



Published in final edited form as:

Am J Audiol. 2014 September 1; 23(3): 326–336. doi:10.1044/2014_AJA-14-0020.

Children's Understanding of Instructions Presented in Noise and Reverberation

Dawna E. Lewis, Ph.D.¹, Crystal M. Manninen, Au.D.^{1,2}, Daniel L. Valente, Ph.D.¹, and Nicholas A. Smith, Ph.D.¹

¹Boys Town National Research Hospital, Omaha, Nebraska

²University of Nebraska-Lincoln, Lincoln, Nebraska

Abstract

Purpose—This study examined children's ability to follow audio-visual instructions presented in noise and reverberation.

Method—Children (8–12 years) with normal hearing followed instructions in noise or noise plus reverberation. Performance was compared for a single talker (ST), multiple talkers (MT) and multiple talkers with competing comments (MTC). Working memory was assessed using measures of digit span.

Results—Performance was better for children in noise than for those in noise plus reverberation. In noise, performance for ST was better than for either MT or MTC and performance for MT was better than for MTC. In noise plus reverberation, performance for ST and MT were better than for MTC but there were no differences between ST and MT. Digit span did not account for significant variance in the task.

Conclusions—Overall, children performed better in noise than in noise plus reverberation. However, differing patterns across conditions for the two environments suggested that the addition of reverberation may have affected performance in a way that was not apparent in noise alone. Continued research is needed to examine the differing effects of noise and reverberation on children's speech understanding.

Keywords

Noise; speech perception; children

Introduction

For children, understanding speech is a complex process that is impacted by the acoustic environment. The effects of noise and reverberation on children who are developing speech perception abilities have been documented for a variety of stimuli and conditions (Bradley &

Dawna E. Lewis, Boys Town National Research Hospital, Phone: 402-498-6607, Fax: 402-498-6351, dawna.lewis@boystown.org.

The content of this paper is the responsibility and opinions of the authors and does not represent the views of the National Institutes of Health. The first author serves on the Pediatric Advisory Board of Phonak but there are no conflicts with the work presented here.

No other conflicts of interest are noted by any authors.

Sato, 2008; Elliott, 1979; Fallon, Trehub & Schneider, 2000, 2002; Johnson, 2000). The type of communication interaction (e.g., distance between talkers and between talker(s) and listener(s), number of talkers, whether talkers are visible) as well as linguistic context and complexity can impact understanding (Fallon et al., 2002; Ross, et al., 2011; Ryalls & Pisoni, 1997; Stelmachowicz, Hoover, Lewis, Kortekaas, & Pittman, 2000). Children's ability to understand speech in noise and reverberation also may be influenced by development of memory, attention, and language skills (Hnath-Chisolm, et al., 1998; Mayo, Scobbie, Hewlett, & Waters, 2003; Nittrouer & Boothroyd, 1990; Nittrouer & Miller, 1997; Oh, Wightman, & Lutfi, 2001; Wightman, Callahan, Lutfi, Kistler, & Oh, 2003).

In noise, children require a greater signal-to-noise ratio (SNR) to achieve performance similar to adults on speech recognition tasks (Hall, Grose, Buss, & Dev, 2002; Johnson, 2000; Litovsky, 2005; McCreery et al., 2010; Wightman & Kistler, 2005). Age effects also have been shown for listening in reverberation, with younger children requiring shorter reverberation times than older children and adults for similar speech perception (Neuman & Hochberg, 1983). In conditions that include both noise and reverberation the detrimental effects may be synergistic (Crandell & Smaldino 2000; Finitzo-Heiber & Tillman, 1978; Johnson, 2000; Neuman, Wroblewski, Hajicek, & Rubenstein, 2010; Wroblewski, Lewis, Valente, & Stelmachowicz, 2012; Yang & Bradley, 2009). Because real-world listening most often takes place in environments that contain both noise and reverberation, understanding their combined effect on children's speech perception across a variety of conditions is important.

To date, many of the studies examining children's speech-perception abilities have used relatively simple tasks such as recognition of phonemes, words, or sentences (Finitzo-Hieber & Tillman, 1978; Johnson, 2000; Neuman, et al., 2010; Yang & Bradley, 2009). These tasks provide important information regarding speech recognition; however, they do not address how children perform during more complex listening tasks. As listening tasks become more complex, higher demands are placed on cognitive resources and the negative consequences of noise and reverberation may be intensified (Kjellberg, 2004; Klatt, Lachmann, & Meis, 2010; Shield & Dockrell, 2008). Because their cognitive abilities are still developing, children may need to rely more on acoustic-phonetic (bottom-up) processing for comprehension during such tasks. If the effort required for phonological coding of the speech signal in unfavorable listening environments is increased, fewer resources may be available for other cognitive tasks (Klatt, Hellbruck, Seidell, & Leistner, 2010; Klatt, Lachmann et al., 2010; McCreery & Stelmachowicz, 2013; Picard & Bradley, 2001).

Within everyday environments, multiple cues are available to assist listeners in their ability to selectively attend to speech in the presence of other, irrelevant, sounds. For example, understanding of speech in both adults and children may be improved by both spatial separation of signals (Ihfeldt & Shinn-Cunningham, 2008; Litovsky, 2005) and the availability of audiovisual cues (Massaro & Cohen, 1995; Ross, Saint-Amour, Leavitt, Javitt, & Foxe, 2007; Sumbly & Pollack, 1954; Wightman, Kistler, & Brungart, 2006). However, the benefit of these cues may vary depending on a variety of factors. Spatial separation may be more beneficial in situations requiring attention to one source than in those requiring attention to multiple sources (Shinn-Cunningham & Ihfeldt, 2004). The

addition of visual information has been shown to improve speech perception in noise (Erber, 1969, 1975; Sumbly & Pollack, 1954), as well as under some conditions when the auditory signal is not degraded by noise (e.g., second language for the listener, accented speech by the talker, syntactic and/or semantic complexity of the speech) (Arnold & Hill, 2001). Developmental factors may impact the ability to benefit from audiovisual input, with adults showing greater speech-perception benefit from the addition of visual information than children, and younger children showing less benefit than older children (Massaro, 1984; Massaro, Thompson, Barron, & Laren, 1986; Ross et al., 2011; Wightman et al., 2006).

Classrooms represent one environment in which children must use multiple developing skills in their understanding of speech. Research has consistently shown that the noise and reverberation in classrooms varies and many do not meet national standards (American National Standards Institute [ANSI], 2010; Bradley, 1986; Knecht, Nelson, Whitelaw, & Feth, 2002; Nelson, Smaldino, Erler, & Garstecki, 2008; Rosenberg et al., 1999). In these potentially poor acoustic environments, children complete multifaceted tasks. They must comprehend new information and use short- and long-term memory skills to store that information either for immediate use (e.g., following directions, answering questions) or later retrieval (e.g., taking tests). In complex environments such as these, children must ignore irrelevant sounds and focus on signals of interest (Neuman et al., 2010). This can be especially difficult during activities that involve multiple talkers in which the listener may be forced to either separate the pertinent speech signal from a background of competing signals or to follow a variety of talkers that may change frequently. In such instances, both auditory and visual attention to non-target talkers could serve as distractors (Ricketts & Galster, 2008; Valente, Plevinsky, Franco, Heinrichs-Graham, & Lewis, 2012).

