL Unaided, open triangles) and SNR (Easy, Moderate, Hard).

($\beta = 0.03$, $z = 4.25$, $p < 0.001$), the difference between Moderate and Easy did not reach significance ($\beta = 0.01$, $z = 1.88$, $p = 0.15$).

CNH Versus CHL Aided

Word Recognition • Figure 5 displays word recognition performance accuracy scores as a function of SNR and hearing status. Analysis revealed a significant main effect of hearing status (χ^2 [1] = 60.31, $p < 0.001$), with word recognition scores significantly higher in CNH (M = 62%, SD = 7%) compared with CHL Aided (M = 33%, SD = 8%). The analysis also revealed a main effect of SNR (χ^2 [2] = 126.33, $p < 0.001$) but no interaction (χ^2 [2] = 4.30, $p = 0.12$). Post hoc tests comparing SNR conditions collapsed across hearing status revealed a systematic increase in word recognition as the SNR improved (Easy > Moderate, $\beta = 0.48$, $z = 7.23$, $p < 0.001$; Moderate > Hard, $\beta = 0.82$, $z = 11.17$, $p < 0.001$). Mean word recognition improved in Easy versus Moderate by 11 percentage points, and an additional 16 percentage points in Moderate versus Hard. Overall, word recognition was significantly better in the single- (M = 47%) versus dual-task (M = 44%) condition (χ^2 [1] = 6.71, $p = 0.01$). However, there was no interaction between task (single versus dual) and hearing status (χ^2 [1] = 1.31, $p = 0.25$) or between task and SNR (χ^2 [2] = 3.33, $p = 0.19$). There was also no significant three-way interaction between task, hearing status, and SNR (χ^2 [2] = 1.35, $p = 0.51$).

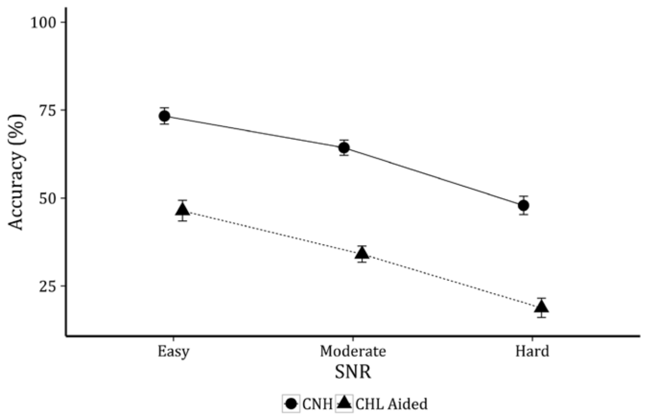

Fig. 5. Across-group mean performance accuracy (%) scores (± 95% CIs) as a function of hearing status (CNH, filled black circles; CHL Aided, filled triangles) and SNR (Easy, Moderate, Hard).

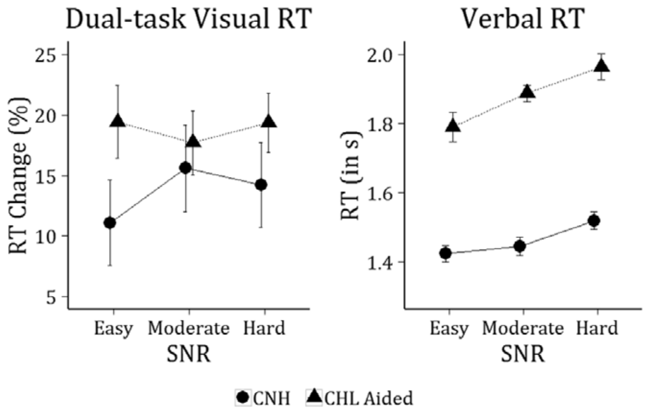

Fig. 6. Across-group means (± 95% CIs) for dual-task visual RTs (left panel) and verbal RTs (right panel; in seconds) as a function of hearing status (CNH, filled black circles; CHL Aided, filled triangles) and SNR (Easy, Moderate, Hard).

Dual-Task Visual RTs • Figure 6 (left panel) displays dual-task visual RTs as a function of SNR and hearing status (CNH versus CHL Aided). Analysis of dual-task visual RTs revealed no main effect of hearing status (χ^2 [1] = 2.01, $p = 0.16$) or SNR (χ^2 [2] = 0.46, $p = 0.79$) and no interaction (χ^2 [2] = 4.93, $p = 0.09$). Analyses on baseline RTs revealed no significant main effects of hearing status or SNR, and no interaction (all ps > 0.05).

