Quantifying the effects of background speech babble on preschool children's novel word learning in a multi-session paradigm: a preliminary study

Meital Avivi-Reich

Megan Y. Roberts

Tina M. Grieco-Calub

The Roxelyn and Richard Pepper Department of Communication Sciences and Disorders

Northwestern University, IL, United States

Conflict of Interest: The authors whose names are listed above certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Funding: The first author is funded by the Lambert Translational Postdoctoral Fellowship.

Corresponding author:

Meital Avivi-Reich
Address: Northwestern University, 2240 Campus Drive
Evanston, IL 60208
Phone: 847-467-4156
Fax: 847-491-4975
Email: m.avivireich@northwestern.edu

Abstract

Purpose: The present study tested the effects of background speech babble on novel word learning in preschool children with a multi-session paradigm.

Method: Eight 3-year-old children were exposed to a total of eight novel word-object pairs across two story books presented digitally. Each story contained four novel CVC nonwords. Children were exposed to both stories, one in quiet and one in the presence of 4-talker babble presented at 0 dB signal-to-noise ratio. After each story, children's learning was tested with a referent selection task and a verbal recall (naming) task. Children were exposed to and tested on the novel word-object pairs on five separate days within a two-week span.

Results: A significant main effect of session was found for both referent selection and verbal recall. There was also a significant main effect of exposure condition on referent selection performance, with more referents correctly selected for word-object pairs that were presented in quiet compared to pairs presented in speech babble. Finally, children's verbal recall of novel words was statistically better than baseline performance (i.e., 0%) on sessions 3-5 for words exposed in quiet, but only on session 5 for words exposed in speech babble.

Conclusions: These findings suggest that background speech babble at 0 dB SNR disrupts novel word learning in preschool-age children. As a result, children may need more time and more exposures of a novel word before they can recognize or verbally recall it.

Introduction

Young children's naturalistic environments are often not acoustically ideal for listening (Crukley, Scollie, & Parsa, 2011). A common source of background noise in these environments is "other" talkers, whose speech may interfere with children's ability to perceive and process the speech of a target talker (e.g., Elliott, 1979; Fallon, Trehub, & Schneider, 2000; Grieco-Calub, Saffran, & Litovsky, 2009; Leibold, Bonino, & Buss, 2016). If background speech disrupts perception and processing of *known* words, then it is reasonable to expect that it disrupts the *acquisition of novel words.* The purpose of the present study was to test novel word learning in preschool-age children in the presence of a 4-talker speech babble.

Novel word learning depends on the ability of children to encode the phonological representation of the word form and store it in long-term memory for future access (e.g., Gathercole & Baddeley, 1993). Speech babble may disrupt this process through various mechanisms. First, the spectral energy of the speech babble may overlap with that of the target speech, causing *energetic masking* (Pollack, 1975). As a result, children may perceive only a partial representation of the novel words on any given exposure. Second, speech babble may distract children's attention from the target speech, causing *informational masking*. Speech babble contains meaningful content that engages children's attention and working memory, which are essential cognitive processes that are necessary for linguistic and semantic processing of the target speech (e.g., Brungart, Simpson, Ericson, & Scott, 2001; Freyman et al., 2004; Schneider, Li, & Daneman, 2007). Third, speech babble may require children to exert more cognitive effort to resolve the target speech, which results in fewer cognitive resources to process and store the speech input into memory (Kahneman, 1973; Rönnberg et al., 2013; Avivi-Reich, Daneman, & Schneider, 2014). For example, speech babble may disrupt children's ability to phonologically encode, rehearse, and retain spoken material even when the listener can correctly recognize what was said (Murphy, Craik, Li, & Schneider, 2000; Pichora-Fuller, Schneider, & Daneman, 1995).

When studying the possible effects of speech babble on children's ability to perceive and recall spoken words, there are at least four previous findings that should be considered. First, children demonstrate greater speech perception difficulties when listening in the presence of background speech: young children require higher (easier) signal-to-noise ratios (SNRs) to achieve the same speech perception performance as teenagers and young adults (Choi, Lotto, Lewis, Hoover, & Stelmachowicz, 2008; Elliott, 1979; Fallon et al., 2000). Second, young children have a relatively limited working memory capacity compared to that of young adults (Choi et al., 2008; Gathercole, 1998; Irwin-Chase & Burns, 2000). Because background speech disrupts verbal working memory in children (Elliott, 2002; Grieco-Calub, Collins, Snyder, & Ward, 2018), children may have fewer working memory resources to process novel word forms that are presented in speech babble. Third, attentional control skills develop with age and are not fully developed even in 10-to-11-year-old children (Cowan, Fristoe, Elliott, Brunner, & Saults, 2006). These skills are required when attempting to inhibit competing speech stimuli while attending to the target speech. Fourth, children's prior knowledge is not as expanded as that of adults and, therefore, children are likely to be less skilled at using top-down knowledge-driven processes to restore speech perception (Newman, 2006). Thus, children's ability to learn novel words may be compromised by the presence of noise, and the effect should be larger for younger children.

