The effects of noise and reverberation on listening effort for adults with normal hearing

Erin M. Picou[a], **Julia Gordon**[b], and **Todd A. Ricketts**[a]

[a]Vanderbilt University Medical Center

[b]University of Pittsburgh

Abstract

Objectives—The purpose of this study was to investigate the effects of background noise and reverberation on listening effort. Four specific research questions were addressed related to listening effort. These questions were: A) with comparable word recognition performance across levels of reverberation, what are the effects of noise and reverberation on listening effort? (B) what is the effect of background noise when reverberation time is constant? (C) what is the effect of increasing reverberation from low to moderate when signal-to-noise ratio is constant? (D) what is the effect of increasing reverberation from moderate to high when signal-to-noise ratio is constant?

Design—Eighteen young adults (mean age 24.8 years) with normal hearing participated. A dual-task paradigm was used to simultaneously assess word recognition and listening effort. The primary task was monosyllable word recognition and the secondary task was word categorization (press a button if the word heard was judged to be a noun). Participants were tested in quiet and in background noise in three levels of reverberation (T30 < 100 ms, T30 = 475 ms, and T30 = 834 ms). Signal-to-noise ratios used were chosen individually for each participant and varied by reverberation to address the specific research questions.

Results—As expected, word recognition performance was negatively affected by both background noise and by increases in reverberation. Furthermore, analysis of mean response times revealed that background noise increased listening effort, regardless of degree of reverberation. Conversely, reverberation did not affect listening effort, regardless of whether word recognition performance was comparable or signal-to-noise ratio was constant.

Conclusions—The finding that reverberation did not affect listening effort, even when word recognition performance was degraded, is inconsistent with current models of listening effort. The reasons for this surprising finding are unclear and warrant further investigation. However, the results of this study are limited in generalizability to young listeners with normal hearing and to the signal-to-noise ratios, loudspeaker to listener distance, and reverberation times evaluated. Other populations, like children, older listeners, and listeners with hearing loss have been

Address correspondence to: Erin M. Picou, 1215 21st Ave South, Room 8310, Nashville, TN 37232, 615-936-5258 (w), 615-936-6914 (f), erin.picou@vanderbilt.edu.

previously shown to be more sensitive to reverberation. Therefore, the effects of reverberation for these vulnerable populations also warrant further investigation.

Introduction

"Listening effort," which is often defined as the cognitive resources necessary for speech recognition (e.g., Fraser et al. 2010; Hicks & Tharpe 2002; Picou & Ricketts 2014), naturally varies with the difficulty of a listening situation. In listening situations that are more difficult, listeners must use additional cognitive resources and listening effort increases. While data are limited, the consequences of sustained increases in listening effort may include communicative disengagement (Hétu et al. 1993; Hétu et al. 1988), mental fatigue (Hornsby 2013), reduced academic/vocational involvement (Kramer et al. 2006; Nachtegaal et al. 2009), and decreased well-being (Hua et al. 2013).

Because "cognitive resources" are difficult to directly measure, investigators have used several indirect measurement techniques to assess listening effort. For example, some investigators have used physiological indices to infer a listener's exerted effort. These methodologies rely on a body's natural changes that occur with changes in effort, like increased pupil dilation (e.g., Koelewijn et al. 2014; Zekveld et al 2010) and perspiration (Mackersie & Cones 2011). Other investigators have used memory paradigms (e.g., McCoy et al. 2005; Rabbitt 1991). These paradigms rely on the assumption that human cognitive capacity is fixed (Kahneman 1973); as more cognitive resources are deployed to assist with speech recognition, there are fewer cognitive resources available for the processes involved with rehearsal and recall of heard information. Another general type of measurement tool, response-time based measures, also rely on the assumption of finite cognitive resources. In this case, listening effort is inferred when response times slow. Investigators have measured the time it takes a listener to respond to the speech (e.g., Gatehouse & Gordon 1990) or response time to a secondary task during a dual-task paradigm (e.g., Fraser et al. 2010; Picou & Ricketts 2014). Although the preceding list is not all encompassing, it does highlight the diversity of measurement techniques. While there is no consensus on the "best" technique for measuring listening effort, each of the aforementioned techniques is presumed to be a valid indicator of listening effort.

Conversely, subjective indices of listening effort are less consistently accepted as indirect measures of cognitive resources. Subjective measures which can include standardized questionnaires (e.g., Gatehouse & Noble 2004) or rating scales of effort (e.g., Fraser et al. 2010; Picou et al. 2011) are intuitive and often have good face validity. However, the relationship between subjective and objective indices of listening effort is unclear and the two measures can sometimes yield different results (cf. Fraser et al. 2010). Therefore, subjective and objective methodologies are both valuable measurement techniques, but the results from one type of study may not generalize to the other, and the two may in fact reflect two distinct constructs. Since there is some disagreement as to what constitutes "listening effort," changes in objective listening effort in the current manuscript are operationally defined as changes in response times for a secondary task during a dual-task paradigm.