Verbal RTs • Figure 6 (right panel) displays verbal RTs as a function of SNR and hearing status (CNH versus CHL Aided). Analysis revealed a significant main effect of hearing status (χ^2 [1] = 18.33, $p < 0.001$), with verbal RTs significantly slower in CHL Aided compared with CNH. The analysis also revealed a main effect of SNR (χ^2 [2] = 29.84, $p < 0.001$) but no interaction (χ^2 [2] = 2.28, $p = 0.32$). Post hoc tests comparing SNR conditions collapsed across hearing status revealed that verbal RTs became significantly faster as the SNR improved (Easy > Moderate, $\beta = 0.02$, $z = 3.16$, $p = 0.004$; Moderate > Hard, $\beta = 0.03$, $z = 4.06$, $p < 0.001$).

CHL Unaided Versus CHL Aided

Word Recognition • Figure 7 displays unaided and aided word recognition scores as a function of SNR for the CHL. Analysis revealed no significant main effect of amplification (χ^2 [1] = 0.47, $p = 0.49$). A significant main effect of SNR was found (χ^2 [2] = 50.56, $p < 0.001$), but no interaction (χ^2 [2] = 1.78, $p = 0.41$). Post hoc tests comparing SNR conditions collapsed across amplification revealed a systematic increase in word recognition as the SNR improved (Easy > Moderate, $\beta = 0.60$, $z = 6.36$, $p < 0.001$; Moderate > Hard, $\beta = 0.78$, $z = 8.19$, $p < 0.001$). Mean word recognition performance improved in Easy versus Moderate by 14 percentage points, and an additional 14 percentage points in Moderate versus Hard. Overall, word recognition was significantly better in the single- (M = 36%) versus dual-task (M = 32%) condition (χ^2 [1] = 4.18, $p = 0.04$). However, there was no interaction between task (single versus dual) and amplification (χ^2 [1] = 0.29, $p = 0.59$) or between task and SNR (χ^2 [2] = 1.16, $p = 0.56$). There was also no significant

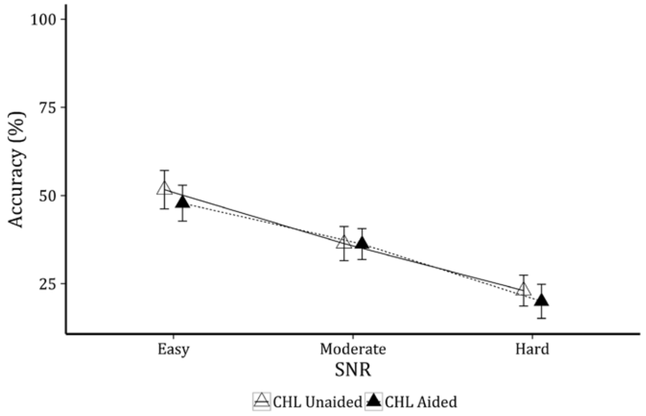

Fig. 7. Across-group mean performance accuracy (%) scores (± 95% CIs) as a function of amplification (CHL Unaided, open triangles; CHL Aided, filled triangles) and SNR (Easy, Moderate, Hard).

three-way interaction between task, amplification, and SNR (χ^2 [2] = 1.18, $p = 0.55$).

Dual-Task Visual RTs • Figure 8 (left panel) displays dual-task visual RTs as a function of SNR and amplification (CHL Unaided versus CHL Aided). Analysis of dual-task visual RTs revealed a larger increase in listening effort when unaided; however, the main effect of amplification was not significant (χ^2 [1] = 3.04, $p = 0.08$). Likewise, the main effect of SNR (χ^2 [2] = 2.51, $p = 0.28$) and the interaction effect (χ^2 [2] = 5.33, $p = 0.07$) were not significant. Analyses of unaided and aided baseline RTs as a function of SNR revealed no significant main effects or interactions (all ps > 0.05).

Verbal RTs • Figure 8 (right panel) displays verbal RTs as a function of SNR and amplification (CHL Unaided versus CHL Aided). Analysis revealed a significant main effect of SNR (χ^2 [1] = 11.86, $p = 0.003$), but no main effect of amplification (χ^2 [2] = 0.96, $p = 0.33$), and no interaction (χ^2 [2] = 0.72, $p = 0.70$). Post hoc tests comparing SNR conditions collapsed across amplification revealed that verbal RTs became faster as the SNR improved. However, while the difference between Hard and Moderate was significant ($\beta = 0.02$, $z = 3.39$, $p = 0.002$), the difference between Moderate and Easy did not reach significance ($\beta = 0.01$, $z = 2.04$, $p = 0.10$).