To better understand children's speech understanding in their everyday environments, performance for realistic listening tasks needs to be assessed. The goal of the current study was to understand how different acoustic environments (noise alone versus noise and reverberation) impact the ability of children with normal hearing to complete a complex listening task when both auditory and visual information are available. For this task, children were asked to follow audiovisual instructions presented in three talker conditions; a single talker (ST), multiple talkers speaking one at a time (MT), and multiple talkers with competing comments from other talkers (MTC). Testing was completed in one of two acoustic environments: noise or noise plus reverberation.

Simulated acoustic environments were used in the current study to serve as an alternative to testing in real rooms, which would introduce issues of repeatability, consistency of noise sources, and logistics. The use of recorded talkers provided a consistent level of experimental control over stimulus presentation that would not be available in a real room. Although such a simulation only generally represents listening in a real-world environment, it provides information regarding children's comprehension in plausible acoustics when both auditory and visual information are available to the listener.

We hypothesized that performance within each acoustic environment would decrease as the talker conditions became more complex and that performance in noise plus reverberation would be poorer than in noise alone. Alternatively, it was possible that the availability of

both auditory and visual information would reduce differences in performance between the two environments. The results of this study will add to the body of research describing the effects of acoustic environment on children's comprehension of speech.

Method

Participants

Fifty children with normal hearing, 25 females and 25 males (10 children per year of age from 8–12 years) participated in this study. Half of the children completed the task in each acoustic environment (5 per year of age for a total of 25 children per environment). Participants had their hearing screened at 15 dB HL for octave frequencies from 250 to 8000 Hz. Based on parent report, children were typically developing monolingual speakers of American English with normal or corrected-to-normal vision and no reported physical or cognitive barriers to education. Exclusion criteria included diagnoses of learning disability, attention deficit disorder, color blindness, or children who did not know left from right. The current study was approved by the Institutional Review Board of Boys Town National Research Hospital (BTNRH). Written assent and consent to participate were obtained for each child and they were compensated monetarily for participation.

Prior to the experimental task, each child was assessed using the digit-span subtest of the Wechsler Intelligence Scale for Children-III (WISC-III; Wechsler, 1991) to evaluate verbal working memory. Each child was instructed to repeat a dictated series of digits, spoken live-voice by a female researcher, in serial order (forward condition) and another series of digits in reverse serial order (backward condition). In general, measures of digit span assess the capacity an individual's verbal working memory. Specifically, forward digit span requires short-term storage of the stimulus whereas backward digit span requires both short-term storage and processing (Pisoni, Kronenberger, Roman, & Geers, 2010; St. Clair-Thompson, 2010). The series began with two digits and increased in length by one digit after two trials at each length. The test was discontinued for both conditions when the child could not successfully repeat back two consecutive trials of the same length. Even though we expected working memory skills in a group of typically developing children with normal hearing to be near or within normal limits, the measures of total, forward and backward digit span were conducted to rule out working memory skills as a confounding factor in performance on the experimental task used in this study (Pisoni & Cleary, 2003), particularly for children who might score at the high or low end of performance on the experimental task.

Simulated Environments

To create a plausible audiovisual (AV) instruction task, individual video recordings were made 1) of a single child talker and 2) of four different child talkers (8–12 years old) providing instructions for placing objects on a mat that was situated in front of the child listener who was to perform the task (Figure 1). The children were instructed to speak as they would in normal conversation. To ensure that the total length of an instructional task was the same across conditions, the time between individual instructions within a condition varied. Video recordings were made in a separate sound-treated meeting room at BTNRH. Four Advanced Video Coding High Definition video cameras (JVC Evario hard disk

camera, GZ-MG630) were positioned in the relative location where the child would sit during the experiment, directly facing an individual talker. An omni-directional lavalier microphone (Shure, ULX1-J1 pack, WL-93 microphone) was positioned on each talker and talkers were amplified and recorded with the individual video recordings. This technique provided the greatest direct-to-reverberant energy ratio for the talkers to help minimize the reflections that were present in the recording room (Valente et al., 2012).

Two acoustic environments were created: *noise* and *noise plus reverberation*. For both environments, speech was presented at 65 dB SPL at the listener's location as measured with a Larson Davis 824 sound level meter. Multi-talker noise (20 talkers; Auditec, Inc.) was mixed with the audiovisual recordings and presented at a +5 dB SNR. This noise represents a listening environment in which there would be a large number (20 in this case) of other talkers, who were not the target talkers and whose speech would not be clearly understood, at locations close to the talkers of interest. The SNR was chosen to simulate noise levels plausible in an occupied classroom (Dockrell & Shield, 2006; Finitzo, 1988; Jamieson, Kranjc, Yu, & Hodgetts, 2004).

The noise-only environment was designed to examine the effects of noise on performance in otherwise near-ideal acoustical conditions. The room model was created following the methods described in Wroblewski et al. (2012). Specifically, the noise-plus-reverberation environment was created by convolving room impulse responses (RIRs) with the recordings of speech and multi-talker babble. RIRs were created to reflect a model of an occupied rectangular classroom (9.3 m [W] × 7.7 m [D] × 2.8 m [H]) with reverberation time across octave frequencies based on average classroom measurements by Sato and Bradley (2008). The average reverberation time (RT) was approximately 0.6 sec (octave frequencies from 125 to 4000 Hz). This RT is the maximum recommended for typical, medium-sized classrooms (ANSI, 2010). The ray-tracing module of CATT-Acoustic Software ("CATT-Acoustic™", 2006) was utilized to generate an impulse response (with an omni-directional sound source and receiver) using 250,000 rays and a 1.5 second ray truncation time. The sound source was located 1 meter from the receiver in the room model and both the noise and reverberation used the same RIR. Using the standard linear convolution function in MATLAB®, the entirety of each channel of the speech and noise were both convolved with the resultant impulse response. The talkers' voices were equated for level in the sound booth. The multi-talker babble was convolved with the impulse response in the same way as the recordings of the speech.

Test materials

Materials for the experimental task consisted of 16 objects (see Appendix A) varying on several dimensions (e.g., size, color, shape). Instructions required the child to place 10 of these objects in a specified location/orientation relative to the edges of a rectangular mat or to other items. Sentences that served as instructions followed the same format in all conditions. As an example, the instructions for the first condition are listed in Appendix B. The instructions selected for this task were a modification of verbal instructions used by Bunce (1989) to assess children's referential communication skills. They were chosen for the current study because they were age-appropriate for the study population and provided a

range of complexity. The sentence structure of the instructions was scripted so that the range of complexity across sentences was similar for all conditions. For the recordings of the first two experimental conditions (ST, MT), each child read his/her instructions and all talkers followed that instruction for placing objects on mats before the next instruction was given. The children practiced the reading of those instructions several times so that their speech would be smooth and without error during the recording. For the third condition (MTC) a script was created and all of the talkers memorized their parts during multiple practice sessions prior to the recording (As an example, a short portion of the script for this condition can be found in Appendix C). This condition differed from the other two conditions in that the talkers were making comments in addition to the instructions for the task. Each of the final recordings were completed in a single take with all talkers present and free to look at each other during filming to ensure a natural flow to the interaction. The audio portions of the recordings were equalized for sound pressure level across all talkers.