Only a small number of studies have examined the effect of background noise, including speech babble, on novel word learning in preschool-age children (e.g., Han, Storkel, &

Bontempo, 2019; Dombroski & Newman, 2014; McMillan & Saffran, 2016). Somewhat inconsistent with the ample evidence demonstrating a negative noise effect on speech recognition and memory, young children's ability to learn novel words does not seem to be significantly affected by the presence of background noise, especially when the target speech is louder than the competing sound source(s) (Dombroski & Newman, 2014; McMillan & Saffran, 2016). Children as young as 2.5-years-old and as old as 10-years-old demonstrate similar novel word recognition (referent selection) abilities for words they have learned, independently of whether there was noise present during exposure (Blaiser, Nelson, & Kohnert, 2014; Riley & McGregor, 2012). Studies which examined the effect of background noise on novel word production (e.g., verbal recall) have revealed an inconsistent impact of noise on learning. For example, Riley and McGregor (2012) tested word learning in 9- to 10-year-olds and found that words learned in speech-shaped noise produced less accurate utterances compared to similar words learned in quiet. However, in a study which examined the effect of multitalker background babble on new word recognition and production in 2.5- to 6-year-old children with and without hearing loss, no significant effect was found (Blaiser et al., 2014). These inconsistent results likely reflect the varied experimental methodologies implemented across studies. Differences across studies include: (1) the number of exposures to novel label-object pairs; (2) the number of experimental sessions; (3) the type of background noise (e.g., speech, non-speech); and (4) how children's learning is tested (referent selection, verbal recall). In addition, the majority of studies implement a between-subjects manipulation, which does not allow control of intersubject variability across conditions and may have confounded the condition effect (e.g., quiet versus noise). Moreover, most of the studies which examined the effect of noise on novel word learning in children were conducted over a single session, which might have not allowed enough time for consolidation between exposures.

The current study attempts to extend prior work by exploring the effects of speech babble on novel word learning in preschool-aged children with a multi-session, within-subjects design. Each child learned four novel words in one story presented in quiet and another four novel words in a second story presented in the presence of speech babble. Children's ability to remember the new novel words during both referent selection and verbal recall tasks was investigated. We hypothesized that speech babble interferes with novel word learning. We predicted that children would learn more novel words in the quiet condition than in the speech babble condition. We did not predict learning of control novel words that were not introduced during the exposure phase.

Methods

Participants

Eight native-English monolingual toddlers (3 males and 5 females), between 36 and 41 months of age, participated in this study (mean age ± s.d. = 38.75 ± 1.75 years). Participants were recruited from the Child Studies Group Database Registry at Northwestern University, which is located in a north suburb of Chicago, IL. Children had no history of visual impairment, hearing loss, language delay, or developmental disorders per parental report. All children had normal receptive and language vocabulary. This was confirmed with the Receptive One-Word Picture Vocabulary Test (ROW-PVT) and Expressive One-Word Picture Vocabulary Test (EOW-PVT; Martin & Brownell, 2010), in which all participants were found to be ranked in the 45^{th} percentile or over. Informed consent was obtained from each parent or guardian prior to any testing. Each participant participated in five sessions and was monetarily compensated. All procedures were approved by the Institutional Review Board at Northwestern University.

Exposure Stimuli

Auditory stimuli. The novel words were sixteen monosyllabic nonwords that were selected from an established corpus of native English words and nonwords (Storkel, 2013). The novel words were nonwords composed of sounds acquired in early-to-mid childhood and had a consonant-vowel-consonant (CVC) structure. The novel words were arranged into four subsets (A_1: *Nup, Mab, Gok, Tej*; A_2: *Jeg, Tib, Pom, Kun*; B_1: *Gim, Tak, Bop, Nuj*: B_2: *Nog, Pid, Mut, Keb*). In each subset, no vowel or consonant appeared more than once. In addition, in each set of eight novel words (A, B), each of the CV or VC combinations were unique. Each of the novel words' phonotactic probability and neighborhood density was calculated based on child corpora of spoken American English (Storkel & Hoover, 2010) and are presented in Appendix A.

