Background Speech Disrupts Working Memory Span in 5-Year-Old Children

Tina M. Grieco-Calub,1,2 Maya-Simone Collins,1 Hillary E. Snyder,1 and Kristina M. Ward1

Objectives: The present study tested the effects of background speech and nonspeech noise on 5-year-old children's working memory span.

Design: Five-year-old typically developing children (range = 58.6 to 67.6 months; n = 94) completed a modified version of the Missing Scan Task, a missing-item working memory task, in quiet and in the presence of two types of background noise: male two-talker speech and speech-shaped noise. The two types of background noise had similar spectral composition and overall intensity characteristics but differed in whether they contained verbal content. In Experiments 1 and 2, children's memory span (i.e., the largest set size of items children successfully recalled) was subjected to analyses of variance designed to look for an effect of listening condition (within-subjects factor: quiet, background noise) and an effect of background noise type (between-subjects factor: two-talker speech, speech-shaped noise).

Results: In Experiment 1, children's memory span declined in the presence of two-talker speech but not in the presence of speech-shaped noise. This result was replicated in Experiment 2 after accounting for a potential effect of proactive interference due to repeated administration of the Missing Scan Task.

Conclusions: Background speech, but not speech-shaped noise, disrupted working memory span in 5-year-old children. These results support the idea that background speech engages domain-general cognitive processes used during the recall of known objects in a way that speech-shaped noise does not.

Key words: Background noise, Children, Missing-item task, Two-talker speech masker; Working memory span.

(Comm 2018;XX:00–00)

INTRODUCTION

Children's environments often contain background noise, such as environmental sounds, machinery, and other people talking, that disrupt their performance on focal tasks. For example, background noise impairs infants' and children's performance on an array of auditory tasks. Background noise disrupts children's speech perception (Mills 1975; Elliott et al. 1979; Fallon et al. 2000; Johnson 2000; Hall et al. 2002; Leibold et al. 2016), reduces infants' recognition of their own name (Newman 2005), slows toddlers' real-time language processing (Grieco-Calub et al. 2009), interrupts children's mapping of novel label-object pairs (Riley & McGregor 2012; McMillan & Saffran 2016), and disrupts auditory working memory (Olsen & Sullivan 2014). In the course of their day, however, children engage in a number of activities that are nonauditory, such as visually attending to objects in their environment, visually processing and mentally manipulating verbal content (e.g., reading, recalling visually-presented information), or executing motor sequences (e.g., writing). Although there is evidence that background noise disrupts performance on nonauditory tasks such as these in older children and adults, the effect of background noise on younger children's performance on similar tasks has not been examined thoroughly. The purpose of the present study is to test the effect of background noise on a nonauditory task in 5-year-old children—a group of children who are younger than cohorts previously studied.

The effects of background noise on nonauditory tasks in older children and adults have largely been studied within the context of the irrelevant sound effect (ISE). The ISE is an example of cross-modal interference whereby, despite the lack of a perceptual overlap, background noise—an auditory stimulus— disrupts performance on focal tasks relying on other sensory modalities (e.g., visual; Salamé & Baddeley 1982; Martin et al. 1988; Jones et al. 1992; LeCompte 1995; Elliott 2002; Beaman 2005; Perham & Macpherson 2012). The task most frequently used in previous studies of the ISE is serial recall of visual-verbal input. During visual-verbal serial recall tests, individuals are provided with a set of items they encode visually (e.g., 5, 7, 3, 4), temporally store, and then subsequently recall verbally either in the same order (e.g., “5, 7, 3, 4”; forward span) or in reverse order (e.g., “4, 3, 7, 5”; backward span). The ability of background noise to exert cross-modal interference suggests that background noise directly engages domain generalized cognitive processes, such as working memory and attention, which are also necessary for focal task performance. Given that cognitive processing matures throughout development, there is reason to suspect that younger children's task performance may be particularly susceptible to disruption by background noise.

There are three potential mechanisms through which background noise may disrupt performance on serial processing of visual-verbal content despite the absence of a perceptual overlap of the stimuli to be recalled (i.e., visual) and the background noise (i.e., auditory). According to the interference-by-content account, background noise interferes with recall when the noise shares perceptual, phonological, or semantic space with the information to be manipulated (e.g., Neely & LeCompte 1999). In other words, interference stems from an overlap in the perceptual encoding of the visual stimulus and background noise. Alternatively, the interference-by-process account suggests that background noise exerts interference because of overlapping demands on the cognitive processes associated with focal task performance and obligatory processing of the sound (e.g., Hughes et al. 2007; Marsh et al. 2009). For example, background speech does not necessarily interfere with encoding of the items in the visual domain to be recalled, but rather interferes with the cognitive processes needed for serial rehearsal, resulting in poorer recall (e.g., Jones et al. 1992; Sörvist 2010). A third potential mechanism is attentional capture. Temporally dynamic acoustic stimuli (i.e., those that vary more widely in intensity over time or contain abrupt onsets/offsets or deviants)
can draw individuals’ attention during nonauditory tasks (Elliott 2002; Klatte et al. 2010; Elliott et al. 2016; Vachon et al. 2017). The latter two mechanisms have been described as a duplex-mechanism account of auditory distraction (e.g., Hughes et al. 2005, 2007, 2013).