Procedure

A simulation of an AV instruction-following task was created in a sound booth at BTRNH with internal dimensions of 2.73 m [W] × 2.73 m [D] × 2.0 m [H]. Acoustic measurements for the booth were as follows: background noise level = 37 dBA; T30 (mid) = 0.06 sec; early decay time (EDT) = 0.05 sec.

During the experimental task, the child was seated at a table facing an array of four LCD monitors with loudspeakers (M-Audio Studiophile AV 40) mounted on top (Figures 1 and 2). Screens 1 and 4 had dimensions of 27.5 cm × 43 cm (Dell 2007 WFPb). Screens 2 and 3 had dimensions of 30 cm × 47.5 cm (Samsung Syncmaster 225BW). The arrangement placed the child in the near-field of the loudspeakers and acoustic effects of the booth are assumed to be negligible.

Prior to testing, children were familiarized with the objects to be used during the task and provided with examples of the format of the instructions. The children followed sets of 10 AV instructions for placing the various objects on a mat, in three conditions. They were allowed to look at talkers as much or as little as they felt necessary to complete the task. Based on pilot testing suggesting that some children could be unwilling to continue if the initial task was too difficult, the presentation order of conditions remained the same across all subjects, from least to most complex. The first condition (single talker; ST) consisted of one talker presenting instructions on a video monitor directly across from the child (screen 2). The same talker was used throughout this condition for all environments and participants but was not a talker in the other conditions. This condition was the least complex and served as the baseline for the child's ability to follow instructions in each acoustic environment. In the second condition (multiple talkers; MT), four different talkers were on the four separate monitors across from the child. They took turns providing instructions for the task. All talkers were visible at all times. The order in which the talkers spoke was varied so the child could not anticipate who would provide the next instruction. The third condition (multiple talkers with competing comments; MTC) was similar to the second in that the same four children provided instructions. However, in this condition there were interruptions and distractions from the other talkers during the same time an individual talker was providing

the target instruction. This condition simulated an experience in which listeners are required to follow multiple talkers and separate target speech from distracting speech. Each condition took less than 5 minutes to complete and there was a short break between conditions.

Half of the children performed the task in the noise-only environment and half performed the task in the noise-plus-reverberation environment. To ensure an equal distribution across ages for the two environments, children within each age group (8, 9, 10, 11, 12 years), were alternately assigned to either environment.

Scoring

Forward, backward and total (combination of forward and backward scores) digit-span scores were recorded for each child. The scores were then converted to scaled scores (Wechsler, 1991) to allow comparisons across age.

The three listening tasks (ST, MT, MTC) were scored based on multiple aspects of each of the 10 instructions: specific item, position, step (1–10), and orientation (where applicable). A point was given for each correct aspect and divided by the total number of possible points to achieve an overall percent correct score for each task. For example, if the fifth of the 10 instructions was “Place the small pig in the upper left corner of the mat facing the large chip”, it was worth a total of four points (item=small pig; position=upper left corner; step=5 (i.e., fifth instruction); orientation= facing large chip). If an object was placed on the mat incorrectly, but a subsequent object was correctly placed in relation to it, the latter item’s placement was scored as correct (e.g., if the child had placed a small chip on the mat instead of the large chip but the pig was oriented correctly in relation to that item, the latter would be scored correct for orientation). This gave the child credit for understanding the subsequent instruction, even when a previous one was incorrect. In addition, if an incorrect object was placed on the mat or a correct object was positioned incorrectly, its score would be modified if the participant corrected the error during a later step in the process (e.g., after the modification, the position or item might be correct but the step [1–10] during which the item was placed would now be incorrect). This gave partial credit for the child’s ability to process and correct errors.

For scoring purposes, a video camera was placed above the child to record his/her actions in response to the instructions. A researcher initially scored the tasks as they were administered. That researcher re-checked scores and modified them, as needed, based on a review of the video recording. Reliability of scoring was further assessed by having a second researcher score video-recorded responses from 10% of the participants. Results revealed that the two researchers agreed on 99.6% of the scores.

Results

Figure 3 displays percent correct performance across the two acoustic environments. Separate box plots represent the single-talker (ST), multiple-talkers (MT) and multiple-talkers with competing comments (MTC) conditions. Results reveal different patterns of performance across both condition and acoustic environment. To examine these patterns, results were analyzed using a mixed model analysis of variance (ANOVA) with talker (ST,

MT, MTC) as the within-subject factor and environment (noise, noise plus reverberation) as the between-subject factor. Scores were converted into rationalized arcsine units (RAU; Studebaker, 1985) for statistical analyses to equalize variance across the range of scores. This analysis revealed significant main effects of talker ($F_{(2,96)} = 40.348, p < .001, \eta_p^2 = .457$) and environment ($F_{(1,48)} = 13.890, p = .001, \eta_p^2 = .224$) and an interaction between talker and environment ($F_{(2,96)} = 10.789, p < .001, \eta_p^2 = .184$).

To examine the patterns of significant differences in talker across environment while controlling for Type I error, Fisher's least significant difference was used to calculate a minimum mean significant difference (MMD) of 5.429. Performance was significantly better for children in the noise-only environment than for those in the noise-plus-reverberation environment for both the ST and MTC conditions but not for the MT condition.

In the noise-only environment, performance in the ST condition was significantly better than in either the MT or the MTC conditions and performance in the MT condition was significantly better than in the MTC condition (ST: $M = 108.15, SD = 13.38$; MT: $M = 94.25, SD = 12.19$; MTC: $M = 86.50, SD = 11.39$). It is possible that the high scores in the ST condition represented ceiling or near-ceiling performance for children in this group. In the noise-plus-reverberation environment performance in the ST and MT conditions was significantly better than in the MTC condition. However, performance in the ST condition was not different from that in the MT condition (ST: $M = 88.05, SD = 18.0$; MT: $M = 91.96, SD = 13.21$; MTC: $M = 76.11, SD = 10.50$).

To examine how individual data followed the group results, changes in scores across conditions for individual children in each environment are compared in Figure 4. The columns represent the noise-only (left) and noise-plus-reverberation (right) environments. Rows represent comparisons between ST and MT (top), ST and MTC (middle) and MT and MTC (bottom) conditions. In general, the individual results agree with the averages across environments. For most comparisons, individuals perform better in the less complex of the two conditions being compared. The exception is the comparison between ST and MT in noise plus reverberation. Recall that the ST condition was intended to serve as a baseline for the children's ability to follow instructions in each of the two acoustic environments. In noise-only, the majority of children performed better in this condition than in either the MT or MTC conditions. However, in noise plus reverberation a large number of children performed more poorly in the ST than in the MT condition.

To examine the relationship between performance on the instruction-following game and age, separate regressions were conducted between performance in the ST, MT, and MTC conditions as the dependent variables and age (in months) as the independent variable for the noise-only and noise-plus-reverberation environments (Table 1). Age accounted for a significant part of the variance in performance in the MT and MTC conditions but only approached significance in the ST condition. The latter may be related, at least in part, to the high scores in the ST condition, which may have limited potential variance. Positive slopes indicated improving performance with age. Steiger's Z-test was used to examine whether age was differently correlated for the two significant criterion variables (MT, MTC). Results

revealed no significant difference for any comparison ($p > .05$), indicating that the correlation was equivalent between age and performance in both of these talker conditions.