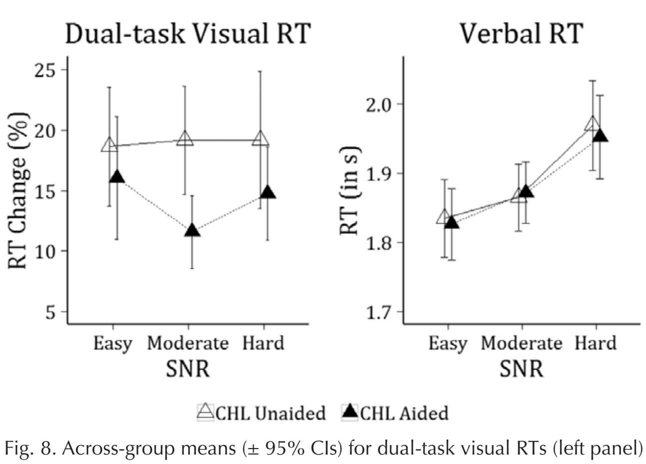

Fig. 8. Across-group means (± 95% CIs) for dual-task visual RTs (left panel) and verbal RTs (right panel; in seconds) as a function of amplification (CHL Unaided, open triangles; CHL Aided, filled triangles) and SNR (Easy, Moderate, Hard).

DISCUSSION

The goal of this study was to examine the effects of SNR, hearing loss, and amplification on listening effort in school-age CNH and CHL. We quantified listening effort using two behavioral measures: verbal RTs and dual-task visual RTs. We predicted slower verbal RTs and visual RTs in more challenging SNRs across both CNH and CHL (hypothesis 1). Verbal RTs were found to be slower in more challenging SNRs for both CNH and CHL, but there was no effect of SNR on dual-task visual RTs for CNH or CHL (Figs. 4, 6, 8). We also predicted slower dual-task visual RTs (i.e., greater dual-task cost) and verbal RTs in CHL versus CNH (hypothesis 2). Overall, verbal RTs were significantly slower in CHL versus CNH. However, no effect of hearing status (CNH versus CHL) was found on dual-task visual RTs. Finally, we predicted slower dual-task visual RTs and verbal RTs in unaided versus aided conditions (hypothesis 3); however, we found no effect of amplification (CHL Aided versus CHL Unaided) on either dual-task visual RTs or verbal RTs. This is the first study to compare verbal RTs and dual-task measures of listening effort in the child population. The current findings extend the literature by demonstrating an effect of hearing loss on verbal RTs in school-age children.

Effect of SNR on Listening Effort

Previous research demonstrates that verbal RTs are sensitive to the effects of SNR (Gustafson et al., 2014; Lewis et al., 2016) in school-age children. Consistent with this previous work, we found that verbal RTs were generally slower in more challenging SNRs for both CNH and CHL. Previous studies using the dual-task paradigm, however, have shown mixed results when determining the effects of SNR on listening effort in children (Howard et al., 2010; Hicks and Tharpe, 2002). Findings from the current study corroborate Hicks and Tharpe (2002) by showing no significant main effect of SNR on dual-task visual RTs across both CNH and CHL.

Discrepant results using the dual-task paradigm could reflect the substantial differences in task methodologies across studies. For instance, a variety of different secondary tasks have been used in previous studies using the dual-task paradigm to investigate listening effort (Gagne, et al., 2017). Digit recall was used as the secondary task by Howard et al. (2010), who found increased listening effort in more negative SNRs. Conversely, the secondary task used in the current study and by Hicks and Tharpe (2002) was a simple button press in response to a visual probe. Two recent studies have examined the effect of secondary task complexity (deep versus surface-level processing) on sensitivity to listening effort in children using the dual-task paradigm (Picou et al., 2017; Hsu et al., 2017). Neither study investigated digit recall as a secondary task, as in Howard et al. (2010). However, their findings suggest that secondary tasks requiring deeper levels of semantic processing (e.g., extracting word meaning) are no more sensitive to the effects of noise than simple (e.g., response to visual probe) secondary tasks. In contrast, relatively few studies have opted to use verbal RTs as a measure of listening effort in children (Lewis et al., 2016; Gustafson et al., 2014). Our finding of increased listening effort with worsening SNR is consistent with these previous studies.

Another potential source of discrepancy among studies examining the effect of SNR on listening effort using dual-task paradigms is the performance level in the primary task. It has been proposed that an effort "threshold" exists whereby effort generally increases until a task becomes too difficult, at which point effort will likely decrease (Pichora-Fuller et al., 2016). Consistent with this notion, research investigating listening effort using pupillometry in adults with hearing loss shows that listening effort plateaus, and even decreases, when a task starts to exceed an individual's cognitive capacity due to increasing acoustic demands (Ohlenforst et al., 2017b; Zekveld et al., 2011). A recent study by Wu et al. (2016) investigating the psychometric functions of dual-task paradigms in adult listeners suggests that listening effort is likely to be maximal at SNRs corresponding to a word recognition score of 30–50% correct. That is, secondary task performance decreased (indicating more listening effort) until speech recognition scores reached 30–50% accuracy. When primary task performance was better, or poorer, than 30–50%, secondary task performance increased (indicating a reduction in listening effort). The results of our study are consistent with this pattern. Specifically, the favorable SNRs used by Hicks and Tharpe (2002), who found no effect of SNR on listening effort, yielded average word recognition performance > 85%. Conversely, speech recognition in the study by Howard et al. (2010), where significant effects of SNR on listening effort were found, averaged between 50% and 80% correct. The poor word recognition scores in our study suggest that the SNRs and reverberant conditions were so difficult that some participants might have experienced cognitive overload and/or disengaged from the primary task—resulting in no effect of SNR on listening effort. This is especially apparent in our finding that dual-task visual RTs did not differ significantly (and in some cases even appeared to improve numerically) between Moderate and Hard SNR conditions.