The existing theoretical framework on the mechanisms of the ISE provides an opportunity to explore the effects of background noise on nonauditory task performance in children. The observation that adults have smaller ISEs than children is attributed to the fact that the cognitive processes necessary for serial recall are more mature in adults than in children (Elliott 2002; Elliott & Briganti 2012; Elliott et al. 2016). Specifically, the integral cognitive skills that are expected to improve with development include recoding and recall of serial information, use of rehearsal strategies, working memory span, and attentional control (Cowan 1995; McCormack et al. 2000; Gathercole et al. 2004; Tam et al. 2010). Given this rationale, if we narrow our focus to only children, we would expect younger children to have larger ISEs (i.e., greater disruption in performance due to background noise) than older children. In the auditory literature referenced above, this has been repeatedly observed: background noise is more disruptive in younger children than in older children and adults (Wightman & Kistler 2005; Leibold & Buss 2013). In addition, both in the auditory literature and in the ISE literature, background noise containing speech, and therefore verbal content, tends to exert greater interference than nonspeech noise (Hall et al. 2002; Wightman & Kistler 2005; Leibold & Buss 2013). These observations support the idea that meaningful background noise (e.g., speech) engages domain-general cognitive processes used during speech perception in noise (e.g., working memory and selective attention) in a way that speech-shaped noise does not. However, the extent to which background noise with different content disrupts performance on nonauditory tasks has not been directly tested in younger children.

One reason for the gap in knowledge regarding the ISE in younger children is related to the high task demands of traditionally implemented serial recall tasks. For example, younger children often encounter output interference on serial recall tasks, whereby their ability to retain the most recent items in the set suffers while verbally recalling earlier items (e.g., Conlin & Gathercole 2006). In addition, children’s immature rehearsal abilities and poor attentional control may limit their ability to recall serially presented items (Elliott et al. 2016). In contrast, missing-item tasks may prove to be advantageous when evaluating the effects of background noise on cognitive processes, specifically working memory, in younger populations. During these tasks, individuals are shown a set of items, which are subsequently hidden from vision. All but one item are returned to vision, and individuals are asked to recall only the item that is still hidden. The limited recall demands of this task are advantageous when testing children who may be less able or willing to retrieve all items in their memory span. In a recent study by Elliott et al. (2016), the researchers used serial recall, probed-recall (e.g., “what number came after X?”), and a missing-item recall task to test the extent to which children’s attentional control and rehearsal strategies influenced their performance. They found that although adults and children experienced disruption due to interference-by-process (i.e., disruption of rehearsal strategies), the poorer performance observed in children was primarily attributed to their immature attentional control. More generally speaking, this work highlighted the feasibility of using a missing-item task in children and adults to evaluate the effects of background noise on working memory span (see also LeCompte 1996; Beaman & Jones 1997).

In the present study, we implemented an adapted version of the Missing Scan Task (MST; Buschke 1963), which was recently implemented by Roman et al. (2014) to assess visual-verbal working memory span in typically developing children between 3 and 6 years of age. Roman et al. (2014) found that children’s age accounted for 24% of the variance in their performance on the MST (i.e., the span score). This relationship is not surprising given the rapid cognitive development that occurs over this age range. In addition, children’s receptive language and their performance on standardized verbal working memory and visual working memory subtests of the NEPSY-II (Korkman et al. 2007) predicted an additional 19% and 14%, respectively, of the variance in children’s performance on the MST. This suggests that the MST is a valid test of memory ability in young children and may be a good test to quantify the effects of background noise.

The purpose of the present study was to explore the relation between background noise and working memory, as measured using a modified version of the MST, in 5-year-old children. Specifically, this study (1) examined whether background noise disrupts children’s working memory span; and (2) tested the extent to which the content of the background noise (i.e., speech or nonspeech noise) differentially alters task performance. In Experiment 1, we implemented conditions during which children completed the MST in the presence of either (1) background noise containing two-talker male speech or (2) background noise containing speech-shaped noise with the same spectral composition and overall intensity as the two-talker speech stimulus but lacking verbal content (Elliott et al. 1979; Papso & Blood 1989; Wightman & Kistler 2005; Leibold & Buss 2013; Corbin et al. 2016). We hypothesized that background noise interferes with working memory in young children and that background noise containing verbal content is more disruptive to performance than background noise lacking verbal content. Therefore, we predicted that children’s span scores would decline in the presence of both types of background noise but that a greater decline would be observed in the presence of two-talker speech than speech-shaped noise.