Figure 5 displays total, forward, and backward digit-span scores. Raw scores are in the left panel and scaled scores are in the right panel. No children were >2 SD below the mean for total, forward or backward scaled scores.

Individual percent correct scores on the instruction task in relation to total raw digit-span scores for each talker condition and environment are shown in Figure 6. To examine the relationship between performance on the AV instruction task and working memory, separate multiple regressions were conducted between performance in the ST, MT, and MTC conditions as the dependent variables and raw digit-span score and age (in months) as independent variables (Table 2). In the ST condition, neither variable accounted for a significant part of the variance in performance. In the MT and MTC conditions, digit span did not account for additional variability in performance beyond age.

Discussion

In the current study, children with normal hearing followed AV instructions presented by one or more talkers in noise and reverberation. It was hypothesized that performance would decrease as complexity of the task increased and when reverberation was added to noise alone. Overall, children who listened in noise performed better than those who listened in noise plus reverberation. However, patterns across talker conditions differed for the two environments. In the noise-only environment, group results followed the predicted pattern. Performance in the baseline ST condition was best, with the majority of subjects performing at or near ceiling. While performance in the MT condition was significantly poorer than in the ST condition, the impact of the multiple talkers on the ability to follow directions was small for a number of children (see Figure 4), possibly as a result of high performance levels in both conditions at the SNR used in the current study. Performance was poorest in the MTC condition, where interruptions and distracting speech could potentially interfere with a listener's ability to follow the target directions. It should be noted that performance levels within conditions as well as relationships across conditions could differ at SNRs other than the one used in the current study.

In noise plus reverberation, children's scores were generally poorer across all conditions compared to those obtained by children in the noise condition. Overall poorer performance suggested that the combination of noise and reverberation interfered with children's ability to perform our experimental task, even when there was both auditory and visual information. The current findings agree with previous research that has shown a synergistic detrimental effect on speech recognition when noise and reverberation are present together in a room (Crandall & Smaldino 2000; Finitzo-Hieber & Tillman 1978; Neuman et al., 2010; Wroblewski et al., 2012; Yang & Bradley, 2009). The pattern of performance across talker conditions in the noise plus reverberation environment was unexpected. Performance in the ST condition was not better than in the MT condition but performance in both the ST and MT conditions was better than in the MTC condition. Our expectation had been that performance with a single talker would be best regardless of acoustic environment. It is

possible that the use of the same presentation order of talker conditions for all subjects may have resulted in an over-estimation of later performance as the result of a) a practice effect for the task itself or b) enhanced performance based on experience listening in the acoustic environment. In the latter, the addition of reverberation could have resulted in a speech signal that was more difficult to understand than in noise alone even in the ST condition. In noise, the ST condition may have been easy enough to represent best performance for most subjects. However, in noise plus reverberation, this may not have been the case for many listeners. In this environment, the mean scores for the ST condition were lower and the range of scores was much greater. Many listeners may have benefited from the experience of listening to reverberated speech in noise during the ST condition, resulting in improved performance in the MT condition (see Figure 4). Potential benefits of previous listening in reverberation are supported by recent work with adults ((Brandewie & Zahorik, 2010; Brandewie & Zahorik, 2013), which showed subsequent improvements in speech perception with even short exposures to the speech signal in a reverberant environment.

An order effect could be expected to result in improved performance across the three conditions as listeners gained experience with the task. In the noise-only environment, it is possible that the degradation in performance that occurred in the MT and MTC conditions would have been even greater without this potential effect. In the noise-plus-reverberation environment, a potential order effect cannot be ruled out, particularly for the comparison of the ST and MT conditions where the majority of subjects demonstrated improved performance in the MT condition. In the ST and MTC as well as the MT and MTC comparisons most subjects demonstrated poorer performance in the more complex of the two conditions. As stated above, that difference may be an underestimation if there is an order effect. Further research will be necessary to examine this issue. As in noise alone, however, the poorest performance in noise and reverberation was in the MTC condition. Poorer overall performance in this condition for both environments suggests an effect of the task complexity. Poorer overall performance in the noise-plus-reverberation environment relative to that in noise alone (e.g., 56% of children had scores \geq 75% compared to only 16% of children in noise) suggests that the combination of both noise and reverberation may have a greater impact on children's ability to perform tasks that require high level cognitive processes. Measuring performance on complex tasks in noise alone may not reflect real-world performance where both noise and reverberation are present.

Working memory, as measured using digit span, did not predict children's ability to follow instructions once accounting for age-related variance. This finding could indicate that the measure of digit span chosen to assess working memory was not sensitive to potential differences for this group of children. The performance in this group who, for the most part, functioned in the normal range on the working memory task may not have differentially impacted their ability to perform the experimental task. Alternatively, it is possible that working memory skills did not play a significant role in performance on the experimental task. The latter possibility is more unlikely, given the nature of the task.

The use of a simulated environment and recorded talkers in the current study presents with some limitations relative to testing under "live" conditions. It is unclear, for example, how well such tasks represent real-world interactions where talkers may vary their speaking level

or rate as the acoustic environment becomes more adverse and how conversational interactions may be altered (McKellin, Shahin, Hodgson, Jamieson, & Pichora-Fuller, 2011). However, simulating the acoustic environment and tasks provided important experimental control that would not be available in a real-world activity. As such, it allowed controlled examination of the effects of acoustic environment on a task requiring both speech recognition and higher level cognitive processing required to follow multiple-talker instructions for which both auditory and visual information are available. The current study used a novel task to assess children's understanding of instructions in noise and reverberation. Only one set of instructions and a single group of talkers was recorded for each condition. It is possible that results might vary if the same task was repeated with different talkers in either the single- or multiple-talker conditions. In addition, the current study did not allow within-subject repeatability measures for the task. Both of these issues should be addressed in future studies.

Understanding speech in everyday life is impacted by many factors. Although much research has examined the effect of noise and reverberation using speech recognition tasks, less is known about effects on tasks requiring higher levels of cognitive processing. The current study represents a step in this direction by examining children's ability to follow directions in two different acoustic environments. The results suggest that, even though noise influences performance as the task becomes more complex, greater effects are seen when noise and reverberation are combined. Current and future work in our laboratory will continue to examine children's speech understanding across a variety of realistic tasks and acoustic environments. Results of such studies will provide information that can assist in optimizing speech perception and multi-modal learning for children with normal hearing as well as children with hearing loss or those with other special listening needs.

Acknowledgments

The authors thank Roger Harpster for assistance in video recording. They also thank Ryan McCreery, Marc Brennan, and Jody Spalding for review of earlier versions of the manuscript. This work was supported by NIH grants R03 DC009675, T32 DC000013, and P30 DC004662.