Effect of Hearing Loss on Listening Effort

Previous studies that have examined listening effort in CHL versus CNH show discrepant results (Hicks & Tharpe, 2002; Lewis et al., 2016; McFadden & Pittman, 2008; Steel et al., 2015). Some studies report overall increased listening effort in CHL versus CNH as revealed using the dual-task paradigm (Hicks & Tharpe, 2002) as well as behavioral and physiological measures (Steel et al., 2015). On the other hand, both verbal RTs (Lewis et al., 2016) and dual-task paradigms (McFadden & Pittman, 2008) have also shown no overall effect of hearing loss on listening effort. Findings from the current study broadly support Hicks and Tharpe (2002) and Steel et al. (2015) by showing an effect of hearing loss on listening effort. However, verbal RTs (and not dual-task visual RTs) appear to be more sensitive to the detrimental effect of hearing loss.

The lack of an effect of hearing loss on dual-task visual RTs is consistent with McFadden and Pittman (2008) but not Hicks and Tharpe (2002). Potential reasons for these discrepant results are discussed in the previous subsection (e.g., the possibility of an effort "threshold" and/or differences in task methodology). Verbal RT findings from the current study are not consistent with Lewis et al. (2016) who used a similar verbal RT paradigm to assess the effect of hearing loss on listening effort in school-age children. This could relate to differences in the range of hearing loss severity included across studies. That is, CHL in Lewis et al. (2016) had relatively less severe (unilateral or mild bilateral) hearing loss than the CHL included in the current study, who had mild-to-moderate bilateral hearing impairments, as well as those studied by Steel et al. (2015), who were cochlear implant users. It has been suggested that listening effort could be increased in children with greater degrees of hearing loss but not in those with minimal or mild hearing loss (Lewis et al., 2016). To determine whether degree of hearing loss was related to listening effort in our sample of CHL ranging from mild to moderate, we conducted further correlation analyses. Specifically, we investigated the relationship between degree of hearing loss (PTA) and both measures of listening effort (dual-task visual RTs and verbal RTs). No significant correlations were found between PTA and measures of listening effort in CHL Unaided or CHL Aided (all *p*s > 0.05). These findings suggest that degree of hearing loss (PTA) does not have a strong effect on listening effort as measured using either verbal RTs or dual-task visual RTs. It is possible, however, that the current study sample (children with mild-to-moderate hearing loss) was not sufficiently heterogeneous to uncover systematic variance as a function of PTA. Furthermore, it is also possible that our decision to use different SNR combinations across children (to reduce floor/ceiling effects) may have reduced our ability to detect differences associated with degree of hearing loss.

Another potential reason for the discrepant verbal RT findings between this study and Lewis et al. (2016) relates to the choice of background noise and/or listening conditions. CHL are known to be more adversely affected by reverberation than CNH (Finitzo-Hieber & Tillman, 1978). It is possible that the combination of multi-talker babble and a reverberant test environment (RT60 ~ 450 ms) in the current study resulted in listening conditions that were more adverse, and thus required more effort, than those used in Lewis et al. (2016), which included testing in double-walled and sound-treated test booths. As such, mean word recognition was ~80% for CHL in Lewis et al. (2016) and ~50% in the current study for the CHL Unaided group in the Easy SNR condition. While adult studies have shown that the addition of reverberation to background noise causes increases in listening effort that are minimal (Rennies et al., 2014; Picou et al., 2016), further research is needed to determine if the combination of background noise and reverberation has a greater impact on listening effort in children than adults, particularly those with hearing loss.