**EXPERIMENT 1**

**Materials and Methods**

**Participants** • Forty-nine children (27 females; 61.3 ± 1.8 months, mean ± SD) were recruited to participate. An additional seven children were recruited but did not complete the task and, therefore, are not included in this dataset. Recruitment occurred primarily through a child registry that is administered by the authors and secondarily through advertisement in the Evanston, IL community. All children were typically developing, monolingual native English speakers, and had normal hearing and vision, per parental report. To ensure all children had age-appropriate receptive vocabulary, the Receptive One-Word Picture Vocabulary Test, 4th edition (ROWPVT-4; Martin & Brownell 2010) was administered. All children
TABLE 1. Summary of the participants’ age, standardized receptive vocabulary score, and task performance per experiment.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Age (mo)</th>
<th>Receptive Vocabulary (Standard Score)</th>
<th>Animals Included in the Task (Out of 34)</th>
<th>Span Score (Quiet)</th>
<th>Span Score (Quiet)</th>
<th>Span Score (Noise)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Quiet only</td>
<td>61.8±2.2</td>
<td>117±11.5</td>
<td>32.4±1.7</td>
<td>4.5±1.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-talker speech (n = 16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order 1 (n = 8)</td>
<td>61.1±0.9</td>
<td>111.6±9.2</td>
<td>33.6±0.7</td>
<td>5.3±1.9</td>
<td>3.8±2</td>
<td></td>
</tr>
<tr>
<td>Order 2 (n = 8)</td>
<td>62.0±2.2</td>
<td>122.4±7.4</td>
<td>32.4±2.1</td>
<td>6.1±1.4</td>
<td>4.8±1.2</td>
<td></td>
</tr>
<tr>
<td>Speech-shaped noise (n = 16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order 1 (n = 8)</td>
<td>60.7±1.0</td>
<td>117.6±7.4</td>
<td>32.9±1.1</td>
<td>4.3±1.3</td>
<td>3.0±0.8</td>
<td></td>
</tr>
<tr>
<td>Order 2 (n = 8)</td>
<td>60.4±1.7</td>
<td>115.2±4.9</td>
<td>33.1±1.1</td>
<td>3.4±1.1</td>
<td>4.5±1.1</td>
<td></td>
</tr>
<tr>
<td>2 Group 1 (n = 15)</td>
<td>62.4±2.0</td>
<td>116.9±8.0</td>
<td>33.0±2.1</td>
<td>4.5±1.2</td>
<td>3.8±1.4</td>
<td>4.1±1.3</td>
</tr>
<tr>
<td>Group 2 (n = 15)</td>
<td>61.1±2.5</td>
<td>119.1±12.2</td>
<td>31.3±3.1</td>
<td>4.5±1.4</td>
<td>4.0±1.5</td>
<td>3.3±0.8</td>
</tr>
<tr>
<td>Group 3 (n = 15)</td>
<td>62.0±2.4</td>
<td>121.2±6.2</td>
<td>31.9±1.3</td>
<td>5.1±1.6</td>
<td>4.5±1.1</td>
<td>5.1±1.4</td>
</tr>
</tbody>
</table>

Values are given as mean ± SD.

obtained a standardized score within or above one standard deviation of the mean, indicating receptive vocabulary within the normal range for their age. The mean receptive vocabulary scores for children, grouped by condition, are listed in Table 1. Before study enrollment, a parent or legal guardian completed an informed consent process. Testing occurred within a single session that lasted no longer than 1.5 hours. Children were monetarily compensated for their time, as well as incrementally provided with additional incentives (i.e., stickers) throughout participation to maintain their engagement. All procedures were approved by the Institutional Review Board at Northwestern University.

Stimuli • Thirty-four plastic toy animals were used as stimuli. The animals in the set were heterogeneous and included domesticated animals (e.g., rabbit), farm animals (e.g., pig), jungle animals (e.g., tiger), ocean animals (e.g., dolphin), zoo animals (e.g., penguin), and forest animals (e.g., black bear). Only animals that children were able to verbally label spontaneously were included in the set of animals used during testing, as described in the Procedure section.

Background Noise • Two types of background noise were used in the present study: two-talker speech and speech-shaped noise. The two-talker speech stimulus consisted of sentences from the Harvard Institute of Electrical and Electronics Engineers corpus (Rothauser & Maiwald 1969) spoken by a male talker. Sentences were concatenated without overlap to form two independent streams, which were then overlaid and offset in the time domain by 1 second to minimize silent temporal gaps. This manipulation resulted in a continuous stream of two-talker male speech (Fig. 1A, top and bottom). We selected this stimulus because it has been previously shown to interfere with other verbal tasks, such as online speech processing (Grieco-Calub et al. 2009) and novel word learning (McMillan & Saffran 2016). The speech-shaped noise stimulus was generated by multiplying the spectral envelope of the two-talker speech by broadband Gaussian noise (Fig. 1B, top and bottom). This manipulation effectively equated the spectral composition of the two types of background noise (Fig. 1C) but removed any meaningful verbal content in the speech-shaped noise. Both maskers were scaled to 60 dB SPL and matched for average root-mean-square amplitude. The output level of each masker was verified in the sound field using a sound-level meter with the microphone positioned 48” from the loudspeaker to approximate the location of the child during the testing phase, as described in the Testing Environment section.