References

- American National Standards Institute. Acoustical Performance Criteria, Design Requirements and Guidelines for Schools, Part 1: Permanent Schools. New York: ANSI; 2010. ANSI 12.60-2010/Part 1
- Arnold P, Hill F. Bisensory augmentation: A speech reading advantage when speech is clearly audible and intact. *British Journal of Psychology*. 2001; 92:339–353.
- Bradley JS. Predictors of speech intelligibility in rooms. *Journal of the Acoustical Society of America*. 1986; 80(3):837–845. [PubMed: 3760337]
- Bradley JS, Sato H. The intelligibility of speech in elementary school classrooms. *Journal of the Acoustical Society of America*. 2008; 123(4):2078–2086. [PubMed: 18397015]
- Brandewie E, Zahorik P. Prior listening in rooms improves speech intelligibility. *Journal of the Acoustical Society of America*. 2010; 128:291–299. [PubMed: 20649224]
- Brandewie E, Zahorik P. Time course of perceptual enhancement effect for noise-masked speech in reverberant environments. *Journal of the Acoustical Society of America*. 2013; 134:EL265–EL270. [PubMed: 23927235]

- Bunce BH. Using a barrier game format to improve children's referential communication skills. *Journal of Speech and Hearing Disorders*. 1989; 54:33–43. [PubMed: 2915525]
- Carhart R, Tillman TW, Greetis ES. Perceptual masking in multiple sound backgrounds. *Journal of the Acoustical Society of America*. 1969; 45:694–703. [PubMed: 5776931]
- Crandell CC, Smaldino JJ. Classroom acoustics for children with normal hearing and with hearing impairment. *Language Speech and Hearing Services in Schools*. 2000; 31:362–370.
- Dockrell J, Shield BM. Acoustical barriers in classrooms: The impact of noise on performance in the classroom. *British Educational Research Journal*. 2006; 32(3):509–525.
- Elliott L. 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*. 1979; 66:651–653. [PubMed: 489836]
- Erber NP. Interaction of audition and vision in the recognition of oral speech. *Journal of Speech and Hearing Research*. 1969; 12:423–425. [PubMed: 5808871]
- Erber NP. Auditory-Visual Perception of Speech. *Journal of Speech and Hearing Disorders*. 1975; 40:481–492. [PubMed: 1234963]
- Fallon M, Trehub SE, Schneider BA. Children's perception of speech in multitalker babble. *Journal of the Acoustical Society of America*. 2000; 108(6):3023–3029. [PubMed: 11144594]
- Fallon M, Trehub SE, Schneider BA. Children's use of semantic cues in degraded listening environments. *Journal of the Acoustical Society of America*. 2002; 111(5):2242–2249. [PubMed: 12051444]
- Finitzo-Hieber T, Tillman T. Room acoustics effects on monosyllabic word discrimination ability for normal and hearing-impaired children. *Journal of Speech Language Hearing Research*. 1978; 21:440–458.
- Finitzo, T. Classroom Acoustics. In: Roeser, R.; Downs, M., editors. *Auditory Disorders in School Children*. 2. New York: Theime Medical; 1988. p. 221-233.
- Hall JW, Grose JH, Buss E, Dev M. Spondee recognition in a two-talker masker and a speech-shaped noise maker in adults and children. *Ear & Hearing*. 2002; 23(2):159–165. [PubMed: 11951851]
- Hnath-Chisolm TE, Laipply E, Boothroyd A. Age-related changes on a children's test of sensory-level speech perception capacity. *Journal of Speech Language Hearing Research*. 1998; 41:94–106.
- Ihlefeld A, Shinn-Cunningham B. Disentangling the effects of spatial cues on selection and formation of auditory objects. *Journal of the Acoustical Society of America*. 2008; 124(4):2224–2235. [PubMed: 19062861]
- Jamieson DG, Kranjc G, Yu K, Hodgetts W. Speech intelligibility of young school-aged children in the presence of real-life classroom noise. *Journal of the Acoustical Society of America*. 2004; 115:508–517.
- Johnson CE. Children's phoneme identification in reverberation and noise. *Journal of Speech Language Hearing Research*. 2000; 43:144–157.
- Kjellberg A. Effects of reverberation time on the cognitive load in speech communication: Theoretical considerations. *Noise & Health*. 2004; 7(25):11–21. [PubMed: 15703145]
- Klatte M, Hellbrück J, Seidel J, Leistner P. Effects of classroom acoustics on performance and well-being in elementary school children: A field study. *Environment and Behavior*. 2010; 42:659–692.
- Klatte M, Lachmann T, Meis M. Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting. *Noise & Health*. 2010; 12(49): 270–282. [PubMed: 20871182]
- Knecht HA, Nelson PB, Whitelaw G, Feth L. Background noise levels and reverberation times in unoccupied classrooms: Predictions and measurements. *American Journal of Audiology*. 2002; 11:65–71. [PubMed: 12691216]
- Litovsky RY. Speech intelligibility and spatial release from masking in young children. *Journal of the Acoustical Society of America*. 2005; 117(5):3091–3099. [PubMed: 15957777]
- Massaro DW. Children's perception of visual and auditory speech. *Child Development*. 1984; 55:1777–1788. [PubMed: 6510054]
- Massaro DW, Cohen MM. Modeling the Perception of Bimodal Speech. *International Congress of Phonetic Sciences*. 1995:106–113.

- Massaro DW, Thompson LA, Barron B, Laren E. Developmental changes in visual and auditory contributions to speech perception. *Journal of Experimental Child Psychology*. 1986; 41:93–113. [PubMed: 3950540]
- Mayo C, Scobbie JM, Hewlett N, Waters D. The influence of phonemic awareness development on acoustic cue weighting strategies in children's speech perception. *Journal of Speech Language Hearing Research*. 2003; 46:1184–1196.
- McCreery RW, Ito R, Spratford M, Lewis DE, Hoover B, Stelmachowicz PG. Performance-intensity functions for normal-hearing adults and children using Computer Aided Speech Perception Assessment. *Ear & Hearing*. 2010; 3(1):95–101. [PubMed: 19773658]
- McCreery RW, Stelmachowicz PG. The effects of limited bandwidth and noise on verbal processing time and word recall in normal-hearing children. *Ear & Hearing*. 2013; 34:585–591. [PubMed: 23446226]
- McKellin WH, Shahin K, Hodgson M, Jamieson J, Pichora-Fuller MK. Noisy zones of proximal development: Conversation in noisy classrooms. *Journal of Sociolinguistics*. 2011; 15(1):65–93.
- Nelson EL, Smaldino J, Erler S, Garstecki D. Background noise levels and reverberation times in old and new elementary school classrooms. *Journal of Educational Audiology*. 2008; 14:12–18.
- Neuman A, Hochberg I. Children's speech perception of speech in reverberation. *Journal of the Acoustical Society of America*. 1983; 74:215–219.
- Neuman AC, Wroblewski M, Hajicek J, Rubenstein A. Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear & Hearing*. 2010; 31(3):1–9.
- Nittrouer S, Boothroyd A. Context effects in phoneme and word recognition by young children and older adults. *Journal of the Acoustical Society of America*. 1990; 87:2705–2715. [PubMed: 2373804]
- Nittrouer S, Miller ME. Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America*. 1997; 101:2253–2266. [PubMed: 9104027]
- Oh EL, Wightman F, Lutfi RA. Children's detection of pure-tone signals with random multitone maskers. *Journal of the Acoustical Society of America*. 2001; 109:2888–2895. [PubMed: 11425131]
- Picard M, Bradley JS. Revisiting speech interference in classrooms. *Audiology*. 2001; 40(5):221–244. [PubMed: 11688542]
- Pisoni D, Cleary M. Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear & Hearing*. 2003; 24(18):106S–120S. [PubMed: 12612485]
- Pisoni D, Kronenberger W, Roman A, Geers A. Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation. *Ear & Hearing*. 2010; 32(1):60S–74S. [PubMed: 21832890]
- Ricketts TA, Galster J. Head angle and elevation in classroom environments; implications for amplification. *Journal of Speech Language Hearing Research*. 2008; 51(2):516–525.
- Rosenberg G, Blake-Rahter P, Heavner J, Allen L, Redmond B, Phillips J, Stigers K. Improving classroom acoustics (ICA): A three-year FM sound field classroom amplification study. *Journal of Educational Audiology*. 1999; 7:8–28.
- Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ. Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex*. 2007; 17:1147–1153. [PubMed: 16785256]
- Ross LA, Molholm S, Blanco D, Gomez-Ramirez M, Saint-Amour D, Foxe JJ. The development of multisensory speech perception continues into the late childhood years. *European Journal of Neuroscience*. 2011; 33:2329–2337. [PubMed: 21615556]
- Ryalls B, Pisoni D. The effect of talker variability on word recognition in preschool children. *Developmental Psychology*. 1997; 33:441–452. [PubMed: 9149923]
- Sato H, Bradley JS. Evaluation of acoustical conditions for speech communication in working elementary school classrooms. *Journal of the Acoustical Society of America*. 2008; 123(4):2064–2077. [PubMed: 18397014]