Effect of Amplification on Listening Effort

Although research investigating the impact of amplification on listening effort in the adult population has increased rapidly in recent years (see Ohlenforst et al. 2017a, for a review), studies investigating the effect of assistive hearing technology on listening effort in children are relatively sparse (Gustafson et al., 2014; Stelmachowicz et al., 2007). Gustafson et al. (2014) found that verbal RTs were slower when digital noise reduction was switched off versus when it was switched on, suggesting reduced listening effort associated with the use of digital noise reduction technology. In the current study, no difference in dual-task visual RTs or verbal RTs were found when CHL were aided versus unaided. The current findings show a trend toward a larger dual-task visual RTs (i.e., increased listening effort) in unaided versus aided conditions (Fig. 8); however, the effect of amplification failed to reach statistical significance ($p = 0.08$). Although dual-task visual RTs were relatively more sensitive than verbal RTs in demonstrating some aided benefit in CHL, overall the use of hearing aids did not reduce listening effort as hypothesized.

There are several possible reasons for the lack of any statistical differences between aided and unaided listening in terms of word recognition and/or listening effort. First, any potential benefit from amplification in terms of increased audibility (A paired t-test was conducted based on data from CHL for whom we had both aided and unaided SII estimates of audibility (n = 20). As expected, SII estimates of audibility were significantly higher (i.e., better) in the aided versus the unaided condition, $t(19) = 5.38$, $p < 0.001$) was minimized in the noisy and reverberant test conditions, particularly in the more challenging SNR conditions. Because children's hearing aid fittings were "as worn" and not set to match prescriptive targets, it is feasible that aided benefit for some children was not optimal. That is, the quality of each child's hearing aid fitting may have influenced listening effort findings. To examine the relationship between adequacy of the hearing aid fitting and listening effort, we conducted further correlation analyses between better-ear average RMS error (i.e., match to DSL targets) and each measure of listening effort. Overall, we found no significant correlations between fit-to-target RMS error and dual-task visual RTs ($r = -0.11$, $p = 0.50$) or verbal RTs ($r = -0.01$, $p = 0.97$). This suggests that any variations in match to DSL targets in our participants' "as worn" hearing aid fittings had a minimal effect on listening effort in the aided condition.

We also explored the impact of unaided/aided audibility on listening effort using SII values obtained from the Verifit. Correlation analyses revealed no significant association between unaided or aided SIIs and any measure of listening effort (all *p*s > 0.05). Likewise, unaided/aided changes in SII were not significantly correlated with changes in listening effort across both measures (all *p*s > 0.05). While we were able to investigate the relationship between audibility in quiet conditions (at 65 dB SPL) and listening effort, we were unable to examine the extent to which audibility in noise (i.e., the test conditions used in the study) impacted listening effort when CHL were aided versus unaided. It is plausible that the noisy and reverberant test conditions served to reduce aided/unaided differences in audibility in some CHL (i.e., the background noise elevated hearing thresholds to a similar level in aided and unaided conditions).

Dual-Task Visual RTs Versus Verbal RTs

Results from the current study suggest that verbal RTs are more sensitive to the effects of SNR and hearing loss than dual-task visual RTs in school-age children. This finding is consistent with a study investigating the sensitivity of dual-task visual RTs versus verbal RTs in the adult population, which showed that verbal RTs were more sensitive to the effect of speech intelligibility than dual-task visual RTs (Pals et al., 2015). In the current study, dual-task visual RTs showed relatively greater sensitivity to the effect of amplification than verbal RTs. However, the effect did not reach statistical significance and should therefore be interpreted with caution. Dual-task visual RTs also showed more variability around the mean than verbal RTs (see CIs in Figs. 4, 6, 8). This increased variability may be at least partly attributed to the use of a relative change value that includes pooled variance from both baseline and dual-task conditions. It may also be the case that extraneous factors (e.g., cognitive ability) contribute to overall performance variability during the dual-task experiment, which requires higher-level attentional and cognitive processing compared with speech recognition alone.

Verbal RT measures have shown promise in both the adult (Gatehouse & Gordon, 1990; Pals et al. 2015) and pediatric (Gustafson et al., 2014; Lewis et al., 2016, Steel et al., 2015) populations as a potential marker of listening effort. The dual-task paradigm has been used to demonstrate effects of SNR, hearing loss, and amplification in the adult population; see Gagné et al. (2017) for an extensive review. However, the interpretation of dual-task data is considered more problematic in pediatric populations (Choi et al., 2008; McFadden & Pittman, 2008). The ability to multi-task is clearly an important skill to master in classroom learning environments (e.g., listening to the teacher's instructions while taking notes; Howard et al. 2010). However, given the unpredictable nature of attention allocation in children and the dual-task method's reliance on assumptions regarding cognitive resource allocation, this approach may not be an ideal technique for objectively measuring listening effort in children. On the other hand, interpretation of verbal RT data is relatively straight-forward as it reflects a behavioral response to a single task without additional sources of distraction and reflects mental processing limited to the auditory modality. Verbal RTs may also represent a more viable alternative to dual-task and/or physiological measures of listening effort that can be more readily used as part of a test battery in the audiology clinic. For example, verbal RT information can be obtained based on simple speech recognition tests already routinely performed in the clinic, thus removing the need for additional testing (e.g., incorporating multiple tasks) and/or costly equipment (e.g., for physiological recordings). A reliable and sensitive measure of listening effort could help to identify CHL who are candidates for different types of rehabilitation (e.g., assistive devices or cognitive training) and could also be used to help optimize hearing aid settings.