Fig. 1. A, top: waveform of two-talker speech. Bottom: first 1000 msec of the waveform in (A). B, top: waveform of speech-shaped noise. Bottom: first 1000 msec of the waveform in (B). C, Spectrograms of the two-talker speech (black) and speech-shaped noise (gray).
Children were then familiarized to the MST in the absence of any background noise (i.e., quiet). First, children were shown two animals and asked to label them aloud (i.e., “What animal is this?”). Children were given approximately 10 sec to look at and verbally label each animal. The researcher then placed both animals behind the barn out of view from the children, waited for approximately 2 sec, and removed one animal from the barn, leaving the other hidden. Children were then asked to recall the label of the animal that remained hidden behind the barn (i.e., “Which animal is still hiding?”). The researcher repeated this procedure with feedback until all children correctly performed the MST.

**Test Phase** • The test phase immediately followed the familiarization phase. During testing, children were shown a set of animals, starting with a baseline set size of three animals (Fig. 2, left). Each trial during testing followed the same procedure as described above, where children were asked to label the animals, the animals were placed behind the barn (Fig. 2, center), all but one animal was returned (Fig. 2, right), and children were asked to produce the label of the animal that was still hiding. If children correctly recalled the hidden animal, the procedure was repeated with an incremental increase in the set size of animals (i.e., 3 → 4 → 5…). If children incorrectly recalled the hidden animal, another trial of the same set size occurred. The test phase continued until children were unable to recall the hidden animal on two consecutive trials of the same set size. During the noise conditions, both the two-talker speech and speech-shaped noise were presented continuously throughout the entire duration of the test block. Children’s span score for each condition was measured as the largest set size for which they correctly recalled the hidden animal.

**Results**

Seventeen of the children completed the modified MST in quiet alone to ensure that our implementation of the MST yielded similar results to existing studies (e.g., Roman et al. 2014). Individual and average data are presented in Figure 3. Children’s span scores ranged from 2 to 8 items and averaged 4.5 ± 1.6 items (mean ± SD). This range was similar to the range previously observed in similarly-aged children (Roman et al. 2014).

Thirty-two children completed the MST twice, once in quiet and once in the presence of either two-talker speech (Fig. 4A; n = 16) or speech-shaped noise (Fig. 4B; n = 16). The mean span scores, grouped by condition, are listed in Table 1. A mixed
analysis of variance (ANOVA) was used to test for a within-subjects effect of listening condition (quiet; background noise) and between-subjects effects of background noise type (two-talker speech; speech-shaped noise) and task order (quiet first; background noise first) for each group. The results of the ANOVA revealed statistically significant main effects of listening condition \[ F(1,28) = 4.57; \ p = 0.041; \ \eta^2_p = 0.14 \] and background noise type \[ F(1,28) = 12.24; \ p = 0.002; \ \eta^2_p = 0.30 \]. The following effects were not found to be statistically significant: a main effect of task order \[ F(1,28) = 3.39; \ p = 0.076; \ \eta^2_p = 0.11 \], a listening condition × background noise interaction \[ F(1,28) = 3.84; \ p = 0.06; \ \eta^2_p = 0.12 \], and a listening condition × order interaction \[ F(1,28) = 3.18; \ p = 0.086; \ \eta^2_p = 0.10 \].

These results suggest that children’s span scores were disrupted by the presence of background noise but that two-talker speech and speech-shaped noise had differential effects on performance. To further probe these relations, we performed separate mixed ANOVAs to test for a within-subjects effect of listening condition (quiet; background noise) and a between-subjects effect of task order for each group. In the group of children who were exposed to two-talker speech, the mixed ANOVA revealed a significant main effect of listening condition \[ F(1,14) = 7.5; \ p = 0.016; \ \eta^2_p = 0.348; \text{Fig. 4A, top} \], suggesting that two-talker speech disrupted children’s working memory span on the MST relative to their performance in quiet. Neither the main effect of test order \[ F(1,14) = 2.2; \ p = 0.159; \ \eta^2_p = 0.137 \] nor the listening condition × order interaction \[ F(1,14) = 0.01; \ p = 0.907; \ \eta^2_p = 0.001 \] were statistically significant (Fig. 4A, bottom). In contrast, in the group of children who were exposed to speech-shaped noise, the mixed ANOVA failed to reveal a main effect of listening condition \[ F(1,14) = 0.02; \ p = 0.895; \ \eta^2_p = 0.001; \text{Fig. 4B, top} \]. Although the main effect of test order was also not statistically significant \[ F(1,14) = 1.52; \ p = 0.238; \ \eta^2_p = 0.098 \], there was a statistically significant listening condition × test order interaction \[ F(1,14) = 6.53; \ p = 0.023; \ \eta^2_p = 0.318; \text{Fig. 4B, bottom} \].

The presence of a listening condition × test order interaction suggests that children’s memory span between the quiet and speech-shaped noise conditions was dependent on the order to which they were assigned. Specifically, children consistently performed better on the first task. There were no differences in age \[ t(14) = 0.42; \ p = 0.48 \], receptive vocabulary \[ t(14) = 0.76; \ p = 0.46 \], or the number of animals verbally labeled during familiarization \[ t(14) = -0.44; \ p = 0.66 \] between children who were assigned to Order 1 and children who were assigned to Order 2 (Table 1). Therefore, it is unlikely that these factors contributed to the observed order effect. One final observation was that children in Order 1 tended to have higher span scores in quiet (i.e., first task) than in the presence of speech-shaped noise. Although there was not a statistical difference in performance between the two conditions \[ t(7) = 1.93; \ p = 0.095 \], the small sample size of the subgroup and low statistical power...
limits us from fully understanding the effects of speech-shaped noise on children’s span score.