- Shield BM, Dockrell JE. The effects of environmental and classroom noise on the academic attainments of primary school children. *Journal of the Acoustical Society of America*. 2008; 123(1):133–144. [PubMed: 18177145]
- Shinn-Cunningham, BG.; Ihlefeld, A. Selective and divided attention: extracting information from simultaneous sound sources. *Proceedings of the International Conference on Auditory Display*; 6–9 July 2004; 2004.
- St Clair-Thompson HL. Backwards digit recall: A measure of short-term memory or working memory? *European Journal of Cognitive Psychology*. 2010; 22:286–296.
- Stelmachowicz P, Hoover B, Lewis D, Kortekaas R, Pittman A. The relation between stimulus context, speech audibility, and perception for normal-hearing and hearing-impaired children. *Journal of Speech Language Hearing Research*. 2000; 43:902–914.
- Studebaker GA. A “rationalized” arcsine transform. *Journal of Speech Language Hearing Research*. 1985; 28:452–465.
- Sumby WH, Pollack I. Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*. 1954; 26(2):212–215.
- Valente DL, Plevinsky H, Franco J, Heinrichs-Graham E, Lewis D. Experimental investigation of the effects of the acoustical conditions in a simulated classroom on speech recognition and learning in children. *Journal of the Acoustical Society of America*. 2012; 131(1):232–246. [PubMed: 22280587]
- Wechsler, D. *Wechsler Intelligence Scale for Children, Third Edition (WISC-III)*. San Antonio, TX: The Psychological Corporation; 1991.
- Wightman FL, Callahan MR, Lutfi RA, Kistler D, Oh E. Children’s detection of pure-tone signals: Informational masking with contralateral maskers. *Journal of the Acoustical Society of America*. 2003; 113:3297–3305. [PubMed: 12822802]
- Wightman FL, Kistler DJ. Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *Journal of the Acoustical Society of America*. 2005; 118(5):3164–3176. [PubMed: 16334898]
- Wightman F, Kistler D, Brungart D. Informational masking of speech in children: Auditory-visual integration. *Journal of the Acoustical Society of America*. 2006; 119(6):3940–3949. [PubMed: 16838537]
- Wroblewski M, Lewis DE, Valente D, Stelmachowicz PG. Effects of reverberation on speech recognition in stationary and modulated noise by school-aged children and young adults. *Ear & Hearing*. 2012; 33(6):731–744. [PubMed: 22732772]
- Yang W, Bradley JS. Effects of room acoustics on the intelligibility of speech in classrooms for young children. *Journal of the Acoustical Society of America*. 2009; 125(2):922–933. [PubMed: 19206869]

Appendix A

Test Materials

small chip	large chip
small bone	large bone
small heart	large heart
small pig	large pig
blue car	green car
yellow crayon	red crayon
blue airplane	red airplane
white dice	red dice

Appendix B

1. Put the large pig in the center of the mat facing you.
2. Put the red dice in the top right corner of the mat.
3. Put the yellow crayon in the top left corner of the mat pointing toward the red dice.
4. Put the green car in the bottom right corner of the mat facing the large pig.
5. Put the red airplane on the mat halfway between the large pig and the green car. Have the front of the airplane facing the left side of the mat.
6. Put the small bone on the top edge of the mat about 1 in. away from the yellow crayon.
7. Put the large chip in the bottom left corner.
8. Put the blue airplane halfway between the large pig and the yellow crayon. Have it point to you.
9. Put the small heart on the bottom edge of the mat about 2 in. from the green car. Have the bottom of the heart point to the top of the mat.
10. Put the large bone halfway between the large pig and the red dice.

Appendix C

(subset of script for the MTC condition)

Talker B: Put the small pig in the top left corner of the mat.

Talker D: I don't have any wigs.

Talker B: I didn't say wig. Pay attention.

Talker D: I'm paying attention but you need to speak (emphasize)...more clearly.

Talker C: Put the large bone in the top right corner of the mat with the front facing the small pig.

Talker B: Hey, if the pig ate the bone it would be a big pig.

Talker A: That rhymes! (while Talker D* starts) A big pig with a red wig.

Talker B to Talker A: (Talking to A while Talker C** starts) Maybe it would be a bigger pig than the big pig we already have.

Talker D*: Excuse me! Put the large chip in the bottom left corner of the mat.

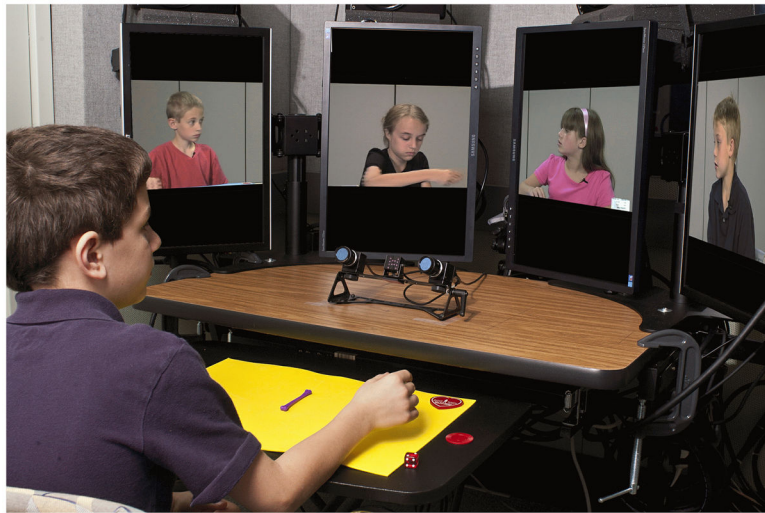


Figure 1.
Set-up for the audiovisual experimental task.