Study Limitations and Future Directions

Several aspects of this study may limit the generalizability of results. Due to a programming error, we were unable to collect reliable accuracy data for responses to visual targets. In a recent study, Grieco-Calub et al. (2017) used the dual-task paradigm to assess multi-tasking abilities in a group of school-age CNH. The primary task was a speech recognition task. Speech was noise-band vocoded with 4, 6, or 8 spectral channels. The secondary task was a visual monitoring task. A decrement in secondary task (visual monitoring) accuracy, but not RTs, was found in the more challenging (4 and 6 spectral channel) listening conditions. It is therefore possible that a dual-task decrement in secondary task performance accuracy may have been present in the current study. While this may be the case, other dual-task studies in the literature show that secondary task accuracy and secondary task RTs often show a similar pattern of results (Anderson Gosselin & Gagne, 2011; Fraser et al., 2010). Because dual-task visual RTs (and not accuracy) are more commonly used in the literature to assess listening effort, an examination of their sensitivity was warranted.

Given the limited experimental control over hearing aid settings in the current study, we are unable to draw firm conclusions regarding the effect of well-fitted amplification on listening effort in children. Although the main effect of SNR on listening effort shown here suggests that better audibility yields less listening effort, we found limited evidence that SII estimates of audibility (in quiet) or hearing aid fitting accuracy (fit-to-target RMS error) influence listening effort in CHL. This suggests that there are likely additional factors (independent of audibility and hearing aid fitting) that contribute to increased listening effort in CHL versus CNH. Future research should examine if currently available methods of SNR improvement (e.g., directional microphones, remote microphone systems) reduce listening effort in CHL. The effect of different types of background noise (e.g., single- versus multi-talker masker) on verbal RTs should also be explored in school-age children. Previous research suggests that CHL may be especially susceptible to the negative effects of speech maskers (Leibold et al., 2013). In recent years, listening-related effort and fatigue have also been measured in pediatric populations using physiological techniques, such as pupillometry (McGarrigle et al., 2017) and event-related potentials (Key et al., 2017; Gustafson et al., 2018). Inclusion of physiological markers of listening effort may help to elucidate factors independent of audibility that influence listening effort. Future research should also consider the impact of the duration of hearing loss and/or the age at which a child is fitted with a hearing device on subsequent listening effort.

CONCLUSIONS

This is the first study to compare two behavioral measures of listening effort in children. Verbal RTs appear more sensitive than dual-task visual RTs to the effects of SNR and hearing loss in school-age children. These findings suggest that inhibiting distracting noise requires more cognitive resources when the background noise level is increased, and this processing speed decrement is generally more pronounced in CHL versus CNH. Given the fast-paced nature of typical classroom exchanges, any slowing in mental processing will likely have a detrimental impact on learning and language development; both of which rely heavily on accurate speech comprehension.

ACKNOWLEDGMENTS

The authors would like to acknowledge the contributions of Hilary Davis and the many graduate students who assisted in participant recruitment and data collection. The authors also thank Wes Grantham for the development of the coding software, as well as Nick Bennett, Elizabeth Geller, Stone Dawson, Quela Royster, and Beth Suba for their help with coding and reliability checks.

This work was supported by the Dan and Margaret Maddox Charitable Trust, and the Institute of Education Sciences (IES), U.S. Department of Education, through Grant R324A110266 to Vanderbilt University (Nashville, TN; awarded to Fred H. Bess, PI). The content expressed are those of the authors and do not necessarily represent the views of the Institute of Education Sciences, or the U.S. Department of Education.

The authors declare no other conflict of interest.

Address for correspondence: Ronan McGarrigle, Department of Psychology, University of Chester, Parkgate Road, Chester, United Kingdom. E-mail: r.mcgarrigle@chester.ac.uk.

Received June 9, 2017; accepted April 27, 2018.

REFERENCES

Amlani, A. M., & Russo, T. A. (2016). Negative effect of acoustic panels on listening effort in a classroom environment. *J Am Acad Audiol*, *27*, 805–815.

Anderson Gosselin, P., & Gagné, J. P. (2011). Older adults expend more listening effort than young adults recognizing speech in noise. *J Speech Lang Hear Res*, *54*, 944–958.

ANSI S3.5, 1997 (R2017). American National Standard Methods for the Calculation of the Speech Intelligibility Index. New York, NY: American National Standard Institute.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J Mem Lang*, *59*, 390–412.

