Multitasking During Degraded Speech Recognition in School-Age Children

Tina M. Grieco-Calub[1,2], Kristina M. Ward[1], and Laurel Brehm[3]

Abstract

Multitasking requires individuals to allocate their cognitive resources across different tasks. The purpose of the current study was to assess school-age children's multitasking abilities during degraded speech recognition. Children (8 to 12 years old) completed a dual-task paradigm including a sentence recognition (primary) task containing speech that was either unprocessed or noise-band vocoded with 8, 6, or 4 spectral channels and a visual monitoring (secondary) task. Children's accuracy and reaction time on the visual monitoring task was quantified during the dual-task paradigm in each condition of the primary task and compared with single-task performance. Children experienced dual-task costs in the 6- and 4-channel conditions of the primary speech recognition task with decreased accuracy on the visual monitoring task relative to baseline performance. In all conditions, children's dual-task performance on the visual monitoring task was strongly predicted by their single-task (baseline) performance on the task. Results suggest that children's proficiency with the secondary task contributes to the magnitude of dual-task costs while multitasking during degraded speech recognition.

Introduction

Speech recognition requires bottom-up sensory processing and recruitment of cognitive resources including selective attention, working memory, and lexical memory (e.g., Baddeley & Hitch, 1974; Pichora-Fuller, Schneider, & Daneman, 1995). The demand placed on these cognitive resources increases during the recognition of speech that has been degraded with signal processing, background noise, and listener hearing loss (e.g., Broadbent, 1958; Downs & Crum 1978; Pals, Sarampalis, & Başkent, 2013; Pichora-Fuller et al., 1995; Rabbitt, 1966; Rakerd, Seitz, & Whearty, 1996; Rönnberg, 2003; Rönnberg et al., 2013; Sarampalis, Kalluri, Edwards, & Hafter, 2009; Wild et al., 2012). As speech recognition demands increase, it is expected that individuals will reallocate cognitive resources to aid in resolving the degraded speech input. According to the limited capacity model (Kahneman, 1973), individuals have a finite pool of cognitive resources from which to allocate to any given task. Therefore, the reallocation of cognitive resources to resolve degraded speech input has an associated cost. Specifically, if the speech recognition task requires an increase in cognitive resource allocation,

there are fewer remaining resources, otherwise known as the cognitive spare capacity, to maintain performance on concurrent tasks. This has practical implications for individuals who are attempting to multitask during speech recognition, which is a common occurrence in daily life (e.g., driving while having a conversation; taking notes while listening to a teacher, etc.).

Multitasking during speech recognition reflects a common challenge for school-age children: Processing a visual scene or writing notes while a teacher is talking is typical for a general education classroom. Excessive classroom noise and reverberation, however, can degrade the fidelity of the teacher's speech (e.g.,

[1]The Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL, USA
[2]Interdepartmental Neuroscience Program, Northwestern University, Evanston, IL, USA
[3]Linguistics Department, Northwestern University, Evanston, IL, USA

Corresponding Author:
Tina M. Grieco-Calub, Northwestern University, 2240 Campus Drive, Room 2-246, Evanston, IL 60208, USA.
Email: tinagc@northwestern.edu

Crandall & Smaldino, 2000; Klatte, Hellbrück, Seidl, & Leistner; Shield & Dockrell, 2003). Compared with older children and adults, younger children require a higher fidelity speech input to preserve accurate perception (Eisenberg, Shannon, Martinez, Wygonski, & Boothroyd, 2000; Elliott et al., 1979; Johnson, 2000; Jones, Moore, & Amitay, 2015; Leibold & Buss, 2013; Papso & Blood, 1989; Wightman & Kistler, 2005). As a result, degraded auditory input may cause children to expend relatively greater cognitive resources to resolve the speech content. In addition, due to their cognitive immaturity, children may have fewer resources to allocate across multiple tasks or an inability to allocate those resources appropriately (Choi, Lotto, Lewis, Hoover, & Stelmachowicz, 2008; Irwin-Chase & Burns, 2000). Thus, school-age children are expected to have increased difficulty multitasking during degraded speech recognition.

Dual-task paradigms are a form of multitasking commonly used in experimental designs to quantify the allocation of cognitive resources during speech recognition. In these paradigms, individuals are instructed to prioritize the primary speech recognition task while they concurrently perform a secondary task that competes for domain-specific (e.g., auditory) and domain-general (e.g., executive function) cognitive processes, depending on the nature of the secondary task. If the primary and secondary tasks utilize overlapping cognitive resources, secondary-task performance is expected to decline as the primary task becomes more difficult (Kahneman, 1973). Dual-task paradigms, therefore, are uniquely useful in addressing both theoretical and practical questions about the cognitive costs of degraded speech recognition. Theoretically, they provide a way to test which cognitive processes may be involved during recognition of degraded speech. Practically, they provide a way to predict how performance on concurrent tasks changes as the speech recognition task becomes more difficult.

The extent to which children reallocate cognitive resources during degraded speech recognition is not well understood. Hicks and Tharpe (2002) evaluated children's spoken word recognition accuracy (auditory-verbal primary task) while also asking them to press a button in response to a light probe (visual-motor secondary task). Children were expected to show progressively slower reaction times to the light probe with decreasing signal-to-noise ratio (SNR) of the speech recognition task (i.e., quiet, +20 dB SNR, +15 dB SNR, +10 dB SNR). Contrary to the authors' predictions, there was no change in reaction time across the four conditions of the primary task, even though statistically significant declines in speech recognition were observed with decreasing SNR. One potential interpretation of this result is that the auditory-verbal and visual-motor nature of the primary and secondary tasks, respectively, did not use overlapping cognitive resources. If this is

true, then one might expect to observe greater declines in performance on a secondary task that requires similar processing as the primary task. Stelmachowicz, Lewis, Choi, and Hoover (2007), however, provide evidence against this idea. School-age children maintained performance on a dual-task paradigm that included a spoken word recognition task with variable spectral bandwidths (auditory-verbal primary task) and a digit span test (visual-verbal secondary task). Children showed no decrement in their digit span or in their speech recognition accuracy with decreasing spectral fidelity of the primary task.

One possibility is that children did not reallocate their cognitive resources during degraded speech recognition in Hicks and Tharpe (2002) and Stelmachowicz et al. (2007) because of their cognitive immaturity. Studies by Irwin-Chase and Burns (2000) and Choi et al. (2008) partially support this interpretation. In Irwin-Chase and Burns' (2000) study, children between 8 and 12 years of age were asked to alter their allocation of attention across primary and secondary tasks that were both visual per preassigned proportions. Example conditions included prioritizing one task over the other (100%–0%) or equally weighting attention on both tasks (50%–50%). Their results showed that the ability to allocate attention across two visual tasks appears to mature between 8 years and 11 years of age, but is not yet adult-like by 11 years of age. Consistent with this finding, Choi et al. (2008) found that similarly aged children performing concurrent speech recognition (auditory-verbal) and digit span (visual-verbal) tasks had difficulty allocating their attention to the task that they were instructed to prioritize. Regardless of instructions, children always prioritized the speech recognition task over the digit span test. Although both of these studies support the idea that the ability to control attention allocation matures within school-age children, they do not address whether attentional control might also vary with tasks that have less overlap in processing.