The results from Experiment 1 suggest that there is something unique about two-talker speech that disrupts working memory span during a free recall task in 5-year-old children. This interpretation, however, is tentative given the inconclusive effect of speech-shaped noise on children’s performance in the MST, for the reasons described above. Specifically, the presence of the listening condition × test order interaction makes it difficult to interpret the results. One possible reason for the interaction is that children who performed the MST in quiet after performing the MST in the presence of speech-shaped noise experienced proactive interference. Specifically, performing the MST in the presence of speech-shaped noise taxed the children’s attentional resources, which ultimately limited children’s ability to subsequently perform the task in quiet. This finding would be consistent with the idea that younger children are more susceptible to proactive interference (Kail 2002), which may be related to smaller memory span (Kane & Engle 2000). It is unclear, however, why this same proactive interference was not observed in the children who were assigned to the two-talker speech condition. Children assigned to the two-talker speech and speech-shaped noise conditions were matched on age (t(50) = −1.75; p = 0.091), receptive vocabulary (t(25.2) = −0.194; p = 0.85 following adjustment for unequal variances), and the number of animals verbally labeled during familiarization (t(30) = 0.0; p = 1.0; Table 1).

Another possible explanation for the difference in performance between children in the two-talker speech and speech-shaped noise conditions is that children in the latter condition were overall poorer performers. Thus, in the presence or absence of proactive interference, they may have experienced floor effects on the MST, which made it difficult to assess the effect of speech-shaped noise.

In summary, the results of Experiment 1 make it difficult to assess the effect of speech-shaped noise on memory span in 5-year-old children. As a result, we are unable to draw a clear distinction between the effects of background speech versus nonspeech noise on recall. To address this limitation, a second experiment was conducted to test the effects of two-talker speech and speech-shaped noise on working memory span while concurrently minimizing proactive interference.

**EXPERIMENT 2**

The purpose of Experiment 2 was to test the effects of two types of background noise (i.e., two-talker speech; speech-shaped noise) on working memory span during administration of the MST. Given the results of Experiment 1, we predicted that children would perform poorer on the MST in the presence of two-talker speech but not speech-shaped noise.

**Materials and Methods**

**Participants** • Forty-five children (22 females; 61.8 ± 2.4 months; mean ± SD) participated in Experiment 2. An additional three children were recruited but did not complete the task and, therefore, are not included in this dataset. All children were typically developing, monolingual native English speakers and had normal hearing and vision per parental report. To ensure all children had age-appropriate receptive vocabulary, the ROWPVT-4 (Martin & Brownell 2010) was administered. All children obtained a standardized score within or above one standard deviation of the mean, indicating receptive vocabulary within the normal range for their age. Mean receptive vocabulary scores for children, grouped by condition, are listed in Table 1. Testing occurred within a single session that lasted approximately 1.5 hours. A parent or legal guardian of each child completed an informed consent process before study enrollment, as approved by the Institutional Review Board at Northwestern University.

**Stimuli** • The stimuli were the same as those described in Experiment 1.

**Procedure** • Experiment 2 incorporated the same stimuli and general procedures as Experiment 1. To investigate the effect of repeated administrations of the MST, all children performed the task three times: twice in quiet and then, for a third time, either in quiet or in the presence of background noise. The type of background noise implemented in the third test block was counterbalanced across children, which resulted in the following conditions: Quiet₁ + Quiet₂ + Quiet (n = 15); Quiet₁ + Quiet₂ + two-talker speech (n = 15); Quiet₁ + Quiet₂ + speech-shaped noise (n = 15). Regardless of the condition, all children received a 15-minute break between the second (i.e., Quiet) and third test blocks during which the ROWPVT-4 was administered. The successive administration of the MST allowed us to test for proactive interference on the MST in the absence of background noise (i.e., Quiet versus Quiet). However, the rationale behind adding a break between Quiet and the third test block was to minimize proactive interference, which ultimately allows us to test for differences in performance based on the condition in the third test block (i.e., quiet, two-talker speech, speech-shaped noise). As in Experiment 1, children’s span score was measured for each test block as the longest set size for which they correctly recalled the hidden animal.

**Results**

Figure 5 illustrates span scores for the three groups of children who performed the MST three times. Group 1 (left) performed the MST in quiet in all three test blocks (Quiet₁, Quiet₂, and Quiet₃). Group 2 (center) performed the MST twice in quiet and then once in the presence of two-talker speech (Quiet₁, Quiet₂, and two-talker speech). Group 3 (right) performed the MST twice in quiet and then once in the presence of speech-shaped noise (Quiet₁, Quiet₂, and speech-shaped noise). The mean span scores, grouped by condition, are listed in Table 1.