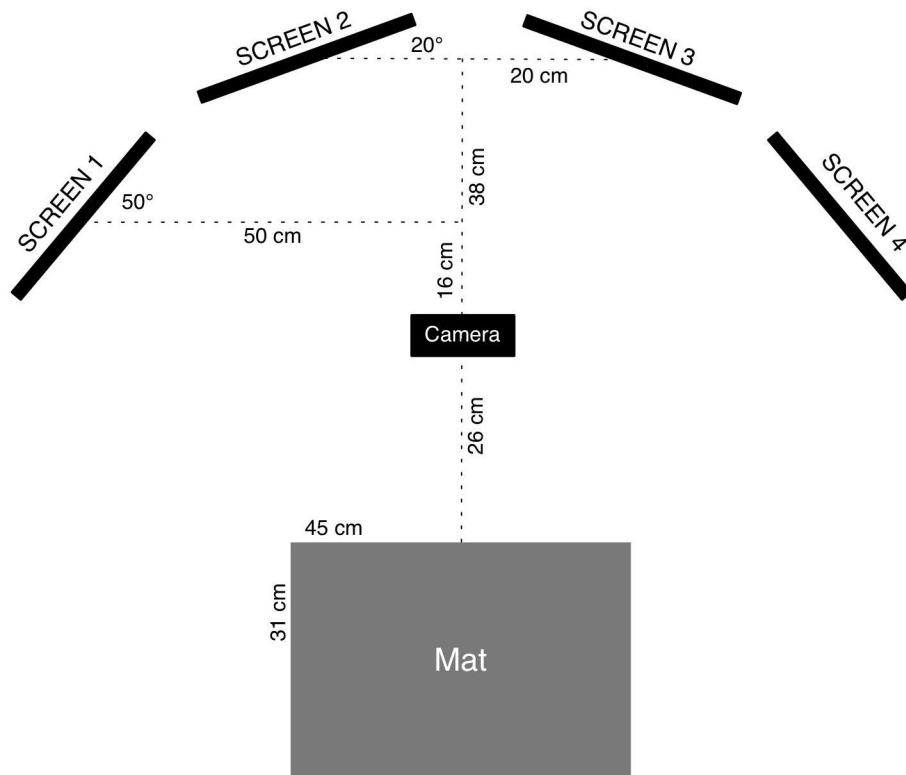


Figure 2.
Spatial arrangement for the experiment.

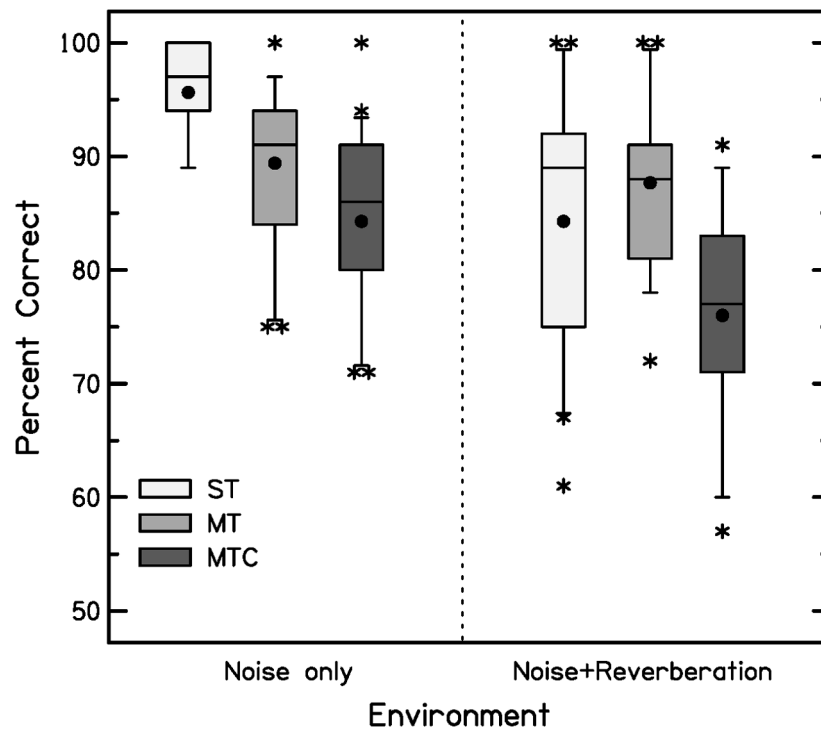


Figure 3. Percent-correct scores for the three talker conditions (single-talker [ST]; multiple talkers [MT], multiple talkers with comments [MTC]) in the noise-only (left) and noise-plus-reverberation (right) environments. Boxes represent the interquartile range and whiskers represent the 5th and 95th percentiles. For each box, the mid-line represents the median and the filled circle represents the mean score. Asterisks represent values that fell outside the 5th or 95th percentiles.

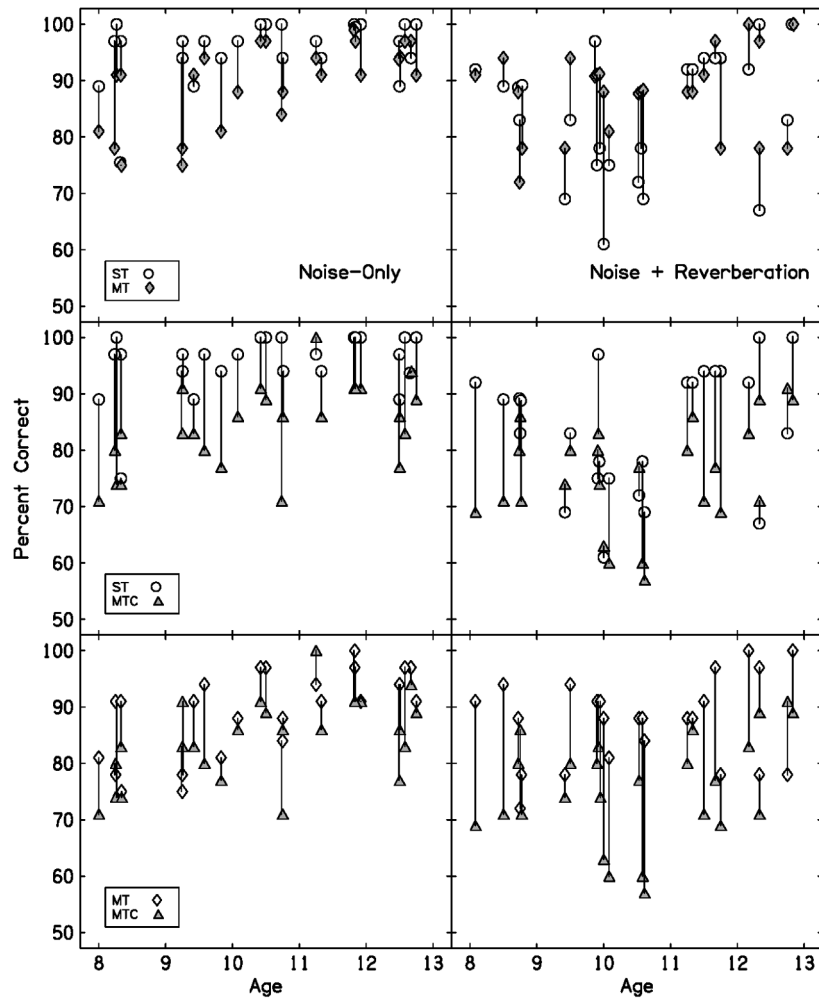


Figure 4. Changes in individual scores when comparing performance across environments (columns) and conditions (rows). Symbols represent the single-talker (ST; open circles), multiple talkers (MT; filled diamonds) and multiple talkers with comments (MTC; filled triangles) conditions. Overlapping symbols are jittered for visibility.