An alternate interpretation of the Hicks and Tharpe (2002) and Stelmachowicz et al. (2007) results is that the primary and secondary tasks in these studies were not sufficiently difficult to elicit cognitive resource reallocation, resulting in consistent secondary-task performance across only modestly difficult primary-task demands. Consistent with the latter interpretation, Howard, Munro, and Plack (2010) observed declines in children's digit span (visual-verbal secondary task) accuracy with decreasing SNR on a speech recognition task (auditory-verbal primary task). Consistent with the assumption of dual-task paradigms, poorer performance on the secondary task suggests a reallocation of cognitive resources to the primary task in the conditions with poorer SNR. A fundamental difference in the Howard et al. (2010) study from the others was that children experienced greater

declines in speech recognition accuracy (~30%–50%) across the conditions of the primary task when compared with performance of age-matched children in previous studies (<20%; Hicks & Tharpe, 2002; Stelmachowicz et al., 2007), which was likely due to the lower SNRs used in Howard et al. (2010). This observation suggests that children's reallocation of cognitive resources away from the secondary task is dependent on the degree to which the primary task is degraded. It is unclear, however, as to whether this also pertains to primary and secondary tasks that have less overlap in their processing demands.

The current study had two objectives. First, the study was designed to quantify school-age children's performance on a bimodal dual-task paradigm, whereby performance on a visual monitoring (secondary) task was quantified during a speech recognition (primary) task that became increasingly more difficult. Although there were myriad secondary tasks that could be utilized in dual-task paradigms, the visual monitoring task was chosen because it elicits competition from domain-general attentional resources rather than inducing domain-specific interference. Second, the study aimed to explore individual factors that influenced children's performance on the secondary task during the dual-task conditions. The predicted results included (a) declines in primary-task accuracy with decreasing spectral fidelity and (b) declines in secondary-task performance coincident with decreasing primary-task accuracy.

The design and analysis of the current study adopted the following assumptions of the dual-task paradigm: (a) Cognitive capacity is limited; (b) The primary and secondary tasks engage overlapping cognitive resources; (c) Individuals are capable of performing the secondary task at levels that are above chance performance during the baseline measures (i.e., in isolation); (d) Primary-task performance is unaffected by the addition of the secondary task; and (e) Cognitive resources are reallocated from the secondary task to the primary task as the primary task becomes more difficult.

Methods

Participants

Twenty-seven children (16 females) between 8 and 12 years old (9.6 ± 1.2 years, mean $\pm SD$) were recruited to participate in the study. All children were native English speakers, had normal or corrected-to-normal vision as measured by a Snellen eye chart, and had normal hearing in both ears. Criterion for normal hearing was pure-tone thresholds of $\leqslant 20$ dB HL at octave frequencies from 0.25 through 8 kHz as well as interoctave frequencies of 3 kHz and 6 kHz. Testing occurred within a single session that took no longer than 2 hours. Children were monetarily compensated for their time. All procedures were approved by the institutional review board at Northwestern University.

Stimuli

Speech stimuli. The speech stimuli consisted of a corpus of 180 Bench-Kowal-Bamford (BKB) phonetically balanced short sentences (e.g., "They are looking at the clock"; "Children like strawberries"; Bench, Kowal, & Bamford, 1979) produced by a male speaker. Sentences were noise-band vocoded using custom MATLAB software in order to generate four conditions that varied in spectral resolution. Noise-band vocoding was chosen for the current study because it enables degradation of the speech signal without adding another *object* to attend, which is the case when there is acoustic competition (i.e., background noise) and which may alter how cognitive resources are distributed during the task. During noise-band vocoding, each sentence was root-mean-square equalized at 65 dB SPL and digitally filtered into the appropriate number of frequency channels (i.e., 4-, 6-, or 8-channels). The output of each filtered waveform was half-wave rectified and low-pass filtered at 30 Hz using a fourth-order Butterworth filter with a slope of -24 dB per octave in order to extract the amplitude envelope of the signal. The envelope was then multiplied by a broadband noise carrier and passed through its respective bandpass filter. This process stripped the sentences of their original temporal fine structure and reduced spectral information while maintaining the slow-changing temporal features of the speech waveform. In addition to the three noise-band-vocoded conditions, testing incorporated an unprocessed condition in order to establish a baseline measure of speech recognition.

Visual stimuli. The visual stimuli were 260 individual grayscale illustrations (281 by 197 pixels) of familiar animate and inanimate objects (e.g., anchor, leaf, snail). The images were scaled to 3" by 4" and individually displayed on a computer monitor using custom software adapted from Wright et al. (2014) with MATLAB PsychToolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997).

Testing Apparatus

Testing was conducted in a sound-attenuating booth with only the participant present. Children sat at a desk in front of a Dell Ultrasharp U2413 24 inch monitor attached to an Apple wired USB keyboard. Stimuli were presented from a Macbook Pro laptop positioned externally to the sound booth. Speech stimuli were directly routed from the laptop to supra-aural Sennheiser

HD 25 SP headphones worn by the child, and the visual stimuli were displayed on the Dell computer monitor. Children indicated their response to visual stimuli by pressing a designated key on the keyboard.

Executive Function Assessment

Ward, Shen, Souza, and Grieco-Calub (2017) showed that older adults experienced greater declines in performance than younger adults on the same visual monitoring task as the one used in the current study during degraded speech recognition. The magnitude of secondary-task performance change was predicted by individuals' executive function skills. Given that the dual-task paradigm was bimodal (i.e., auditory-verbal primary task and visual-motor secondary task), these results suggest that a portion of domain-general cognitive resources were engaged in older adults' dual-task performance and raise the possibility that the same relation may be observed in children. Similar to Ward et al. (2017), children in the current study completed the Flanker Inhibitory Control and Attention Test (McDonald, 2014), which is a standardized psychometric measure of attention and inhibitory control available in the NIH Toolbox. In this task, a row of five arrows was displayed on a computer screen and the participant had to press either the leftward or rightward arrow key as quickly as possible to indicate the directionality of the middle arrow. Test trials consisted of intermixed congruent (i.e., all arrows facing the same direction) and incongruent (i.e., middle arrow facing the opposite direction) trials for a total of 20 trials. Scores were generated by the software program according to the following criteria (per the test's manual): (a) if children achieved accuracy scores of 80% or higher, accuracy and reaction time scores were combined to generate a composite scale score with a mean of 100 and standard deviation of 15; or (b) if children failed to achieve 80% accuracy, only the accuracy score was considered. Computed scores that were unadjusted for age were used in regression models described in the Results section. Children's age was added as an independent predictor variable in these models.

Procedures

Speech recognition (primary) task. Children were presented with BKB sentences of varying spectral fidelity binaurally via headphones at 65 dB SPL. Stimulus levels were confirmed using an ear simulator coupler attached to a sound level meter. There was a fixed silent period of 5 seconds following each sentence. Children were instructed to listen to each sentence and repeat it as accurately as possible; they were encouraged to guess if unsure of what they heard. Responses were scored live

with a point awarded for each of the three or four predetermined keywords within each sentence that children correctly repeated. Prior to testing, each child was familiarized to a preset list of 20 BKB sentences that were distinct from the test set. The familiarization list contained five sentences from each of the four listening conditions; spectral resolution decreased as the list progressed. If children's performance fell below the criterion score of 50% keywords correctly identified in the most degraded (4-channel) condition, they received an additional training block ($N = 20$ children). This block consisted of the same 20 sentences used during familiarization, all presented in the 4-channel condition. Feedback was only provided for the first ten sentences; children's performance on the final 10 sentences was used to determine whether they met the 50% correct criterion after the second training block. Following the second training block, 15 children scored $\geqslant 50\%$, 4 children scored between 40% and 50%, and 1 child scored 12.5%. The latter child performed within two standard deviations of the group performance in the 4-channel condition during the test conditions; therefore, all children were included in the analyses.
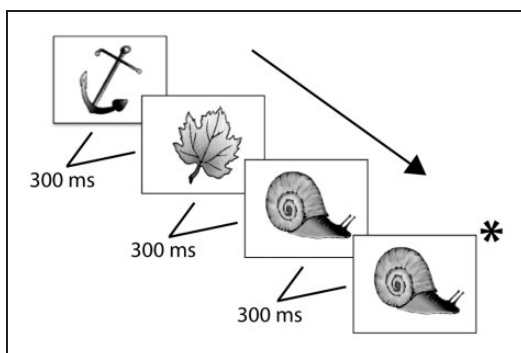
After familiarization, children were presented with 20 BKB sentences in each of the four listening conditions (i.e., unprocessed, 8-, 6-, and 4-channel noise-band vocoded) for a total of 80 sentences. Sentence lists were quasirandomly assigned to the listening conditions across participants to minimize an effect of sentence list. In addition, the presentation order of listening conditions was randomized across participants to minimize an effect of condition. As in the familiarization phase, responses were scored live with a point awarded for each of the sentence-dependent key words correctly repeated. Within each test condition, speech recognition accuracy was calculated as the proportion of keywords correctly identified out of the 62 possible keywords. Children completed the speech recognition task in each condition first in isolation and then concurrently with the secondary task (i.e., the dual-task paradigm).