The subsets of novel words were presented in the context of two age-appropriate stories that were created for this study. Each story included the same four children as characters. In one story, the children were asked to each bring one item to their grandparent's house for a sleepover, and in the second story, the children were asked to bring items to a "Show and Tell" at their daycare. The two stories followed a similar structure and were similar in length (*Sleepover*: 223 words, 1.38 min; *Show and Tell*: 227 words, 1.45 min). Within each story, each novel word was presented three times. Complete stories are presented in Appendix B.

Each story was recorded twice: the "Sleepover" story contained the novel word subset A_1 or A_2, and the "Show and Tell" story contained the novel word subset B_1 or B_2. A female talker read each story in an infant-directed register while she was audio-recorded in a double-walled sound-attenuating chamber with PRAAT software using a Blue Snowball microphone at a 44.1kHz sampling rate and 16-bit precision. The RMS of all story recordings was set to 60 dB

SPL using PRAAT software[1]. A total of eight auditory versions of the two stories were created: four in quiet and four in the presence of speech babble.

For the speech babble condition, each recorded story was mixed with 4-talker babble. The number of competing talkers used in the current study falls within the range of talker number found within the literature (Grieco-Calub et al., 2009; McMillan & Saffran, 2016; Dombroski & Newman, 2014). To create the 4-talker babble, two female and two male native-English speakers speaking sentences from the IEEE/Harvard corpus (Rothauser & Maiwald, 1969) were overlaid. The RMS of the 4-talker babble was set to 60 dB SPL using PRAAT software, to create a signal-to-noise ratio (SNR) of 0 dB. This challenging SNR is similar to noise levels found in occupied preschool environments (e.g., Crukley et al., 2011).

Visual stimuli: Each novel word was matched to one of sixteen unidentified objects that were selected either from the Novel Object and Unusual Name (NOUN) Database (Horst & Hout, 2016) or from publicly-available images on the internet (see Appendix C). All images were of unfamiliar objects that ten young adults were unable to identify. Each story described above was illustrated into a book by the first author. Each page contained black and white drawings depicting each of the scenes described in the story. Each drawing included only one of the novel objects, except the first and the last drawings (introduction and conclusion) which presented none. The objects were the only colorful items in the scenes, and their width and length were set to be between 1.5 to 2.5 inches. Children were presented with each storybook via a video recording to control for both auditory and visual exposure. A video of a female adult flipping through the pages from an "over-the-shoulder" perspective was recorded for each of the two stories. The exact angle and position of the books as they appeared on the screen were

[1] PRAAT's script does not excludes any pauses when calculating the RMS. Thus, the averaged of the entire sound file was set to 60 dB SPL.

maintained during filming (see Appendix D). These recordings were mixed with the different versions of the audio recordings to create an overall experience similar to an adult reading a book to a child either in quiet or in a background of competing talkers.

Test stimuli

Auditory stimuli: The same female talker who recorded the stories also recorded the sentence, "Where is the _____?", for each of the novel words and four known words (dog, duck, shoe, ball). Recording parameters were similar to those described for the exposure stimuli. Each question was presented in quiet or mixed with the 4-talker babble.

Visual stimuli: Each test question was presented while the child simultaneously saw a display containing four pictures, each located at one corner of the touch screen in a white square (11 cm wide and 11 cm long) on a black background. The four pictures were all objects paired with known words or novel words from the same category (known words/ exposed novel words (target)/unexposed novel words (control)). Visual examples of screen shots taken during the referent selection testing phase are shown in Appendix E.

Procedure

Each child was seated in a chair located in the center of a double-walled sound-attenuated chamber. Stimuli were presented simultaneously over three loudspeakers, which were placed one meter away and at 0° and ±45° azimuth relative to the child's location. The height of the loudspeakers was adjusted to match the ear level of a seated child of an average body height for the age tested. The videos and the images were presented on a 24-inch P2314Tt flat panel Dell touch screen monitor placed 0.5 meters in front of the child. Note that the placement of the monitor did not alter the measured intensity level of the stimuli at the location of the child.

Stimuli were presented and responses were collected by customized server side scripting language software on a Hypertext Pre-processor (PHP) platform.