The first step in the analysis was to quantify children’s test–retest reliability on the first two administrations of the MST. For this analysis, we included the span scores from all three groups because every child performed the MST in quiet during the first two test blocks. Bivariate correlation analysis showed that performance across Quiet₁ and Quiet₂ was statistically significantly correlated (Pearson r = 0.41; p = 0.006; n = 45; Fig. 6). Compared with previous studies, this correlation is on the lower end of the range of test–retest reliability for other working memory tasks (e.g., r = 0.45 to 0.81; Waters & Caplan 2003; r = 0.53 to 0.81; Gathercole et al. 2004). In the present study, visual inspection of the distribution of span scores in each condition suggested that children’s performance declined during the second administration of the MST (i.e., Quiet₂). This observation was statistically confirmed: the average span score in Quiet₁ (4.1 ± 0.2, mean ± standard error) was statistically significantly lower than the average span score in Quiet₂ (4.7 ± 0.2; t(1.4) = 2.5; p = 0.017). The result of this analysis suggests that,
although children were able to maintain somewhat consistent performance on the MST between Quiet₁ and Quiet₂, many of them had smaller span scores in Quiet₃. It is important to note, as it is relevant for the upcoming ANOVA, that this difference was largely driven by children in Group 1. Specifically, paired t-tests showed that changes in span scores between Quiet, and Quiet, were statistically significant for children in Group 1 \( t(14) = 2.32; p < 0.05 \), but not for children in Group 2 \( t(14) = 1.39; p = 0.19 \) or children in Group 3 \( t(14) = 1.02; p = 0.33 \).

The second step in the analysis was to quantify an effect of the type of background noise (i.e., test block 3) on children’s span score. Given that a subset of children experienced a decline in their span scores after the second administration of the MST, there was a possibility that they would continue to do poorly in the third test block, regardless of the absence or presence of background noise. To minimize that possibility, children were given a 15-minute break between Quiet and the third test block. A mixed ANOVA with a priori post hoc testing (Bonferroni corrected for multiple comparisons) was used to test for a within-subjects effect of test block (1; 2; 3) and a between-subjects effect of background noise type for test block 3 (quiet; two-talker speech; speech-shaped noise). The main effect of test block was found to be close to significance at \( p = 0.05 \) \( F(2,84) = 3.08; p = 0.051; \eta^2_p = 0.068 \). Consistent with the earlier finding, span scores in Quiet₂ were lower than in Quiet₁, although the comparison did not quite reach statistical significance at a \( p < 0.05 \) \( (p = 0.058) \). This was likely due to reduced statistical power for this more complex model compared to the t-test used in the analysis described earlier. In addition, span scores in test block 3 were not statistically different from test block 1 \( (p = 0.22) \) or test block 2 \( (p = 1.0) \). However, the interpretation of this finding is not completely straightforward because the conditions in test block 3 varied across the three groups of children included in this omnibus analysis. The main effect of background noise type was found to be statistically significant \( F(2,42) = 4.92; p = 0.012; \eta^2_p = 0.19 \). Post hoc comparisons revealed that span scores from Group 2 were statistically lower than span scores from Group 3 \( (p = 0.014) \), but there was no difference in span scores between Group 2 and Group 1 \( (p = 1.0) \) or Group 3 and Group 1 \( (p = 0.075) \). Finally, a test block \( \times \) background noise interaction was not statistically significant \( F(4,84) = 1.72; p = 0.15; \eta^2_p = 0.076 \).

As in Experiment 1, the omnibus ANOVA revealed differences in children’s performance depending on the group to which they were assigned, but it is difficult to identify the reason for this observation. To further explore these data, we implemented separate repeated-measures ANOVAs with a priori post-hoc testing (Bonferroni corrected for multiple comparisons) for each group with a main effect of test block and a priori post hoc paired comparisons (Bonferroni corrected for multiple comparisons) to test for differences in performance across each test block. The goal of these separate analyses was to isolate and statistically test for differences in children’s behavior during the third test block for each group.

The results of the separate repeated-measures ANOVAs failed to identify a main effect of test block for the group of children who performed the MST in quiet across three consecutive test blocks \( [Group 1; F(2,28) = 1.19; p = 0.32; \eta^2_p = 0.078] \) or for the group of children who were exposed to speech-shaped noise in the third test block \( [Group 3; F(2,28) = 0.92; p = 0.41; \eta^2_p = 0.062] \). The results support the idea that over the course of three administrations, including a break between Quiet₁ and the third test block, children were able to maintain their performance reasonably well between the MST completed in the quiet and in the presence of speech-shaped noise. In contrast, there was a statistically significant main effect of test block for the group of children who were exposed to two-talker speech in the third test block \( [Group 2; F(2,28) = 6.31; p = 0.005; \eta^2_p = 0.31] \). Post hoc analyses showed that children’s span score was statistically lower in the presence of two-talker speech than in
quiet, \( p = 0.004 \)), but not statistically different from the span score in quiet, \( p = 0.23 \), the latter likely reflecting the subtle decline in span scores in quiet, due to proactive interference. The decline in quiet, however, was not enough for performance to be statistically different from performance in quiet, \( p = 0.56 \).