Visual monitoring (secondary) task. This task was modeled after the rapid serial visual presentation task, which has been used historically as a measure of visual attention in the temporal domain (Potter & Levy, 1969). Modified versions of this task have been used successfully in other dual-task paradigms (e.g., Pals et al., 2013). In the current study, grayscale images were presented sequentially on a computer monitor at a rate of 300 milliseconds per image with an interstimulus interval of equivalent duration (Figure 1). Each image was randomly selected from a corpus of 260 images and shown on the 3" × 4" (281 by 197 pixels) center of the screen. Children were instructed to press a key as quickly as possible when the same picture occurred twice in a row

(*, Figure 1). In each block of images per condition, children saw approximately 206 images. Each sequence of images contained duplicate pictures that were randomly distributed within each block. Duplicate pictures comprised ~10% of trials per condition. Performance was quantified by *accuracy*, defined as the proportion of trials in which a repeated image elicited a key press from the total number of repeated images, as well as *reaction time*, or the duration of time between the display of the repeated image and the participant's key press. The majority of children (24 of 27) performed the visual monitoring task in isolation at two time points during the study, immediately prior to the dual-task conditions (Time Point 1) and immediately following the dual-task conditions (Time Point 2), to provide a measure of baseline performance and to rule out an effect of learning or fatigue. The remaining three children only have performance scores from Time Point 1 because their data were collected prior to the decision to add a second measure of single-task performance following the dual-task conditions. All children completed the visual monitoring task first in isolation (i.e., single task) and then in the dual-task conditions. The single-task, baseline condition was performed in the absence of unattended primary-task stimuli. The rationale for this design was to maximize children's performance on the visual monitoring task and to minimize any distraction that the speech stimuli of the primary task might cause (e.g., Choi et al., 2008). Thus, any change in performance on the visual monitoring task in the dual-task conditions would more accurately reflect the demands of the primary task in each condition.

Dual-task paradigm. In addition to the primary- and secondary-task baseline measurements, children performed the two tasks simultaneously as part of the dual-task paradigm. Childre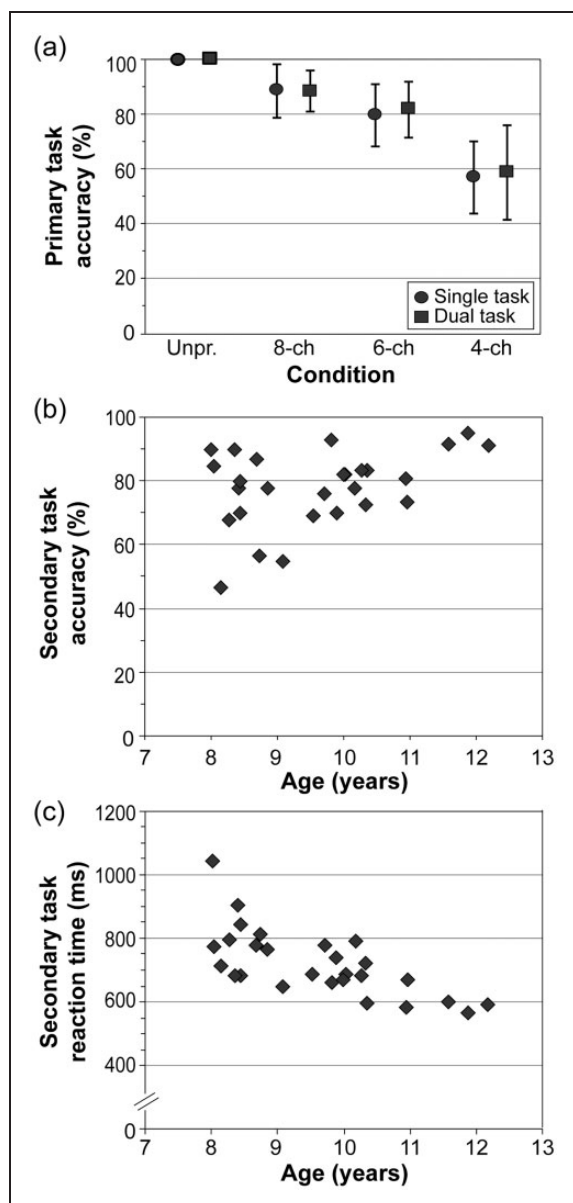n were instructed to listen to each sentence and repeat it aloud as accurately as possible while simultaneously watching the computer monitor for repeating images. They were explicitly told to prioritize the speech recognition task. Similar to the speech recognition task performed in isolation, the dual-task paradigm consisted of a total of 80 sentences presented across 20-sentence lists, 1 list per each listening condition. Similar to the single task, sentence lists and listening conditions were quasirandomized to listening conditions and order of presentation, respectively. Although all sentence lists stemmed from the same corpus, speech material was not redundant between the single and dual-task conditions. Similar to the speech recognition task, the visual monitoring task was administered continuously throughout the dual-task conditions. Repeated pictures could occur within any of the following three contexts relative to the speech recognition task: (a) during the auditory presentation of the target sentence; (b) during the participant's verbal recall of the target sentence; (c) during the remaining period of silence within the 5-second intersentence interval. Thus, although there were moments within the dual-task condition where participants were only actively performing one task, it is expected that the demand for shared cognitive resources was consistently maintained because of the randomness with which repeated images occurred. Performance scores for the speech recognition task (i.e., accuracy) and visual monitoring task (i.e., accuracy and reaction time) were recorded for each listening condition. Accuracy scores were folded-square-root transformed before statistical analysis to minimize floor and ceiling effects without requiring adjustment of extreme data values (i.e., 0% and 100%, respectively; Tukey, 1960).

Results

Single-Task Performance: Speech Recognition (Primary) Task

Speech recognition accuracy on the primary task performed in isolation is illustrated in Figure 2(a) (*circles*). A repeated-measures analysis of variance (ANOVA) showed a main effect of condition, $F(3, 78) = 240.6$, $p < 0.001$, $\eta^2 = 0.902$. Post hoc paired comparisons revealed that children's speech recognition accuracy significantly declined with each decrease in spectral fidelity (e.g., performance in the 8-channel condition was worse than in the unprocessed condition; performance in the 6-channel condition was worse than in the 8-channel condition, etc.; $p < .001$ following Bonferroni correction for multiple comparisons). This result is consistent with prior work showing a reduction in speech recognition in children with decreasing spectral fidelity of the stimulus (Eisenberg et al., 2000).

Figure 1. Temporal sequence of the visual monitoring task. The asterisk (*) represents a duplicate image that would prompt a key press from the child.

Figure 2. (a) Speech recognition (primary) task accuracy in isolation (*circles*) and during dual task conditions (*squares*). Mean ± *SD*. (b) Visual monitoring (secondary) task accuracy in isolation of each child as a function of chronological age. (c) Visual monitoring (secondary) task reaction time of each child as a function of chronological age.