The Ease of Language Understanding (ELU) model (Rönnberg et al. 2013, 2008) provides a theoretical framework for understanding listening effort and for conceptualizing changes in effort. Briefly, the ELU model suggests that a listener compares language input to long term memory stores. If there is a match between the language input and a memory store, then the speech is easily understood. If there is a mismatch between the language input and a listener's memory stores, as might be the case if the signal was degraded or unfamiliar, cognitive resources must be exerted to achieve recognition.

Consistent with this model, previous investigations have found that listening effort increases with increases in background noise (Murph et al. 2000; Picou & Ricketts 2014; Picou et al. 2011; Rabbitt 1968) and at low sensation levels (Gatehouse & Gordon 1990). Conversely, listening effort can be reduced by factors that improve the signal representation and facilitate a match between the language input and long-term memory stores. For example, listening effort can be improved by visual cues (Fraser et al. 2010), hearing aid use (Downs 1982; Picou et al. 2013) and some noise suppression schemes (Desjardins & Doherty 2014; Ng et al. 2013; Sarampalis et al. 2009). Importantly, while changes in listening effort often mirror changes in speech recognition, changes in listening effort may also occur independently of changes in speech recognition (McCoy et al. 2005; Picou et al. 2013; Sarampalis et al. 2009; Surprenant 1999). Because the negative consequences of listening effort may be substantial, and because listening effort may be a distinct construct from speech recognition, it is important to understand how conditions that vary in natural situations affect listening effort.

One factor that has not yet been investigated previously is the effect of reverberation on objective listening effort. While the potential effects on listening effort are not yet known, the consequences of reverberation on speech recognition are well documented. Reverberation impairs speech recognition, particularly at moderate signal-to-noise ratios (SNRs) where performance is neither near floor or near ceiling (Culling et al. 2003; Culling et al. 1994; Duquesnoy & Plomp 1980; Finitzo-Hieber & Tillman 1978; Harris & Swenson 1990; Helfer & Wilber 1990; Náb lek & Pickett 1974b; Neuman et al. 2010; Plomp 1976; Wróblewski et al. 2012). This performance degradation in reverberation is a result of a combination of reverberation effects including reduced ability to utilize interaural timing cues (Darwin & Hukin 2000), reduced unmasking of speech with modulated noise (Wróblewski et al. 2012), reduced ability to exploit differences in fundamental frequency to perceptually segregate talkers (Culling et al. 2003), and reduced binaural advantage of spatially separated speech and noise (Culling et al. 2003; Culling et al. 1994; George et al. 2012; MacKeith & Coles 1971; Moncur & Dirks 1967; Náb lek & Pickett 1974a; Plomp 1976). Furthermore, reverberation exacerbates the negative consequences of background noise on speech recognition performance. Specifically, while reverberation and background noise both impair speech recognition, the combined effect of noise and reverberation is greater than the sum of each of the effects individually (Finitzo-Hieber & Tillman 1978; Gordon-Salant & Fitzgibbons 1995; Harris & Swenson 1990; Náb lek & Mason 1981). Although many auditory skills are negatively affected by reverberation, some auditory skills are preserved in the presence of reverberation. These include the ability to use prosodic cues to separate target talkers (Darwin & Hukin 2000) and spatial release from informational masking (Kidd et al. 2005).

It is important to note, however, that the effects of reverberation vary as a function of timing and the number of reflections. In reverberant environments, the signal reaching a listener's ear includes energy direct from the sound source(s) and also reflected sound energy both from the sound source(s) of interest, as well as any competing noise sources (Beranek 1954). The reflected sound is delayed in time relative to the direct sound. In quiet, if the reflected sound arrives at the listener's ear early, the direct and early reflections are integrated by a listener's auditory system, potentially enhancing speech recognition (Haas 1972; Lochner & Burger 1964; Nábělek & Robinette 1978). Late reflections, and all reflections from background noise, are not integrated with the direct signal of interest (Nábělek & Robinette 1978; Soulodre et al. 1989). Instead, these reflections overlap with the direct sound, causing masking and temporal smearing of the original signal and reducing speech recognition performance. The time delay that defines the boundary between early and late reflections is typically regarded as 0.05 seconds after direct signal presentation (Bradley 1986; Bradley et al. 1999). In addition to not being integrated with the direct energy, the reflected energy may vary spectrally from the direct energy because the duration that reflected energy is present in any reverberant environment varies as a function of frequency. Furthermore, reverberation can significantly alter signal envelope. These resultant differences between the reflected and direct energy may also negatively affect speech recognition performance.

Reverberation is typically quantified by the length of time required for the signal level in a room following an impulse sound to decay a certain amount. For example, T30, a common measure of reverberation time, is the time it takes for the energy to decay from 5 to 35 dB below the initial level; this time is then doubled to approximate the traditional measure of reverberation (RT60), which is the time it takes for a signal to decay 60 dB (ISO 2009). Reverberation times in real rooms vary dramatically, based on room size and other features (e.g., number of reflective surfaces, wall angles). For example, reverberation times measured in typical classrooms range from 200 to 1270 ms (Knecht et al. 2002), whereas concert halls have reverberation times in excess of 3000 ms (Winckel 1962). However, the American National Standards Institutes (2002) recommends reverberation times of less than 600 ms for optimal speech understanding and learning.