The results of Experiment 2 suggest that repeated administration of the MST may result in performance decline, although the effect size is quite small. This observation fits well within the theories of proactive interference and also directed attention fatigue (Kaplan 1995). Specifically, some children’s ability to focus on the MST in quiet, depleted cognitive resources necessary to sustain attention on the MST in quiet. Theoretically, the break provided to the children between quiet, and the third test block served as a “restorative” period that allowed children to replenish attentional resources and resume task performance after the break. This appeared to be the case for children who were assigned to the quiet and speech-shaped noise groups (i.e., Group 1 and Group 3, respectively) because there was no main effect of block, which is visually confirmed in Figure 5. The children assigned to the two-talker speech group (i.e., Group 2) had the lowest span scores in the third test block. This effect does not seem to be a reflection of group differences in age \( F(2, 42) = 1.2, p = 0.3; \eta^2_p = 0.056 \), receptive vocabulary \( F(2, 42) = 0.84, p = 0.44; \eta^2_p = 0.039 \), or the number of animals verbally labeled during familiarization \( F(2, 42) = 2.04, p = 0.14; \eta^2_p = 0.088 \); Table 1). Assuming that the children in Group 2 were also able to restore their attentional resources during the break, the results suggest that their decline in performance in the third test block was due to disruption by the background speech.

**DISCUSSION**

The results of Experiments 1 and 2 support the idea that children’s performance on a missing-item recall task is vulnerable to interference in the presence of two-talker speech, but not speech-shaped noise, despite the two stimuli having equivalent spectral and intensity characteristics. The observation that background speech was effective in disrupting recall of a hidden animal is consistent with prior studies that have investigated the ISE using more traditional serial recall methodologies in older children and adults (e.g., Salamé & Baddeley 1982; Elliott 2002; Beaman 2005) and missing-item tasks in adults (Beaman & Jones 1997; LeCompte 1996) and children (Elliott et al. 2016). The present study extends this prior work by providing novel evidence of the impact of background noise on 5-year-old children’s working memory span.

The present study also extends previous work showing that speech has a special “ability” to disrupt children’s recall of verbal material (Elliott 2002). However, the exact mechanism of this disruption in the present study is unclear. For example, the presence of two-talker speech during the execution of the MST may have disrupted the recoding of the items to be remembered, the mental manipulation of the verbal-visual input necessary to identify the hidden animal, or general attentional abilities. For the first mechanism—recoding—disruption would have occurred due to potential phonological overlap of the animal labels and the background speech. If this was the case, then interference-by-content is a reasonable mechanism by which two-talker speech disrupted children’s span scores given that both the target and background noise contained verbal content, and greater interference is expected when concurrent input overlaps in content (Neely & LeCompte 1999; Bayliss et al. 2003). The methodology implemented in the present study is unable to resolve the issue of whether children represented the animal labels phonologically. For example, it is unlikely that 5-year-old children recode visual items into a phonological code (Palmer 2000). Because children also verbally produced each label, both at familiarization and during exposure, they may have had a phonological code based on the auditory label.

Alternatively, according to the interference-by-process account, the background speech may not have necessarily interfered with recoding of the items to be recalled but rather interfered with the cognitive processes used to remember the animals’ names as they were placed behind the barn (e.g., Jones et al. 1992; Sörqvist 2010). A key observation associated with the interference-by-process account is that dynamic, changing-state noise, rather than steady-state noise, disrupts performance. One reason for the differential effect of changing-state noise is that is has been found to disrupt the rehearsal strategies necessary for serial recall. Fundamentally, the two-talker speech stimulus was more “dynamic” than the speech-shaped noise stimulus in the present study: though both stimuli were continuous and the occurrence of short-duration silent gaps was minimized in the two-talker speech by temporally offsetting the two streams of the male talkers, the two-talker speech had a temporal envelope with larger intensity variations. However, the interference-by-process account is likely not the mechanism at play because, as shown in previous work, the missing item task does not engage rehearsal strategies (Jones & Macken 1993; Elliott et al. 2016). Therefore, the observation that background speech disrupted memory span is probably not due to interference-by-process.

The likely mechanism by which the two-talker speech stimulus disrupted children’s performance on the MST in the present study is attentional capture, which would be consistent with previous studies of the ISE in older children (Elliott 2002; Elliott et al. 2016). Temporally dynamic acoustic stimuli (i.e., those that vary more widely in intensity over time or contain abrupt onsets/offsets or deviants) have been shown to draw individuals’ attention during nonauditory tasks (Elliott 2002; Klatte et al. 2010; Vachon et al. 2017). It is reasonable to suspect that children’s attention was drawn to words within the two-talker speech stream, resulting in a momentary lapse in attention to the MST. Support for this comes from previous work in adults (LeCompte 1996) and children (Elliott et al. 2016): background speech disrupted performance on a missing item task, and there was no measurable difference in the degree of disruption based on whether the speech was changing-state or steady state. The interpretation of these findings supports the idea that background speech disrupts performance even when the task does not require rehearsal strategies. The present study extends this work by showing that two-talker speech disrupts performance similarly as single-talker speech, which has largely been the stimulus of choice in the ISE literature. This finding is consistent with work from the speech perception literature showing that both single-talker and two-talker speech exerts masking effects (Freyman et al. 2001, 2004).