Single-Task Performance: Visual Monitoring (Secondary) Task

The secondary task was a visual monitoring task that required children to decide whether each presented image matched the preceding image. In essence, it was a Go/No-Go task: Children identified each image as repeating (by pressing a key) or nonrepeating (by doing nothing). The visual monitoring task allowed for both

accuracy and reaction time to be calculated. Reaction times presented in the remaining analyses reflect those for correct responses only. Most children also exhibited false alarms, meaning that they indicated a repeating image (by key press) when the image did not repeat. The average false alarm rate (i.e., number of key presses divided by the number of nonrepeating pictures) was stable across conditions: unprocessed: $3.4\% \pm 5.9\%$, mean ± *SD*; 8-channel: $1.3\% \pm 1.8\%$; 6-channel: $2.3\% \pm 5.2\%$; 4-channel: $2.6\% \pm 4.2\%$. Given that the false alarm rate was low and not normally distributed, false alarms were not considered in the following analyses.

Single-task (baseline) accuracy and reaction time on the visual monitoring task were calculated from Time Point 1 for the first three children of the study (see Methods section). For the remaining 24 children who completed the visual monitoring task in isolation at Time Points 1 and 2, single-task (baseline) accuracy and reaction time were calculated as an average across the two time points. The rationale for this design was related to the observation that children's accuracy on the secondary task decreased slightly across the experiment—Time Point 1: $80.6\% \pm 10.3\%$ vs. Time Point 2: $72.8\% \pm 18.5\%$; $t(1,23) = 2.4$, $p < .05$. Reaction time was similar across the two time points—Time Point 1: $702\,ms \pm 86\,ms$ vs. Time Point 2: $732\,ms \pm 137$; $t(1,23) = -1.4$, $p = .18$. The statistical difference in accuracy between the two time points appears to be driven largely by two children whose difference scores (Time Point 1 – Time Point 2) were 41 and 51 points. Removal of these children's data from analysis resulted in no statistical difference in performance for the remaining 22 children, Time Point 1: $80.2\% \pm 10.5\%$; Time Point 2: $75.8\% \pm 15.7\%$; $t(1,21) = 1.7$, $p = .13$, suggesting that the majority of children performed the task with similar accuracy across the two time points. Thus, averaging performance across Time Points 1 and 2 provided a more stable "baseline" estimate of performance against which to compare performance from the dual-task conditions.

Figure 2(b) and (c) illustrates children's baseline accuracy and reaction time, respectively, on the visual monitoring task plotted against their chronological age. Inspection of the data revealed large individual variability in performance. Given that both age and executive function are believed to influence dual-task performance, separate linear regression models were tested to determine how each contributed to children's performance on the visual monitoring task in isolation. In addition to chronological age, Flanker scores (unadjusted for age; McDonald, 2014) were used to account for children's attention and inhibitory control. As a group, children scored within the normal range on the Flanker test, (103.6 ± 10.5) and children's scores did not statistically correlate with age ($R^2 = 0.36$, $p = .067$).

The overall fit of the model evaluating the influence of age and executive function on secondary-task accuracy was not statistically significant at an alpha of .05, $F(2, 24) = 2.79$, $p = .081$. The model fit improved when the Flanker score was removed as a predictor, $F(1, 25) = 5.04$, $p = .034$. The model with age as the only predictor variable—$\beta = 0.41$, $t(25) = 2.25$, $p = .034$, 95%, CI [0.05, 1.26]—accounted for 13.5% (adjusted R^2) of the observed variance in accuracy. This result suggests that > 85% of the variability observed in children's ability to correctly identify duplicate images on the visual monitoring task was unaccounted for in the data.

The overall fit of the model evaluating the influence of age and executive function on secondary-task reaction time—$F(2, 24) = 12.8$, $p < .001$—was better than the model predicting secondary-task accuracy, as it accounted for 47.5% (adjusted R^2) of the observed variance in reaction time. Again, age was the only significant predictor—$\beta = -0.647$, $t(24) = -4.2$, $p < .001$, 95% CI [−83.77, −29.04]; removal of the Flanker score from the model did not change model fit—$F(1, 25) = 24.4$, $p < .001$; adjusted $R^2 = 47.4\%$. Taken together, these results showed that performance on the visual monitoring task was influenced by age; however, the extent to which age predicted performance was dependent on how performance was quantified (i.e., accuracy vs. reaction time). Aspects of executive function quantified by the Flanker test (i.e., attention and inhibitory control) did not predict performance on the visual monitoring task in the children tested in the current study. This result contrasts findings from an earlier study, whereby Flanker scores predicted secondary-task performance in younger and older adults (Ward et al., 2017).

Dual-Task Performance: Speech Recognition (Primary) Task

Average accuracy on the speech recognition task during the dual task is illustrated in Figure 2(a) (*squares*) next to single-task performance (*circles*). A 4 (Condition) × 2 (Task) repeated-measures ANOVA revealed a main effect of condition—$F(2,51) = 331.1$, $p < .001$, $\eta^2 = 0.93$; Greenhouse-Geiser corrected for violation in sphericity. Consistent with the analysis of single-task performance, post hoc (Bonferroni-adjusted) paired comparisons revealed that children's speech recognition accuracy significantly declined with each decrease in spectral fidelity ($p < .001$). Critically, there was no main effect of task, $F(1,26) = 1.1$, $p = .31$, $\eta^2 = 0.04$, suggesting that children performed equally well on the primary task regardless of whether it was performed in isolation or concurrently with the secondary task. This finding suggests that children prioritized their attention on the speech recognition task in each condition regardless of whether they were simultaneously performing the visual monitoring task. This is consistent with results from Choi et al. (2008).
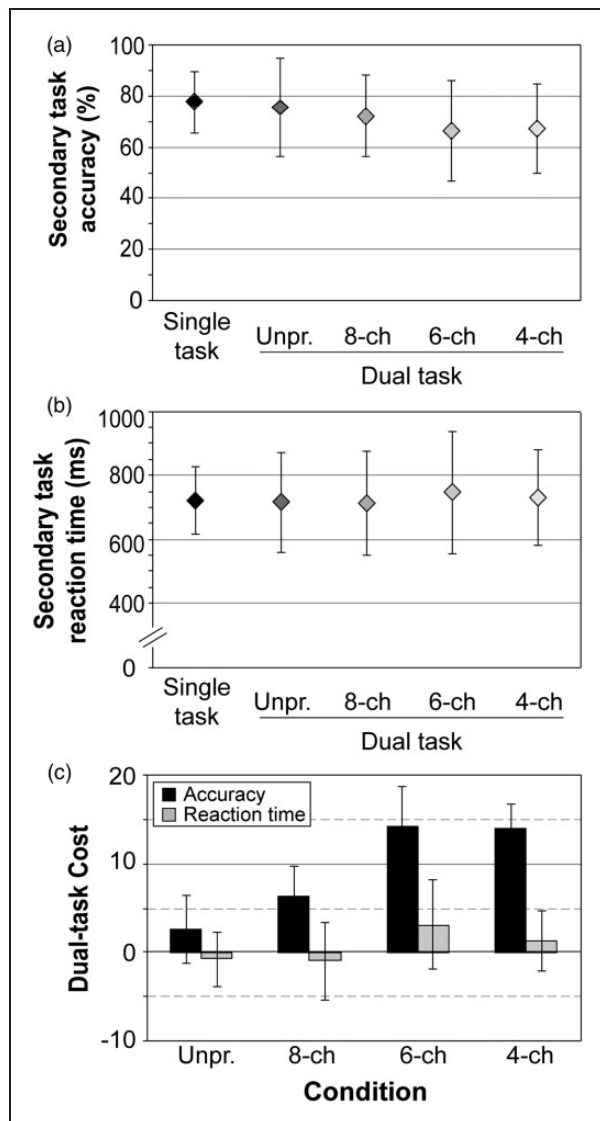
Dual-Task Performance: Visual Monitoring (Secondary) Task

Secondary-task accuracy and reaction time for the single- and dual-task conditions are illustrated in Figure 3(a) and (b), respectively. According to the assumptions of the dual-task paradigm, performance on the secondary task is expected to decline as the primary task becomes more demanding if the two tasks vie for the same cognitive resources (Kahneman, 1973). Consistent with this assumption, a repeated-measures ANOVA revealed a main effect of primary-task condition on secondary-task accuracy—$F(4, 104) = 5.1$, $p = .001$, $\eta^2 = 0.165$. Post hoc (Bonferroni-adjusted) paired comparisons showed that accuracy significantly differed from baseline performance in the 6-channel ($p < .05$) and 4-channel ($p < .01$) conditions. In contrast, there was no main effect of primary-task condition on secondary-task reaction time—$F(2.7, 71.3) = 0.512$, $p = .66$, $\eta^2 = 0.019$; Greenhouse-Geiser corrected for violation in sphericity. Children maintained their reaction time for indicating duplicate images that they identified in all dual-task conditions, even when their accuracy on this task declined.