Barr, D. J., Levy, R., Scheepers, C., et al. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang*, *68*, 255–278.

Bendixen, A., Grimm, S., Deouell, L. Y., et al. (2010). The time-course of auditory and visual distraction effects in a new crossmodal paradigm. *Neuropsychologia*, *48*, 2130–2139.

Bess, F. H., Gustafson, S. J., & Hornsby, B. W. (2014). How hard can it be to listen? Fatigue in school-age children with hearing loss. *J Educ Audiol*, *20*, 34–47.

Broadbent, D. E. (1958). *Perception and Communication*. New York, NY: Pergamon Press.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (2010). *TONI-4, Test of Nonverbal Intelligence*, Austin, TX: ProEd.

Choi, S., Lotto, A., Lewis, D., et al. (2008). Attentional modulation of word recognition by children in a dual-task paradigm. *J Speech Lang Hear Res*, *51*, 1042–1054.

Cox, R. M., Alexander, G. C., Gilmore, C. (1987). Development of the Connected Speech Test (CST). *Ear Hear*, *8*(5 Suppl), 119S–126S.

Finitzo-Hieber, T., & Tillman, T. W. (1978). Room acoustics effects on monosyllabic word discrimination ability for normal and hearing-impaired children. *J Speech Hear Res*, *21*, 440–458.

Fischer, R., Plessow, F., Kiesel, A. (2010). Auditory warning signals affect mechanisms of response selection: evidence from a Simon task. *Exp Psychol*, *57*, 89–97.

Fraser, S., Gagné, J. P., Alepins, M., et al. (2010). Evaluating the effort expended to understand speech in noise using a dual-task paradigm: the effects of providing visual speech cues. *J Speech Lang Hear Res*, *53*, 18–33.

Gagné, J. P., Besser, J., Lemke, U. (2017). Behavioral assessment of listening effort using a dual-task paradigm. *Trends Hear*, *21*, 2331216516687287.

Gatehouse, S., & Gordon, J. (1990). Response times to speech stimuli as measures of benefit from amplification. *Br J Audiol*, *24*, 63–68.

Gomes, H., Molholm, S., Christodoulou, C., et al. (2000). The development of auditory attention in children. *Front Biosci*, *1*, 108–120.

Grieco-Calub, T. M., Ward, K. M., Brehm, L. (2017). Multitasking during degraded speech recognition in school-age children. *Trends Hear*, *21*, 2331216516686786.

Gustafson, S., McCreery, R., Hoover, B., et al. (2014). Listening effort and perceived clarity for normal-hearing children with the use of digital noise reduction. *Ear Hear*, *35*, 183–194.

Gustafson, S. J., Key, A. P., Hornsby, B. W. Y., Bess, F. H. (2018). Speech-processing related fatigue in children with hearing loss: behavioral, subjective, and electrophysiological measures. *J Speech Lang Hear R, 61*, 1000–1011.

Hick, C. B., & Tharpe, A. M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *J Speech Lang Hear Res*, *45*, 573–584.

Houben, R., van Doorn-Bierman, M., Dreschler, W. A. (2013). Using response time to speech as a measure for listening effort. *Int J Audiol*, *52*, 753–761.

Howard, C. S., Munro, K. J., Plack, C. J. (2010). Listening effort at signal-to-noise ratios that are typical of the school classroom. *Int J Audiol*, *49*, 928–932.

Hsu, B. C., Vanpoucke, F., van Wieringen, A. (2017). Listening effort through depth of processing in school-age children. *Ear Hear*, *38*, 568–576.

Hughes, K. C., & Galvin, K. L. (2013). Measuring listening effort expended by adolescents and young adults with unilateral or bilateral cochlear implants or normal hearing. *Cochlear Implants Int*, *14*, 121–129.

Kahneman, D. (1973). *Attention and Effort*: Englewood Cliffs, NJ: Prentice-Hall Inc.

Key, A. P., Gustafson, S. J., Rentmeester, L., et al. (2017). Speech-processing fatigue in children: auditory event-related potential and behavioral measures. *J Speech Lang Hear Res*, *60*, 2090–2104.

Klatte, M., Lachmann, T., Meis, M. (2010). Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting. *Noise Health*, *12*, 270–282.

Larsen, J. B., & Blair, J. C. (2008). The effect of classroom amplification on the signal-to-noise ratio in classrooms while class is in session. *Lang Speech Hear Serv Sch*, *39*, 451–460.

Leibold, L. J., Hillock-Dunn, A., Duncan, N., et al. (2013). Influence of hearing loss on children's identification of spondee words in a speech-shaped noise or a two-talker masker. *Ear Hear*, *34*, 575–584.