The finding that speech, but not nonspeech, disrupted performance is consistent with what has been observed in the auditory masking literature. Specifically, speech maskers disrupt speech perception to a greater extent than nonspeech noise, even when...
effects of overlapping frequency components are accounted for. This phenomenon has been termed informational masking (Pollack 1975; Leek et al. 1991; Brungart 2001; Freyman et al. 2004) and has been shown to be larger in children than in adults (Wightman & Kistler 2005; Leibold & Buss 2013). The present study supports the idea that one possible mechanism for increased informational masking in children is that background speech is better able to disrupt the cognitive processes that are engaged during speech perception in noise, namely working memory and selective attention. Future work is needed to probe the relation between attention, working memory, and speech perception in noise in children.

There are some nuanced aspects of the results that deserve mentioning. First, the results of Experiment 2 suggest that speech-shaped noise does not disrupt children’s ability to recall items, especially when children are allowed a break between test intervals. However, given the results of Experiment 1, we cannot rule out the possibility that broad-band noise stimuli are disruptive to some aspects of cognitive processing. Specifically, performing the MST in the presence of the speech-shaped noise might have elicited proactive interference or directed attention fatigue: inhibiting attention to the speech-shaped noise may have exerted cognitive effort (Kaplan 1995) and had a prolonging negative effect on subsequent performance (e.g., Klatte et al. 2013). However, similar proactive interference was observed between the Quiet, and Quiet, conditions in Experiment 2 and was unrelated to the presence of any background noise. In other words, the subpopulation of children who were assigned to the speech-shaped noise condition of Experiment 1 may have simply been more susceptible to proactive interference than children who were assigned to the two-talker speech condition of Experiment 1. The potential floor effects and low statistical power limit our ability to rule out an effect of speech-shaped noise on performance in certain conditions, and future studies are necessary to resolve this issue. Given that many young children spend the majority of their day in classrooms that contain multiple sources of nonspeech environmental noise (e.g., heating, ventilation, and air conditioning [HVAC] systems, traffic noise; Crandell & Smaldino 2000), the long-term impact of background noise containing nonspeech sounds on attention and task performance is a research area worth pursuing further.

The task implemented in the present study was essentially free item recall—children were not asked to maintain any information about the order of items placed behind the barn. This aspect of the methodology does not rule out the possibility that some children used strategies, such as rehearsing the order of the animals, to facilitate their recall (Beaman & Jones 1998). Anecdotally, the researchers who administered the MST often observed children verbally reciting the names of the animals during the period when the animals were hidden behind the barn. According to the phonological loop model (Baddeley 2003), visually presented stimuli can be recoded into a verbal form to be further processed by the phonological loop via a rehearsal process (e.g., Palmer 2000). This is a reasonable strategy that may have been implemented by the children, particularly because they were tasked with verbally recalling the hidden animal. If true, the results of the present study would extend the findings of Baddeley et al. (1998) who suggested that this rehearsal process, and therefore the phonological recoding of visually presented stimuli, is not developed until 7 years of age. Further studies are necessary to explore the developmental trajectory of this strategy and the conditions under which it may, or may not, be utilized.

Children depend on working memory to execute numerous cognitive functions related to language comprehension, math computations, and reading (e.g., Daneman & Carpenter 1980; Daneman & Merikle 1996; Cain et al. 2004; Gathercole et al. 2008). Understanding how working memory is impacted by background noise in early childhood may be helpful in identifying children who are at risk for delayed development in these areas. The results from the present study provide insight into the deleterious effects of background noise, specifically that contain speech, on working memory in nonauditory domains. This work has implications for how background noise in a classroom, for example, may impact cognitive processing and learning in young children. Both preschool and primary school classrooms are characteristically noisy, with much of the noise being attributed to the occupants. Competing speech from other children and adults may disrupt the online processing of input and learning of new material by engaging the attentional and working memory systems. Future work should focus on children’s ability to perform cognitive tasks in more naturalistic environments. Finally, the results from the present study provide evidence that tasks like the MST are feasible tests of memory span in children who are unable to perform serial recall.

ACKNOWLEDGMENTS

The authors thank the families who participated in the study and Dr. Beverly Wright for helpful comments on the manuscript. Partial funding for this study was provided by an Undergraduate Research Grant from Northwestern University awarded to M.-S.C. for completion of an undergraduate honors thesis. T.M.G.-C., M.-S.C., and H.E.S. designed the experiments. M.-S.C., H.E.S., and K.M.W. performed the experiments. T.M.G.-C. and K.M.W. analyzed the data. T.M.G.-C. wrote the manuscript with comments from M.-S.C., H.E.S., and K.M.W.

The authors have no conflicts of interest to disclose.

Address for correspondence: Tina M. Grieco-Calub, The Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Room 2–246, Evanston, IL 60208, USA. E-mail: tinage@northwestern.edu

Received December 1, 2017; accepted May 31, 2018.

REFERENCES