In addition to absolute performance on the visual monitoring task, dual-task costs were calculated as the percentage of change in accuracy and reaction time on the visual monitoring task when performed in isolation versus when performed during the dual-task conditions. The mathematical equation used for this calculation is:

$$\text{Dual-task cost} = ([\text{Single Task} - \text{Dual Task}]/ \text{Single Task}) \times 100 \qquad (1)$$

Variations of this formula, originally developed by Somberg and Salthouse (1982), have been used previously to quantify *listening effort*, which relates to the cognitive costs of speech recognition (e.g., Kemper et al. 2009; Gosselin & Gagné 2011; Desjardins & Doherty 2013; Ward et al., 2017). For the current study, *dual-task cost* is the most fitting terminology because the purpose of the study was to quantify changes in secondary-task performance as a function of spectral fidelity of the target speech. Dual-task costs were calculated separately from children's accuracy and reaction time on the visual monitoring task (Figure 3(c)). To facilitate the comparison of dual-task costs as measured by accuracy and reaction time along the same scale (i.e., higher values indicating greater costs), the scores derived from changes in reaction time were each subtracted from zero, effectively inverting the sign. Dual-task costs derived from accuracy and reaction time were subjected

Figure 3. (a) Average (±SD) visual monitoring (secondary) task accuracy in isolation (*black*) or during the dual task conditions (*shades of gray*). (b) Average (±SD) visual monitoring (secondary) task reaction time in isolation (*black*) or during the dual task conditions (*shades of gray*). (c) Dual-task costs quantified by accuracy (*black*) or reaction time (*gray*). Positive values (>0) indicate greater dual-task costs.

to independent one-way ANOVAs with a within-subject factor of condition. Results identified a main effect of condition for dual-task costs in accuracy, $F(3, 78) = 4.4$, $p < .01$, $\eta^2 = 0.145$. Post hoc pairwise comparisons showed greater dual-task costs in the 4-channel and 6-channel conditions relative to the unprocessed condition ($p < .05$), though only the 4-channel condition is statistically different at an alpha of .05 following Bonferroni adjustment (unprocessed vs. 4-channel: $p < .05$; unprocessed vs. 6-channel: $p = .059$). All other paired comparisons were not statistically different

($p > .05$). In addition, there were no effects of condition for reaction time—$F(2.3, 60.6) = 0.58$, $p = .59$, $\eta^2 = 0.02$; Greenhouse-Geiser corrected for violation in sphericity.

In summary, the analysis of children's dual-task costs was consistent with the analysis of absolute secondary-task performance. Specifically, children failed to maintain their accuracy on the visual monitoring task as the speech of the primary task became more degraded, even while maintaining consistent reaction time for correct responses. This finding is consistent with the idea that children experience cognitive costs during degraded speech recognition that reduce their ability to multitask. The absence of dual-task costs on children's reaction time to identifying duplicate images is not, however, consistent with prior findings in young adults on a similar task (e.g., Pals et al., 2013; Ward et al., 2017). This issue will be further addressed in the Discussion section.

Individual Variability in Visual Monitoring During Speech Recognition

The reason for including the unprocessed condition in the dual task was to ensure that children could multitask when the speech signal was optimized. Therefore, secondary-task performance was not expected to differ from baseline performance in the unprocessed condition, and statistically, it did not. There was, however, large individual variability among children when asked to perform the dual-task condition: Some children experienced declines in their accuracy on the visual monitoring task during the dual-task paradigm. This observation suggests that some children experienced dual-task costs despite the high fidelity of the speech stimulus. This behavior may be related to individual factors, such as the age or cognitive maturity of the participants (e.g., Choi et al., 2008; Irwin-Chase & Burns, 2000), or to factors related to task performance, including their accuracy on the visual monitoring task in isolation or their accuracy on the speech recognition task.

To probe the source of variability in dual-task costs in the unprocessed condition, a linear regression model was developed to test whether accuracy on the visual monitoring task in the unprocessed dual-task condition was related to individual factors such as age, Flanker (age-unadjusted) score, baseline accuracy on the visual monitoring task, and speech recognition accuracy on the primary task. A priori correlations were conducted first to determine potential collinearity among factors. Only age and baseline (single task) accuracy on the visual monitoring task were correlated, though not strongly (see Figure 2(b)). All other correlations were not statistically significant ($p > .05$), and therefore, all of the factors were included in the model. The overall model fit was significant, $F(4, 22) = 6.38$, $p = .001$, accounting for 53.7% of the variance. Variance inflation factors (VIF)

were <2 and tolerance levels were >0.6 for all factors, confirming that collinearity was not present. Baseline accuracy on the visual monitoring task was found to be the only significant predictor at $p < .05$, $\beta = 0.637$, $t(22) = 3.9$, $p = .001$, 95% CI [0.48, 1.56], suggesting that children with higher accuracy on the visual monitoring task at baseline were better able to maintain their performance during the dual-task condition.

This finding raises the possibility that the observed individual variability in secondary-task accuracy during the dual task across the degraded conditions is also related to children's baseline performance on the visual monitoring task. This finding would be consistent with the idea that the relative complexity of the secondary task influences the extent to which performance on this task declines under degraded conditions of the primary task (Picou & Ricketts, 2014). To test this, independent regression models were generated for each degraded condition (i.e., 8-channel, 6-channel, and 4-channel). The dependent variable was visual monitoring task accuracy during each dual-task condition, and the predictor variables included dual-task speech recognition accuracy, baseline (single task) accuracy on the visual monitoring task, age, and Flanker (age-unadjusted) score. A priori correlations were conducted first to determine potential collinearity among the predictor variables. Age correlated with baseline accuracy on the visual monitoring task (Pearson $r = .41$, $p < .05$) and accuracy on the primary task in all degraded conditions (Pearson r for 8-channel = .45, 6-channel = .44, 4-channel = .57, $p < .05$). Baseline accuracy on the visual monitoring task correlated with accuracy on the speech recognition accuracy in the 8-channel and 4-channel conditions (Pearson r for 8-channel = .56, 4-channel = .46, $p < .05$), but not the 6-channel condition (Pearson $r = .32$, $p = .102$).

To isolate the relation between single-task and dual-task accuracy on the visual monitoring task, two models were tested for each condition. In Model 1, baseline accuracy on the visual monitoring task was chosen as the sole predictor. In Model 2, additional factors (i.e., age, dual-task accuracy on the speech recognition task, Flanker score) were added to determine if these factors accounted for any additional variance beyond baseline accuracy on the visual monitoring task in each condition. The overall fits of Model 1 for each condition were significant—8-channel: $F(1, 25) = 14.9$, $p = .001$; 6-channel: $F(1, 25) = 13.1$, $p = .001$; 4-channel: $F(1, 25) = 32.6$, $p < .001$—and accounted for 34.8%, 31.8%, and 54.9% (adjusted R^2) of the variance observed in secondary-task accuracy, respectively. Baseline accuracy on the visual monitoring task was a significant predictor in each model—8-channel: $\beta = 0.611$, $t(25) = 3.86$, $p = .001$, 95% CI [0.42, 1.37]; 6-channel: $\beta = 0.586$, $t(25) = 3.6$, $p = .001$, 95% CI [0.4, 1.48]; 4-channel: $\beta = 0.752$, $t(25) = 5.7$, $p < .001$, 95% CI [0.72, 1.52]. The overall fits of Model

2 for each condition did not result in a significant F-change ($p > .05$), suggesting that all other predictor variables (i.e., age, dual-task speech recognition accuracy, Flanker score) did not significantly account for secondary-task accuracy beyond that of baseline accuracy on the visual monitoring task in each condition.