At each session, children heard each story: one in quiet and the other in speech babble. Stories were counterbalanced across conditions across children. In addition, subsets of novel word-object pairs were counterbalanced across stories. The two subsets of novel words that the child was not exposed to were used as control words to evaluate the effect of exposure on learning. It These assignments were the same at each session for each child. At the beginning of each session, children were instructed to listen to and watch the story as it played on the monitor. After each story, children participated in a closed-set referent selection task and an open-set verbal recall task. The sessions were held on five separate days within a two-week period.

Novel word referent selection task. The task was a four-alternative-forced-choice task with one trial dedicated to each known word and each novel word assigned to the story (12 trials in total). On each trial, children were shown one of the subsets of words (exposed novel words, non-exposed novel words, or known words) on a touchscreen while they heard, “Where is the ______?” and were instructed to point to the object that matched the spoken label. The software program registered only touch which occurred after the end of the question. No feedback was provided to minimize the contribution of exposure during testing to learning. Regardless of exposure condition (quiet or speech babble), half of the trials were presented in quiet and half in speech babble (0 dB SNR). This two-by-two matrix design in which two of the four novel words which are learned in quiet are tested in speech babble and two of the four novel words which are learned in speech babble are tested in quiet was chosen to examine the effect of exposure condition versus test condition. This design resulted in four test conditions for the subsets of novel words used in the stories: 1) words exposed in quiet and tested in quiet (E_QT_Q); 2) words exposed in quiet and tested in speech babble (E_QT_B); 3) words exposed in speech babble and tested in speech babble (E_BT_B); and 4) words exposed in speech babble and tested in quiet (E_BT_Q). In addition, half of the non-exposed novel words and half of the known words were tested in quiet and half in speech babble. The test trials were blocked by test condition (quiet or speech babble). Each block contained trials for exposed novel words, non-exposed novel words, and known words. Order of test condition and trials was set randomly by the software program at the beginning of every test phase.

Novel word verbal recall task. At the end of each video presentation and the referent selection task, the tester presented images of all 4 target objects presented in the video and asked the child to pronounce the name of each object. The child was encouraged to name the object even if he/she was uncertain as to how it is called. Once the child responded, no feedback was provided, and the tester moved on to the next object. The child's production was phonetically written, and the novel word was scored as correct when correctly pronounced as a whole.

Results

Known word referent selection (receptive recognition)

Referent selection accuracy (percent correct) in the presence of the speech babble was near ceiling performance for each of the known words: Ball (95% ± 3.49 SE), Duck (97.5% ± 2.5 SE), Dog (97.5% ± 2.5 SE), and Shoe (100% ± 0 SE). Accuracy was calculated for each of the four known words based on the total number of responses collected from all eight participants across the five experimental sessions. The rationale for including these trials was twofold: first, it allowed us to familiarize the children with the four-alternative-forced-choice task; and second, it gave us a measure of speech-in-noise recognition. The children's high accuracy on this task suggests that they understood the task and that the target words were audible, even with a modest SNR.

Novel word referent selection (receptive recognition)

Figure 1A shows the average number of novel words that were correctly identified during the referent selection task for novel words that were exposed in quiet and for those that were exposed in the presence of speech babble. Individual accuracy data for each word across sessions, conditions, and participants are presented in Figure 1B. These raw data highlight the source of variability in novel word recognition, both within and across participants. Of note, the data support the idea that the majority of children recognize more of the words presented in quiet, versus those presented in the presence of speech babble, by the end of the five sessions.

Single-sample t-tests that examined the accuracy of the exposed novel words on the referent selection task showed that the performance exceeded chance performance (i.e., 25%) on each of the five experimental sessions when the novel words were exposed in quiet (*1*: $t(15) = 3.05$, $p = .008$; *2*: $t(15) = 4.198$, $p < .001$; *3*: $t(15) = 5.056$, $p < .0001$; *4*: $t(15) = 7.319$, $p < .0001$; *5:* $t(15) = 7.889$, $p < .0001$). However, the performance of the eight participants exceeded chance performance for the referent selection task only on the third experimental session when the novel words were exposed in speech babble ($t(15)=2.573$, $p = .021$).

The number of novel words that were correctly identified were subjected to a Repeated-Measures Analysis Of Variance (RMANOVA) with training sessions (1-5), exposure condition (E_B vs. E_Q) and testing condition (T_B vs. T_Q) as within-subjects variables. The RMANOVA revealed a significant main effect of session ($F[4,28] = 4.619$, $p = .005$, $\eta^2p = .398$), as well as a significant main effect of exposure condition ($F[1,7] = 7.859$, $p = .026$, $\eta^2p = .529$). A post-hoc power analysis revealed that the statistical power to detect these effect sizes with a sample size of eight children was 0.9 and 0.99, respectively. Neither a main effect of test condition ($F[1,7] = .135$, $p = .724$, $\eta^2 p = .019$) nor any of the interactions were statistically significant ($p > .05$).