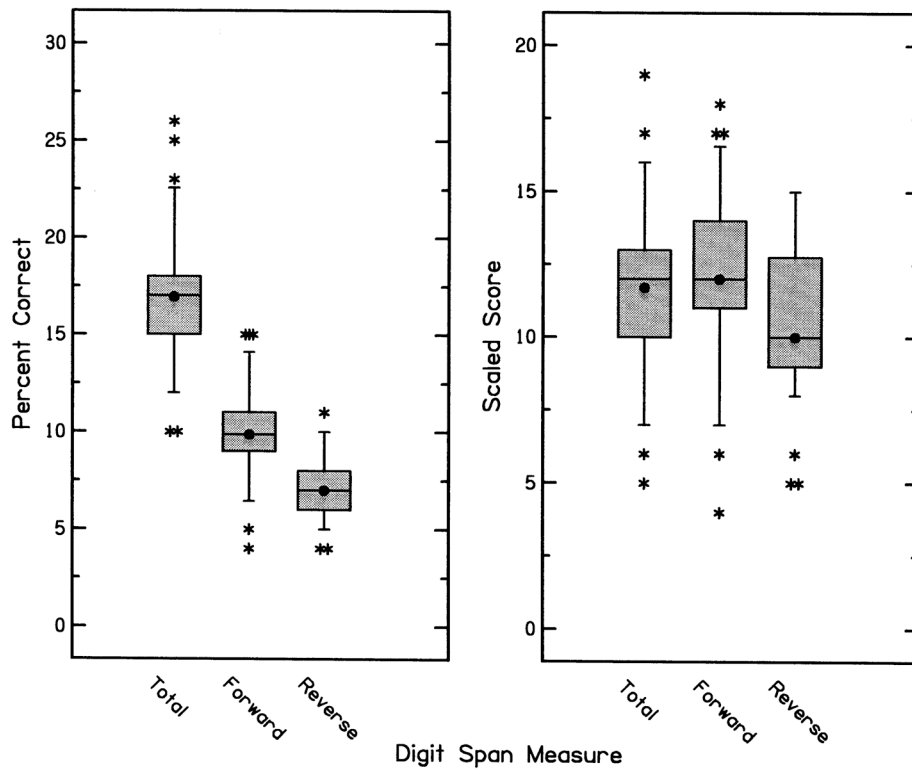


Figure 5. Total (DST), forward (DSF), and backward (DSB) digit-span scores. Raw scores are in the left panel and scaled scores are in the right panel. Boxes represent the interquartile range and whiskers represent the 5th and 95th percentiles. For each box, lines represent the median and filled circles represent the mean scores. Asterisks represent values that fell outside the 5th or 95th percentiles.

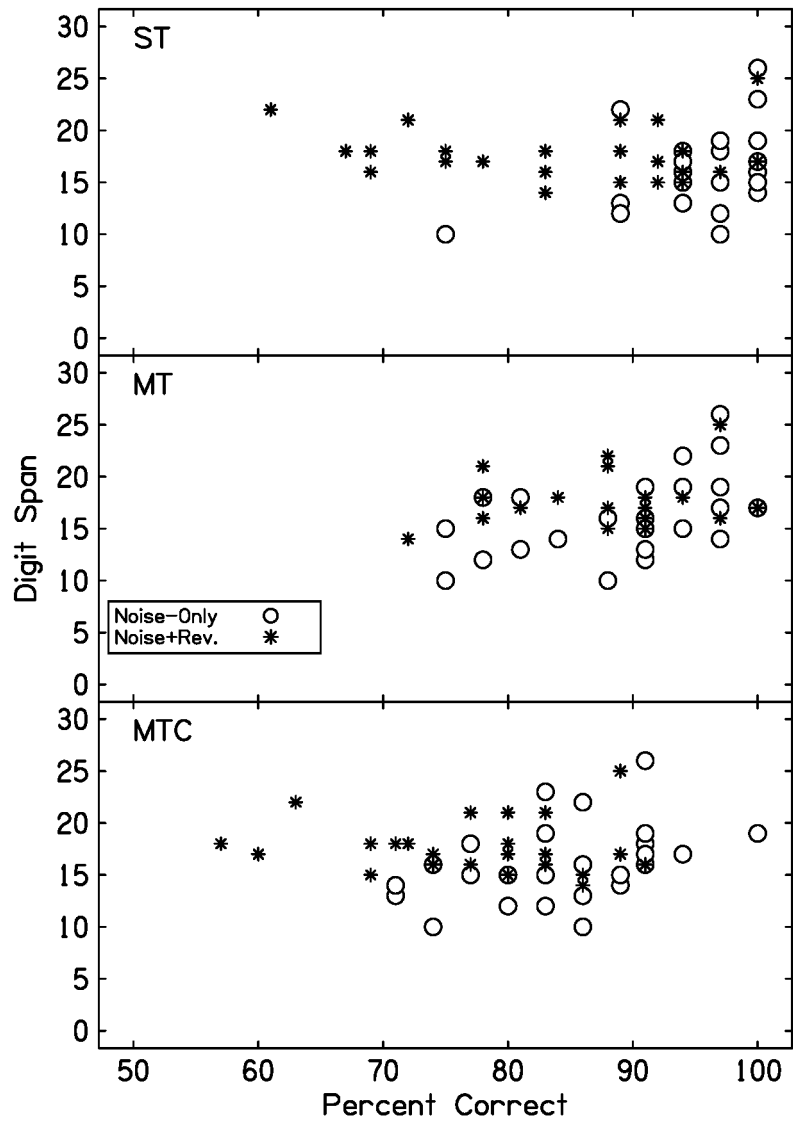


Figure 6. Percent-correct score on the instruction task in relation to total raw digit-span score. Open circles represent the noise-only environment and asterisks represent the noise plus reverberation environment.

Table 1

Regression results examining the relationship between performance on the listening-game task and age.

Condition		Beta	<i>p</i>	R ²
ST	F(1,48) = 3.767	0.270	0.058	0.073
MT	F(1,48) = 13.835	0.473	0.001	0.224
MTC	F(1,48) = 8.085	0.380	0.007	0.144

Table 2

Multiple regression results examining the relationship between performance on the listening-game task and immediate memory.

Condition		Beta	<i>p</i>	R ²
ST	F(2,47) = 1.941; <i>p</i> = 0.115			0.076
	Raw Digit Span Score	-0.063	0.674	
MT		0.292	0.057	
	F(2,47) = 7.372; <i>p</i> = 0.002			0.239
MTC	Raw Digit Span Score	0.131	0.340	
	F(2,47) = 4.084; <i>p</i> = 0.023	0.427	0.003	0.148
	Raw Digit Span Score	-0.067	0.645	
	Age	0.403	0.007	