Lenth, R. V., & Hervé, M. (2014). lsmeans: least-squares means. R package version 2.13. http://CRAN.R-project. org/package=lsmeans. Accessed January 1, 2017.

Lewis, D., Schmid, K., O'Leary, S., et al. (2016). Effects of noise on speech recognition and listening effort in children with normal hearing and children with mild bilateral or unilateral hearing loss. *J Speech Lang Hear Res*, *59*, 1218–1232.

Lo, S., & Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Front Psychol*, *6*, 1171.

Luna, B., Garver, K. E., Urban, T. A., et al. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Dev*, *75*, 1357–1372.

Mackersie, C. L., Boothroyd, A., Minniear, D. (2001). Evaluation of the Computer-assisted Speech Perception Assessment Test (CASPA). *J Am Acad Audiol*, *12*, 390–396.

McFadden, B., & Pittman, A. (2008). Effect of minimal hearing loss on children's ability to multitask in quiet and in noise. *Lang Speech Hear Serv Sch*, *39*, 342–351.

McGarrigle, R., Dawes, P., Stewart, A. J., et al. (2017). Measuring listening-related effort and fatigue in school-aged children using pupillometry. *J Exp Child Psychol*, *161*, 95–112.

Ohlenforst, B., Zekveld, A. A., Jansma, E. P., et al. (2017a). Effects of hearing impairment and hearing aid amplification on listening effort: a systematic review. *Ear Hear*, *38*, 267–281.

Ohlenforst, B., Zekveld, A. A., Lunner, T., et al. (2017b). Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation. *Hear Res*, *351*, 68–79.

Pals, C., Sarampalis, A., van Rijn, H., et al. (2015). Validation of a simple response-time measure of listening effort. *J Acoust Soc Am*, *138*, EL187–EL192.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., et al. (2016). Hearing impairment and cognitive energy: the Framework for Understanding Effortful Listening (FUEL). *Ear Hear*, *37*, 5S–27S.

Picou, E. M., Gordon, J., Ricketts, T. A. (2016). The effects of noise and reverberation on listening effort in adults with normal hearing. *Ear Hear*, *37*, 1–13.

Picou, E. M., Charles, L. M., Ricketts, T. A. (2017). Child-adult differences in using dual-task paradigms to measure listening effort. *Am J Audiol*, *26*, 143–154.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychol Bull*, *86*, 446–461.

Rennies, J., Schepker, H., Holube, I., et al. (2014). Listening effort and speech intelligibility in listening situations affected by noise and reverberation. *J Acoust Soc Am*, *136*, 2642–2653.

Ruscetta, M. N., Arjmand, E. M., Pratt, S. R. (2005). Speech recognition abilities in noise for children with severe-to-profound unilateral hearing impairment. *Int J Pediatr Otorhinolaryngol*, *69*, 771–779.

Sato, H., & Bradley, J. S. (2008). Evaluation of acoustical conditions for speech communication in working elementary school classrooms. *J Acoust Soc Am*, *123*, 2064–2077.

Scollie, S., Seewald, R., Cornelisse, L., et al. (2005). The desired sensation level multistage input/output algorithm. *Trends Amplif*, *9*, 159–197.

Singmann, H., Bolker, B., & Westfall, J. (2015). *afex: Analysis of Factorial Experiments*. R package version 0.13–145.

Steel, M. M., Papsin, B. C., & Gordon, K. A. (2015). Binaural fusion and listening effort in children who use bilateral cochlear implants: a psychoacoustic and pupillometric study. *PloS one*, e0117611. doi:10.1371/journal.pone.0117611.

Stelmachowicz, P. G., Lewis, D. E., Choi, S., et al. (2007). Effect of stimulus bandwidth on auditory skills in normal-hearing and hearing-impaired children. *Ear Hear*, *28*, 483–494.

Stoffels, E. J., & van der Molen, M. W. (1988). Effects of visual and auditory noise on visual choice reaction time in a continuous-flow paradigm. *Percept Psychophys*, *44*, 7–14.

Styles, E. (2006). *The Psychology of Attention*, Oxford, UK: Psychology Press.

Whelan, R. (2008). Effective analysis of reaction time data. *Psychol Rec*, *58*, 475–482.

Wiig, E. H., Secord, W., & Semel, E. M. (2004). *CELF Preschool 2: Clinical Evaluation of Language Fundamentals Preschool*: San Antonio, TX: Pearson, PsychCorp.

Wu, Y. H., Stangl, E., Zhang, X., et al. (2016). Psychometric functions of dual-task paradigms for measuring listening effort. *Ear Hear*, *37*, 660–670.

Zekveld, A. A., Kramer, S. E., Festen, J. M. (2011). Cognitive load during speech perception in noise: the influence of age, hearing loss, and cognition on the pupil response. *Ear Hear*, *32*, 498–510.