--Figure 1 here--

To better understand the course of the novel word learning across the five experimental learning sessions, a post-hoc Helmert contrast analysis for the RMANOVA was conducted. This contrast analysis compares each session to the succeeding sessions (Session 1 vs. Sessions 2-5; Session 2 vs. Sessions 3-5; etc.). When all the conditions were included in the analysis (E_BT_B, E_BT_Q, E_QT_Q, E_QT_B), the results showed a significant difference between the first session and the succeeding sessions (i.e., 2-5; $F[1,7] = 13.496$, $p = .008$, $\eta^2 p = .658$) as well as between the second session and the succeeding sessions (i.e., 3-5; $F[1,78] = 6.333$, $p = .04$, $\eta^2 p = .475$). No significant differences were found between the last three sessions (*p*s > .05).

A RMANOVA with test condition (T_B vs. T_Q) and session (1-5) as within-subjects variables was conducted to test whether any improvement in referent selection was evident for words that children did not hear during the exposure phase. As expected, the results revealed no significant main effect of session ($F[4,28] = 1.144$, $p = .356$, $\eta^2 p = .141$) or test condition ($F[1,7] = .104$, $p = .756$, $\eta^2 p = .015$) as well as no significant interaction between the two ($F[4,28] = .201$, $p = .936$, $\eta^2 p = .028$). In addition, single-sample t-tests which examined the correct referent selection of the unexposed novel words, showed that the performance of all eight participants on all five sessions did not exceed chance performance (i.e., 25%)) when the test condition was quiet (Mean = 26.25, $t(39)= .388$, $p = .7$) or when it was speech babble (Mean = 27.5, $t(39) = .781$, $p = .44$).

Novel word verbal recall (expressive production)

The average number of novel words correctly produced at each session under each of the two exposure conditions conducted (quiet, speech babble) are presented in Figure 2A. Visual inspection of the data suggests that children failed to verbally recall novel words when shown their corresponding images in the first two sessions, regardless of whether the children were exposed to the words in quiet or in the presence of speech babble. However, during the later sessions, children were better able to verbally recall novel words, and they seem to be more successful at producing those words which were exposed in quiet versus those exposed in speech babble. These preliminary observations were confirmed statistically. Single-sample t-tests revealed children recalled words that were exposed in quiet at levels that were statistically better than baseline performance (i.e., 0%) on session 3 ($t(7) = 3.862$, $p = .003$), session 4 ($t(7)=3.055$, $p = .009$), and session 5 ($t(7) = 3.24$, $p = .007$). In contrast, children recalled words that were exposed in the presence of speech babble at levels that were statistically better than baseline performance only on session 5 ($t(7) = 2.376$, $p = .0246$). To evaluate the combined effects of session and condition on performance, a RMANOVA with exposure condition (E_B vs. E_Q) and session (1-5) as within-subjects variables was conducted. The results of that analysis revealed a significant main effect of session ($F[4,28] = 16.22$, $p < .001$, $\eta^2p = .699$); however, both the main effect of exposure condition ($F[1,7] = 5.223$, $p = .056$, $\eta^2p = .427$) and the two-way interaction between session and exposure condition ($F[4,28] = 2.116$, $p = .105$, $\eta^2p =.232$) were not statistically significant.

Individual accuracy data for each word across sessions, conditions, and participants are presented in Figure 2B. These raw data highlight the source of variability in novel word recall, both within and across participants. Overall, participants were better able to verbally recall more words that were exposed in quiet compared to words presented in babble. Seven out of eight of the participants were unable to correctly recall a novel word without cueing until the third experimental session.

---Figure 2 here---

Discussion

The overarching goal of this preliminary study was to determine how background speech babble affects novel word learning in preschool-age children across multiple sessions. The results of the study suggest that speech babble at 0 dB SNR disrupts young children's ability to stabilize their representation of novel words. The following discussion will highlight aspects of the study that are novel contributions to the field as well as limitations that should be considered for future studies.