In summary, the results of the regression models suggest that children's performance on the visual monitoring task during degraded speech recognition is largely influenced by their ability to perform the visual monitoring task in isolation, regardless of the difficulty of the primary-task condition. The observation that secondary-task accuracy declined with poorer spectral fidelity in the primary task suggests, however, that something about the primary task influenced secondary-task performance in school-age children. This relation could not be identified with the battery of tests that were implemented in the current study.

Discussion

The purpose of the current study was to assess school-age children's multitasking abilities during degraded speech recognition. Results showed that children's accuracy on the visual monitoring (secondary) task declined with decreasing spectral fidelity in the primary speech signal. These data are consistent with the idea that children reallocated cognitive resources to the primary task and therefore had fewer cognitive resources for the visual monitoring task in the more difficult conditions of the speech recognition task.

In contrast to the observed declines in secondary-task accuracy, children maintained their reaction time for correct responses across conditions. These data contrast those from young adults, whose reaction time on a visual secondary task slowed with increased demands of the primary task (e.g., Pals et al., 2013; Ward et al., 2017). The children's response pattern, however, is consistent with data from older adults who performed the same task under the same experimental conditions: The older adults in Ward et al. (2017) had statistically significant declines in secondary-task accuracy but failed to reveal statistically significant changes in secondary-task reaction time across degraded primary-task conditions. Taken together, the results across these studies suggest that multitasking during degraded speech recognition may be differentially managed across the lifespan (Forstmann et al., 2011; Garrett, 1922; Ratcliff, Love, Thompson, & Opfer, 2012; Salthouse & Somberg, 1982).

Multitasking in Children

Previous multitasking studies in school-age children have often failed to show changes in secondary-task performance across increased demands of the primary tasks

within dual-task paradigms (e.g., Hicks & Tharpe, 2002; Stelmachowicz et al., 2007). Children's inability to reallocate cognitive resources persisted even when children were instructed to weight attention on one task versus the other (Choi et al., 2008; Irwin-Chase & Burns, 2000). The results from the current study extend our understanding about attention allocation in school-age children in three ways. First, the current study provides support for the idea that children can reallocate their cognitive resources during degraded speech recognition when primary-task conditions are sufficiently difficult to elicit a decline in performance. For example, both the current study and the Howard et al. (2010) study showed changes in secondary-task performance that were coincident with declines of 30% to 50% in children's primary-task accuracy across degraded conditions. In contrast, children in earlier studies had, on average, minimal performance declines (<20%; Hicks & Tharpe, 2002; Stelmachowicz et al., 2007) across the primary-task conditions. Thus, greater performance declines in speech recognition may be required to trigger resource reallocation in children. It is important to note, however, that though resources were successfully taken away from the secondary task in the more degraded conditions, there is no definitive evidence that these resources were reallocated to aid in primary-task performance.

Second, when considering the work by Irwin-Chase and Burns (2000) and Choi et al. (2008), the current results suggest that younger children may be better able to allocate attention across tasks that have fewer overlapping processing demands. Irwin-Chase and Burns (2000) and Choi et al. (2008) showed immature attention allocation on dual-task paradigms consisting of primary and secondary tasks that were either both verbal (Choi et al., 2008) or both visual (Irwin-Chase & Burns, 2000). Immature allocation was not evident in the unprocessed, 8-channel, or 6-channel conditions in the current study because children were able to maintain their primary-task accuracy in the dual-task conditions, which is what they were instructed to do. The only evidence of insufficient cognitive resource reallocation was in the 4-channel condition. Children had statistically similar accuracy on the secondary task in the 6-channel and 4-channel conditions despite declining accuracy on the primary task in those conditions. This observation suggests that children's ability to reallocate their attention in moderately challenging listening situations is compromised under more degraded conditions. Children were either unwilling to sacrifice their performance on the secondary task or unable to appropriately reallocate their attention in the 4-channel condition.

Finally, the results of the current study highlight that proficiency with the secondary task, but not chronological age or performance on a task of attention and inhibition, largely accounted for the change in secondary-task performance as a function of primary-task condition. Specifically, children who had higher accuracy on the visual monitoring task at baseline also had higher accuracy on this task in the degraded dual-task conditions. This finding has functional implications for children who multitask during degraded speech recognition in real-word environments. Namely, greater proficiency on the secondary task in isolation may help to preserve performance when the task is performed concurrently with speech recognition.

One plausible source of variability in results across the multitasking literature is the range of primary and secondary tasks implemented in each study (Picou & Ricketts, 2014). Myriad primary tasks (e.g., word recognition, word categorization, sentence recognition), secondary tasks (e.g., visual probe, auditory digit span, visual-motor task, visual monitoring), and sources of speech degradation (e.g., broadband noise, speech babble, spectral filtering) have likely contributed to the inconsistent observations of cognitive resource reallocation. As a result, it is difficult to generalize results across the literature, particularly because different primary and secondary tasks, and combinations of them, likely demand different cognitive processes. Given that performance on both primary and secondary tasks at baseline appears to relate to the dual-task costs under degraded conditions, proficiency in each task should always be considered when interpreting data from multitasking paradigms.

Secondary-Task Performance

Although the results suggest that children reallocate cognitive resources away from the secondary task during degraded speech recognition, it is notable that this was observed for secondary-task accuracy but not for reaction time. This is consistent with prior studies involving children, in which secondary-task reaction time does not change as a function of speech recognition task demands (e.g., Hicks & Tharpe, 2002; McFadden & Pittman, 2008). This finding may reflect a general response strategy of children, who often prioritize reaction time over accuracy in multitasking situations across a variety of domains (e.g., Karatekin, Couperus, & Marcus, 2004; Rival, Olivier, & Ceyte, 2003) and who may set different response criteria than adults (Ratcliff et al., 2012).

However, the current finding conflicts with the performance of young adults performing similar dual-task paradigms, whose reaction time on the secondary task slows concomitantly with greater primary-task demands (e.g., Downs & Crum, 1978; Pals et al., 2013; Picou, Ricketts, & Hornsby, 2011; Ward et al., 2017). One problem with attempting to reconcile the findings of

the current study with those involving young adults is that secondary-task performance is typically reported as reaction time or accuracy, but not both. Thus, it is difficult to assess differences across these studies because individuals may adopt different response strategies—that is, they differentially weight reaction time or accuracy—depending on their baseline cognitive capacity (Ward et al., 2017), changing task demands of the primary and secondary tasks (e.g., Ratcliff et al., 2012; Van Veen, Krug, & Carter, 2008; Verhaeghen, Steitz, Sliwinski, & Cerella, 2003), and the instructions that were provided (Hale, 1969). Related to this latter idea, children in the current study were told to respond ''. . . as quickly as possible . . .'' when they saw a repeating image when performing the visual monitoring (secondary) task. This instruction may have biased their focus on reaction time to the detriment of accuracy as the speech recognition (primary) task became more degraded. Different instructions, such as ''. . . respond whenever you see a repeating image, even if you are uncertain. . .'' might have resulted in different response patterns.