The results of the present study suggest that background speech babble disrupts children's ability to map novel label-object pairs, as quantified by the referent selection task, and may slow children's ability to recall novel words, as implied by the free recall task. Inspection of the raw data (Figures 1B, 2B) shows large individual variability in performance. However, consistent with our predictions, the data reveal that children recognized and recalled words that were exposed in quiet with higher accuracy than words exposed in the presence of speech babble. These results add to the growing body of research that has inconsistently found an effect of background speech on word learning in preschool-age children (e.g., McMillan, & Saffran, 2016; Dombroski & Newman, 2014). The findings from the present study suggest that one way to better define the effects of background speech on word learning is to assess the effect over multiple sessions. Support for this idea comes from the observation that children needed repeated exposures to solidify their representation of the label-object pair even in quiet. In addition, spacing exposures to new information (i.e., the "spacing effect") leads to better retention of the new information, as it allows time for consolidation (e.g., Williams & Horst, 2014; Vlach, Sandhofer, & Kornell, 2008). Specifically, the spacing effect has been shown to result in better acquisition and generalization of new information compared to learning from sequential exposures in a single session (Williams & Horst, 2014). The observation that consolidation was weaker in the presence of speech babble supports the idea that the speech babble likely disrupted encoding either directly, or by exhausting the attentional resources in aid of speech perception. Future studies are necessary to determine the source of the disruption.

Another benefit of the present study was the implementation of a within-subject design, which served as a critical experimental control that has not been implemented in previous studies of word learning in preschool-age children. For example, the within-subject design provided an opportunity to quantify trajectories of novel word acquisition in different conditions *within* the same children, providing us with more statistical power than between-subjects designs. In addition, the within-subject design reduced the intersubject variability that may have confounded a condition effect (e.g., quiet versus noise) on the referent selection task in previous studies (e.g.,Riley & McGregor, 2012; Dombroski & Newman, 2014). In addition, considering the relatively high inter-subject variability, it is also advisable to use larger sample sizes in future studies.

Aspects of the stimuli should also be considered when interpreting the results within the context of the extant research. First, the SNR implemented in the present study (i.e., 0 dB SNR) is less favorable than the SNRs used in most previous word learning studies (e.g., Dombroski & Newman, 2014; McMillan & Saffran, 2016), but similar to noise levels found in common preschool listening environments (e.g., Crukley et al., 2011). Thus, it is possible that the children in the present study gleaned less of the novel word on each exposure. Learning of the novel words exposed in the presence of the speech babble was evident, albeit at a slower rate in both the referent selection and verbal recall tests than for the novel words exposed n quiet. This observation suggests that children were able to build more accurate representations of the words over time. Additional studies are necessary to probe the relation between exposure and children's representation in greater detail. Second, the use of 4-talker speech babble as the background noise not only caused energetic masking but also informational masking (Brungart et al., 2001). For example, speech babble fluctuates in both intensity and spectral content over time, which may be more distracting to children (e.g., Elliott, 2002; Aubuchon, McGill, & Elliott, 2018). In summary, the segmented glimpses of the novel words during exposure in the presence of 4-talker speech babble—due to perceptual overlap with the speech babble or attentional gaps—likely contributed to the results of the present study.

A final issue to consider when planning future studies is the tasks used to test learning following exposure. Most studies that assess word learning in children familiarize the children with several novel word-object pairs and subsequently test their knowledge of the pairs by asking children to select the correct referent object from a limited array when it is named. This referent selection task demonstrates the children's knowledge of the association between each word form and its referent. However, encoding a partial word form could be sufficient for correct recognition of the object. This lack of task sensitivity is of special interest considering previous findings suggesting that infants only encode partial representations of forms after a limited number of exposures (Kay-Raining Bird & Chapman, 1998). Recent evidence suggests that quantifying children's verbal production of novel words may be a more sensitive measure of the impact of background noise on novel word form representation (Han, Storkel, & Bontempo, 2019). These verbal recall tasks provide valuable information regarding how novel word forms are encoded in memory but may be difficult to implement in younger children who are less willing to verbally respond or who have underdeveloped articulation skills. Thus, task selection will likely influence the extent to which we are able to measure a child's true representation of a novel word form because each task may tap separate aspects of the word learning process. Studies using a multi-tiered approach may provide a more precise measure of word learning (Gordon & McGregor, 2014), which may be necessary when exploring word learning in different types of listening environments.

In light of the present results and the issues raised, it is of great importance to continue investigating the possible effects of background noise on the ability of children to learn novel words. Since most word learning studies take place in quiet laboratory settings, they might lack ecological validity. Such studies may provide valuable information as to the word-learning mechanism (see review by Swingley, 2010) but are limited in their ability to fully capture this complex process as it would occur in common noisy environments. Given the prevalence of noise in young children's environment (American Speech Language Hearing Association [ASHA], 2005), it is crucial to continue exploring the effect of noise on word learning over multiple experimental sessions.

References

American Speech-Language-Hearing Association. (2005). Acoustics in educational settings: Position statement.

Aubuchon, A., Mcgill, C., & Elliott, E. (2018). Auditory distraction does more than disrupt rehearsal processes in children's serial recall. *Memory & Cognition,* Memory & Cognition, 11/29/2018.

Avivi-Reich, M., Daneman, M., & Schneider, B.A. (2014). How age and linguistic competence alter the interplay of perceptual and cognitive factors when listening to conversations in a noisy environment. *Frontiers in Systems Neuroscience*. https://doi.org/10.3389/fnsys.2014.00021

Brungart, D., Simpson, B., Ericson, M., & Scott, K. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *The Journal of the Acoustical Society of America, 110*(5), 2527-2538.

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.

Cowan, N., Fristoe, N. M., Elliott, E. M., Brunner, R. P., & Saults, J. S. (2006). Scope of Attention, Control of Attention, and Intelligence in Children and Adults. *Memory & Cognition*, *34*(8), 1754–1768.

Crukley, J., Scollie, S., Parsa, V. (2011). An exploration of non-quiet listening at school. *Journal of Educational Audiology*, 17, 23-35.

Dombroski, J., & Newman, R. S. (2014). Toddlers' ability to map the meaning of new words in multi-talker environments. *The Journal of the Acoustical Society of America, 136,* 2807–2815. https://doi.org/10.1121/1.4898051

Elliott, L. L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *Journal of the Acoustical Society of America, 66,* 651–653.

Elliott, E.M. (2002). The irrelevant-speech effect and children: Theoretical implications of developmental change. *Memory and Cognition, 30 (3)*, 478–487.

Fallon, M.1., Trehub, S.E., & Schneider, B.A. (2000). Children's perception of speech in multitalker babble. *The Journal of the Acoustical Society of America 108(6),* 3023-3029.

Freyman, RL, Balakrishnan, U., & Helfer, K.S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *Journal of the Acoustical Society of America, 115,* 2246-2256.

Gathercole, S. (1998). The development of memory. Journal of Child Psychology, 39, 3–27.

Gathercole, S. E., & Baddeley, A. D. (1993). Short-term memory may yet be deficient in children with language impairments: A comment on van der Lely & Howard. *Journal of Speech and Hearing Research, 38,* 463–466.

Gathercole, S.E., Hitch, G.J., Service, E., & Martin, A.J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology, 33,* 966–979.

Gordon, K., & McGregor, K. (2014). A spatially supported forced-choice recognition test reveals children's long-term memory for newly learned word forms. *Frontiers in psychology*, 5, 164. https://doi.org/10.3389/fpsyg.2014.00164.

Grieco-Calub, T.M., Collins, M., Snyder, H.E.W, Ward, K.M. (2018). Background Speech Disrupts Working Memory Span in 5-Year-Old Children. *Ear and Hearing*. https://doi.org/10.1097/AUD.0000000000000636

Grieco-Calub, T.M., Saffran, J.R., & Litovsky, R.Y. (2009). Spoken word recognition in toddlers who use cochlear implants. *Journal of Speech Language and Hearing Research, 52(6),* 1390-1400. https://doi.org/10.1044/1092-4388(2009/08-0154).

Han, M., Storkel, H., & Bontempo, D. (2019). The effect of neighborhood density on children's word learning in noise. *Journal of Child Language, 46(1),* 153-169. https://doi.org/10.1017/S0305000918000284

Horst, J.S., & Hout, M.C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods, 48,* 1393-1409. https://doi.org/10.3758/s13428-015-0647-3

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,* 61–85.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kay-Raining Bird, E., & Chapman, R.S. (1998). Partial representations and phonological selectivity in the comprehension of 13- to 16-month-olds. *First Language, 18 (52),* 105 – 127. https://doi.org/10.1177/014272379801805204

Leibold, L.J., Yarnell Bonino, A., & Buss, E. (2016). Masked Speech Perception Thresholds in Infants, Children, and Adults. *Ear and Hearing, 37(3)*, 345-353.

Martin, N.A., & Brownell, R. (2010). Expressive and receptive one word picture vocabulary test-4 (ROWPVT-4). Novato, CA: Academic Therapy.

McMillan, B., & Saffran, J. R. (2016). Learning in complex environments: the effects of background speech on early word learning. *Child Development, 87(6),* 1841-1855.