Taken together, there are two possible interpretations of children's secondary-task performance in the current study. One possibility is that children differ in their multitasking strategies in degraded listening conditions relative to adults: Children attempt to maintain their speed of response at the risk of missing a duplicate picture whereas adults' accuracy and reaction time change concomitantly. Alternatively, the results of the current study may solely reflect the instructions that children were given prior to performing the secondary task. If this latter point is true, then their response strategies may change based on the nature of the secondary task and how they are instructed (Bonnel & Hafter, 1998; Picou & Ricketts, 2014). Additional studies will need to be implemented to better understand these relations.

Relation Between Primary- and Secondary-Task Performances

One assumption of dual-task paradigms is that the magnitude of the decline in secondary-task performance is expected to be somewhat proportional to the increasing cognitive demands of the primary task (Kahneman, 1973). Consistent with this idea, we observed that secondary-task accuracy declined as spectral fidelity decreased. However, post hoc comparisons showed that the declines in secondary-task accuracy did not strictly mirror the declines in speech recognition accuracy across conditions. In other words, though children's speech recognition accuracy significantly decreased with each sequential decrease in spectral fidelity of the speech input, statistically significant changes in secondary-task accuracy relative to baseline were only observed in the 6-channel and 4-channel conditions (see Figure 3(a) and

(c)). There was no change in secondary-task accuracy between the (a) unprocessed and 8-channel conditions; (b) 8-channel and 6-channel conditions, or (c) 6-channel and 4-channel conditions. One possible explanation for this finding is that children actually prioritized the secondary task over the primary task in the noise-band vocoded conditions. If this were true, however, then one might expect a ''dual-task cost'' in primary-task accuracy, whereby speech recognition accuracy would decline from baseline performance during the dual task. The current data do not support this interpretation: Speech recognition accuracy was equivalent between the single and dual tasks in all conditions. Instead, the data are more consistent with the idea that it was easier for children to maintain performance across two sensory modalities (e.g., auditory and visual) versus within two channels of a single modality (e.g., auditory only; Proctor & Proctor, 1979; Treisman & Davies, 1973).

An alternate explanation for these results is that the change in primary-task accuracy needs to reach a minimal threshold before coincident changes in secondary-task performance can be observed (as discussed above and in Howard et al., 2010). In the current study, children recognized a high percentage ($\sim$88%) of speech that was noise-band vocoded with eight spectral channels and the magnitude of difference between the unprocessed and 8-channel conditions was small ($\sim$12%). Alternatively, the visual monitoring task implemented in the current study may have required few cognitive resources, especially for children who were more proficient on the task (Hasher & Zacks, 1979). If this is true, then a more demanding secondary task may be more sensitive to small changes in speech recognition accuracy. In either case, the data are consistent with previous work wherein small changes in primary-task accuracy did not change secondary-task performance (Hicks & Tharpe, 2002; McFadden & Pittman, 2008; Stelmachowicz et al., 2007).

An alternate possibility is that children may have simply been unwilling to commit additional cognitive resources to extract meaning from the degraded speech input. Our data also suggest that there may be a point at which children will not commit any additional resources to support of primary-task performance. For example, there was a $\sim$22% decline in speech recognition accuracy in the 4-channel condition relative to the 6-channel condition. Despite this large decline on the primary task, performance on the secondary task was equivalent across these conditions, suggesting that children were allocating the same amount of cognitive resources to the secondary task regardless of difficulty. This is somewhat unexpected in the context of theoretical frameworks such as the Ease of Language Understanding model, which would predict greater declines in secondary-task performance with increasing mismatches between language input and memory stores (Rönnberg,

2003; Rönnberg et al. 2013). The underlying assumption to this model, however, is that the listener attempts to resolve the mismatch versus offering their "best guess" when repeating what they heard. In contrast, our data suggest that equivalent cognitive resources were allocated to the secondary task in the 6-channel and 4-channel conditions, suggesting an unwillingness or inability to reallocate any additional resources away from the visual monitoring task in the 4-channel condition. This finding would then be more consistent with the work from Irwin-Chase and Burns (2000) and Choi et al. (2008) because it represents a condition in which children may not have reallocated their cognitive resources as instructed.

Limitations of the Study

One unexpected finding was the failure to identify a relation between children's executive function, as quantified by the Flanker task, and their ability to multitask during degraded speech recognition. Due to the length of the protocol (~2 hours inclusive of breaks) and because the dual-task paradigm was bimodal, the only executive function measure used was the Flanker task. The Flanker was specifically chosen for this study because of its ability to capture participants' attention and inhibition abilities, predicted to be crucial in a bimodal paradigm. However, children's Flanker scores were not shown to predict dual-task performance in any of the conditions, suggesting that executive function did not serve as the critical cognitive bottleneck for children's ability to multitask in the current study. Primary-task accuracy was also not predictive of secondary-task performance in the degraded conditions, and a great deal of variance was left unexplained, suggesting the importance of as-yet undiscovered additional factors.

Another potential confound relates to the bimodality of the dual-task paradigm itself. Despite being a visual task, the visual monitoring task may have engaged the phonological loop of working memory (Baddeley, 2000) if children were silently naming the pictures as they viewed them on the screen. The current experimental design is unable to rule out this possibility. Given that it is unclear which cognitive abilities are involved in this particular multitasking paradigm, the results again suggest that other unmeasured individual factors are plausibly important for multitasking.

A further limitation of the current study is that the results may not be generalizable across primary tasks consisting of speech degraded by other methods (e.g., the addition of speech babble or non-speech noise) or secondary tasks with greater overlap with the primary task (e.g., an auditory or verbal secondary task). The large interstudy variability in methodologies across the dual-task literature suggests that additional studies are necessary to pin down the exact cognitive mechanisms

that are involved, the strategies individuals use to maintain speech recognition accuracy under degraded conditions, and what secondary tasks are the most ideal to quantify these strategies.

Finally, there is a question as to whether these results generalize to children's performance in more realistic environments. Noise-band vocoding is not a typical experience for listeners, particularly children. However, the consistency between our results and those of Howard et al. (2010), who used speech babble to degrade the primary task, suggests that noise-band vocoding is a feasible method to use when exploring the cognitive mechanisms involved in children's ability to multitask during degraded speech recognition.

Conclusion

In conclusion, degraded speech recognition in school-age children comes with a domain-general cognitive cost. As a result, children's ability to multitask, even when the two tasks do not compete for domain-specific resources, can be compromised. Secondary-task performance, however, appears to depend on primary-task difficulty, the children's baseline proficiency on the secondary task, and potentially the metric used to quantify performance (i.e., accuracy vs. reaction time). Additional studies are necessary to probe what cognitive resources are involved in multitasking during degraded speech recognition in children. In addition, the primary and secondary tasks chosen for future studies should be those for which baseline proficiency can be obtained. If working with younger children, the use of multimodal tasks may be better in order to minimize developmental effects.

References

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *The Psychology of Learning and Motivation*, *8*, 47–89.

Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology*, *13*, 108–112.

Bonnel, A. M., & Haftser, E. R. (1998). Divided attention between simultaneous auditory and visual signals. *Perception & Psychophysics*, *60*(2), 179–190.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Broadbent, D. E. (1958). *Perception and communication*. Elmsford, NY: Pergamon Press. doi:10.1037/10037-000.

Choi, S., Lotto, A., Lewis, D., Hoover, B., & Stelmachowicz, P. (2008). Attentional modulation of word recognition by children in a dual-task paradigm. *Journal of Speech, Language, and Hearing Research*, *51*(4), 1042–1054.

Crandell, C. C., & Smaldino, J. J. (2000). Classroom acoustics for children with normal hearing and with hearing impairment. *Language, Speech, and Hearing Services in Schools*, *31*(4), 362–370.

Desjardins, J. L., & Doherty, K. A. (2013). Age-related changes in listening effort for various types of masker noises. *Ear and hearing*, *34*(3), 261–272.

Downs, D. W., & Crum, M. A. (1978). Processing demands during auditory learning under degraded listening conditions. *Journal of Speech, Language, and Hearing Research*, *21*(4), 702–714.

Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., & Boothroyd, A. (2000). Speech recognition with reduced spectral cues as a function of age. *The Journal of the Acoustical Society of America*, *107*(5), 2704–2710.

Elliott, L. L., Connors, S., Kille, E., Levin, S., Ball, K., Katz, D. (1979). Children's understanding of monosyllabic nouns in quiet and in noise. *The Journal of the Acoustical Society of America*, *66*(1), 12–21.

Forstmann, B. U., Tittgemeyer, M., Wagenmakers, E. J., Derrfuss, J., Imperati, D., Brown, S. (2011). The speed-accuracy tradeoff in the elderly brain: A structural model-based approach. *The Journal of Neuroscience*, *31*(47), 17242–17249.

Garrett, H. E. (1922). A study of the relation of accuracy to speed. *Archives of Psychology*, *56*, l–104.

Gosselin, P. A., & Gagne, J. P. (2011). Older adults expend more listening effort than young adults recognizing speech in noise. *Journal of Speech, Language, and Hearing Research*, *54*(3), 944–958.

Hale, D. J. (1969). Speed-error tradeoff in a three-choice serial reaction task. *Journal of Experimental Psychology*, *81*(3), 428.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*(3), 356.

Hicks, C. B., & Tharpe, A. M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language, and Hearing Research*, *45*(3), 573–584.

Howard, C. S., Munro, K. J., & Plack, C. J. (2010). Listening effort at signal-to-noise ratios that are typical of the school classroom. *International Journal of Audiology*, *49*(12), 928–932.

Irwin-Chase, H., & Burns, B. (2000). Developmental changes in children's abilities to share and allocate attention in a dual task. *Journal of Experimental Child Psychology*, *77*(1), 61–85.

Johnson, C. E. (2000). Children's phoneme identification in reverberation and noise. *Journal of Speech, Language, and Hearing Research*, *43*(1), 144–157.

Jones, P. R., Moore, D. R., & Amitay, S. (2015). Development of auditory selective attention: Why children struggle to hear in noisy environments. *Developmental Psychology*, *51*(3), 353.

Kahneman, D. (1973). *Attention and effort* (p. 246). Englewood Cliffs, NJ: Prentice-Hall.

Karatekin, C., Couperus, J. W., & Marcus, D. J. (2004). Attention allocation in the dual-task paradigm as measured through behavioral and psychophysiological responses. *Psychophysiology*, *41*(2), 175–185.

Kemper, S., Schmalzried, R., Herman, R., Leedahl, S., & Mohankumar, D. (2009) The effects of aging and dual task demands on language production. *Aging Neuropsychol Cogn*, *16*(3): 241–259.

Klatte, M., Hellbrück, J., Seidel, J., & Leistner, P. (2010). Effects of classroom acoustics on performance and well-being in elementary school children: A field study. *Environment and Behavior*, *42*(5), 659–692.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.

Leibold, L. J., & Buss, E. (2013). Children's identification of consonants in a speech-shaped noise or a two-talker masker. *Journal of Speech, Language, and Hearing Research*, *56*(4), 1144–1155.

McDonald, S. (2014). Special series on the cognition battery of the NIH toolbox. *Journal of International Neuropsychological Society*, *20*(6), 487–651. doi:http://dx.doi.org/10.1017/S1355617714000241.

McFadden, B., & Pittman, A. (2008). Effect of minimal hearing loss on children's ability to multitask in quiet and in noise. *Language, Speech, and Hearing Services in Schools*, *39*(3), 342–351.

Pals, C., Sarampalis, A., & Başkent, D. (2013). Listening effort with cochlear implant simulations. *Journal of Speech, Language, and Hearing Research*, *56*(4), 1075–1084.

Papso, C. F., & Blood, I. M. (1989). Word recognition skills of children and adults in background noise. *Ear and Hearing*, *10*(4), 235–236.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America*, *97*(1), 593–608.

Picou, E. M., & Ricketts, T. A. (2014). The effect of changing the secondary task in dual-task paradigms for measuring listening effort. *Ear and Hearing*, *35*(6), 611–622.

Picou, E. M., Ricketts, T. A., & Hornsby, B. W. (2011). Visual cues and listening effort: Individual variability. *Journal of Speech, Language, and Hearing Research*, *54*(5), 1416–1430.

Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, *81*(1), 10.

Proctor, R. W., & Proctor, J. D. (1979). Secondary task modality, expectancy, and the measurement of attentional capacity. *Journal of Experimental Psychology: Human Perception and Performance*, 5(4), 610.

Rabbitt, P. M. (1966). Recognition: Memory for words correctly heard in noise. *Psychonomic Science*, 6, 383–384.

Rakerd, B., Seitz, P., & Whearty, M. (1996). Assessing the cognitive demands of speech listening for people with hearing losses. *Ear and Hearing*, 17(2), 97–106.

Ratcliff, R., Love, J., Thompson, C. A., & Opfer, J. E. (2012). Children are not like older adults: A diffusion model analysis of developmental changes in speeded responses. *Child Development*, 83(1), 367–381.

Rival, C., Olivier, I., & Ceyte, H. (2003). Effects of temporal and/or spatial instructions on the speed–accuracy trade-off of pointing movements in children. *Neuroscience Letters*, 336(1), 65–69.

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *International Journal of Audiology*, 42(Suppl 1): S68–S76.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B.,...Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theory, data, and clinical implications. *Frontiers in Systems Neuroscience*, 7, 1–17.

Salthouse, T. A., & Somberg, B. L. (1982). Time-accuracy relationships in young and old adults. *Journal of Gerontology*, 37(3), 349–353.

Sarampalis, A., Kalluri, S., Edwards, B., & Hafter, E. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. *Journal of Speech, Language, and Hearing Research*, 52(5), 1230–1240.

Shield, B. M., & Dockrell, J. E. (2003). The effects of noise on children at school: a review. *Building Acoustics*, 10(2), 97–116.

Stelmachowicz, P. G., Lewis, D. E., Choi, S., & Hoover, B. (2007). Effect of stimulus bandwidth on auditory skills in normal-hearing and hearing-impaired children. *Ear and Hearing*, 28(4), 483–944. http://doi.org/10.1097/AUD.0b013e31806dc265.

Tukey, J. W. (1990). The practical relationship between the common transformations of percentages or fractions and of amounts. In C. L. Mallows (Ed.), *The Collected Works of John W. Tukey. Volume VI: More Mathematical* (pp. 211–219). Pacific Grove, CA: Wadsworth & Brooks-Cole.

Treisman, A. M., & Davies, A. (1973). Divided attention to ear and eye. *Attention and Performance*, IV, 101–117.

Van Veen, V., Krug, M. K., & Carter, C. S. (2008). The neural and computational basis of controlled speed-accuracy trade-off during task performance. *Journal of Cognitive Neuroscience*, 20(11), 1952–1965.

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: A meta-analysis. *Psychology and Aging*, 18(3), 443.

Ward, K. M., Shen, J., Souza, P. E., & Grieco-Calub, T. M. (2017). Age-related differences in listening effort during degraded speech recognition. *Ear and Hearing*, 38(1), 74–84.

Wightman, F. L., & Kistler, D. J. (2005). Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *The Journal of the Acoustical Society of America*, 118(5), 3164–3176.

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: The processing of degraded speech depends critically on attention. *The Journal of Neuroscience*, 32(40), 14010–14021.

Wright, B. A., Conderman, J. S., Waggenspack, M. K., & Marrone, N. (2014). Prevention of learning of a nonnative phonetic contrast by prior exposure to the contrasting stimuli while performing an irrelevant visual task. *The Journal of the Acoustical Society of America*, 135(4), 2357–2357.