Murphy, D. R., Craik, F. I. M., Li, K. Z. H.,& Schneider, B.A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychology and Aging, 15,* 323–334.

Newman, R. (2006). Perceptual restoration in toddlers. *Perception and Psychophysics, 68*(4), 625-642.

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.

Pollack I. (1975). Auditory informational masking. *Journal of the Acoustical Society of America, 57(S1),* S5-S5. https://doi.org/10.1121/1.1995329.

Riley, K. G., & McGregor, K. K. (2012). Noise Hampers Children's Expressive Word Learning. *Language, Speech, and Hearing Services in Schools*, *43*(3), 325–337. http://doi.org/10.1044/0161-1461(2012/11-0053)

Rönnberg, J., Lunner, T., Zekveld, A., Sorqvist, P., Danielsson, H., Lyxell, B., . . . Rudner, M. (2013). The ease of language understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience, 7,* 31. https://doi.org/10.3389/fnsys.2013.00031

Rothauser, E., & Maiwald, D. (1969). Digitalized sound spectrograph using FFT and multiprint techniques. *The Journal of the Acoustical Society of America*, *45*(1), 308-308. https://doi.org/10.1121/1.1971447

Schneider, B.A., Li, L., Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations. *Journal of the American Academy of Audiology, 18,* 578-591. https://doi.org/10.3766/jaaa.18.7.4.

Storkel, H. L. (2013). A corpus of consonant-vowel-consonant (CVC) real words and nonwords: Comparison of phonotactic probability, neighborhood density, and consonant age-of-acquisition. *Behavior Research Methods*, *45*(4), 1159-1167. http://doi.org/10.3758/s13428-012-0309-7

Storkel, H. L. & Hoover, J. R. (2010). An on-line calculator to compute phonotactic probability and neighborhood density based on child corpora of spoken American English. *Behavior Research Methods, 42,* 497-506.

Swingley, D. (2010). Fast mapping and slow mapping in children's word learning. *Language Learning and Development, 6,* 179-183.

Vlach, H.A., Sandhofer, C.M., & Kornell, N. (2008). The spacing effect in children's memory and category induction. *Cognition, 109 1*, 163-167.

Williams, S. E., & Horst, J. S. (2014). Goodnight book: sleep consolidation improves word learning via storybooks. *Frontiers in Psychology*, *5*, 184. http://doi.org/10.3389/fpsyg.2014.00184

Figure Captions

Figure 1. Presents the average accuracy (percent correct) of target novel words for which a referent was correctly selected during each of the test sessions (A). The *black line* represents novel words that were exposed in babble and the *grey line* represent novel words that were exposed in quiet. Mean ± SE are shown. Part (B) presents the individual referent selection performance. Each novel word which was correctly identified by each of the eight participants on each of the five experimental session in either quiet (circles) or in babble (squares) is marked. The novel words were color coded (blue, green, yellow and orange). For illustrative purposes, the word order represents the order of acquisition rather than exposure.

Figure 2. Presents the average accuracy (percent correct) of target novel words correctly recalled at each of the test sessions for the two exposure conditions (A). Mean ± SE are shown. Part (B) presents the individual recall performance. Each novel word which was correctly recalled by each of the eight participants on each of the five experimental session in either quiet (circles) or in babble (squares) is marked. The novel words were color coded (blue, green, yellow and orange). For illustrative purposes, the word order represents the order of acquisition rather than exposure.

Supplemental Material Descriptions

Appendix A:

Novel words' phonotactic probability and neighborhood density as calculated using an on-line calculator to compute phonotactic probability and neighborhood density based on child corpora of spoken American English (for further details see Storkel & Hoover, 2010).

Appendix B:

The two stories created for the current study. Within each story, each novel word was presented three times.

Appendix C:

The sixteen unidentified objects used in the current study, which were selected either from the Novel Object and Unusual Name (NOUN) Database (Horst & Hout, 2016) or from publicly-available images on the internet.

Appendix D:

Screen shots of the stories taken during the familiarization phase. The picture on the left is from the story "Sleepover" and the picture on the right is form the story "Show and tell" (See Appendix A).

Appendix E:

Screen shots of the referent selection test. A four-alternative forced choice task was created to assess the children's receptive word recognition. Each screen presented four pictures of objects from the same category (screen A- known words; screens B/C-exposed novel words/ non-exposed novel words). Each image was located at one corner of the touch screen, and the child was asked to point at the correct object when prompted with "Where is the _________?"