Like speech recognition, it is expected that the effect of reverberation on listening effort would also be performance degradation. This expectation is consistent with the ELU model; a distorted speech signal would require explicit cognitive processing to be understood. Because reverberation, particularly late reflections, introduces distortion, it would be predicted that higher levels of reverberation, including those that result in decreases in speech recognition, would increase listening effort. Further, this increase in listening effort may occur even without a change in speech recognition performance. Specifically, if the SNR is improved to equate speech recognition performance between environments with lower and higher reverberation, listening effort may still increase with higher reverberation because more complex listening situations have been suggested to increase reliance on context-dependent, cognitive processes (Pichora-Fuller et al. 1995).

In partial support of this hypothesis, several investigators have found a relationship between reverberation and perceived listening effort in quiet (Rennies et al. 2014; Sato et al. 2008) and in noise (Rennies et al. 2014; Sato et al. 2012). Indeed, perceived difficulty may be

more sensitive to the effects of reverberation than speech recognition performance; ratings of difficulty change with reverberation even if speech recognition does not (e.g., Morimoto et al. 2004; Sato et al. 2008).

However, these findings were related to subjective listening effort. Only a few previous investigations have reported the effects of reverberation on objectively measured cognitive load. These investigations have revealed poorer memory and reduced comprehension of spoken lectures, in at least some conditions, with increases in simulated reverberation, despite no changes in speech recognition performance (Ljung et al. 2009; Valente et al. 2012). Consistent with the ELU model, these results suggest that increasing reverberation increases cognitive load and thus impairs recall performance. However, to date no previous investigations have explicitly examined the effects of reverberation on objective listening effort by manipulating the reverberation time in real rooms.

Purpose

The purpose of this study was to investigate the effects of reverberation and background noise on listening effort for listeners with normal hearing. Listening effort was evaluated objectively using a dual-task paradigm. Because reverberation has previously been shown to impair word recognition performance, it was of interest to evaluate the effects of reverberation both for SNRs that lead to similar word recognition performance and for constant SNRs for which word recognition performance is expected to vary. The listening environments in this investigation were chosen to reflect those commonly found in face-to-face communication including relatively small listener to talker distances and reverberation levels lower than would be found when listening in large auditoriums. There were four specific research questions of interest, each examining the effect of reverberation, word recognition performance, or SNR, while attempting to equate the other factors. Specifically, the questions of interest were:

A. With comparable word recognition performance across levels of reverberation, what are the effects of noise and reverberation on listening effort? Specifically, it was of interest to evaluate the effect of reverberation in quiet and also in noise when word recognition performance was relatively good (~84%). Three levels of reverberation were chosen, all of which reflect reverberation times listeners might be realistically expected to experience in daily life when listening in small to moderately sized rooms. These levels were low (T30 < 100 ms), moderate (T30 = 475 ms), and high (T30 = 834 ms). The SNRs used for testing were based on pilot data acquired from listeners with normal hearing.

B. Within a constant reverberation level (T30 = 475 ms), what is the effect of noise on listening effort? This level of reverberation was chosen as it is presumed to represent a typical listening situation encountered in daily life. Listeners were tested in quiet and at two SNRs, chosen based on pilot data to yield word recognition performance of approximately 84% and 77%.

C. For a constant SNR, what is the effect on listening effort of increasing reverberation from low to moderate? Based on pilot data, word recognition performance was expected to decrease by approximately seven percentage points

when reverberation increased from low to moderate and the SNR was constant. For this SNR, only performance in only low and moderate reverberation was compared; an additional level of reverberation would further degrade word recognition performance by approximately seven percentage points. Thus, any changes in listening effort across all three levels of reverberation at the same SNR would be confounded by relatively large changes in word recognition performance (~14 percentage points from low to high reverberation).

D. For a constant SNR, what is the effect on listening effort of increasing reverberation from moderate to high? Similar to question (C), it was of interest to keep word recognition performance relatively similar across levels of reverberation (~7 percentage point difference). The SNR for this research question was approximately 2 dB more favorable than for question (C), in order to preserve the same approximate word recognition performance change in both questions.

Materials and Methods

Participants

Participants were eighteen adults (16 females) ranging from 22 to 30 years of age ($M = 24.8$, $\sigma = 2.2$). All participants were screened for normal hearing (< 25 dB HL at audiometric octaves from 250 – 8000 Hz) using a clinical audiometer (GSI 61) and were native English speakers. Participants had no history of chronic middle ear disease or neurologic disorder by self-report. Prior to testing, two lists (10 sentences each) of the Bamford-Kowal-Bench Speech-in-Noise test (BKB-SIN; Etymotic Research 2005) were administered to each participant to assess their speech recognition in noise abilities. During BKB-SIN testing, the speech level starts at 70 dB HL and the SNR changes incrementally after each sentence. BKB-SIN stimuli were presented binaurally through headphones consistent with the test instructions. A participant's verbal response was scored by the experimenter and the level at which a participant could understand 50% of the keywords (SNR-50) was calculated, per test administration instructions. The average BKB-SIN SNR-50 score was −3.3 dB ($\sigma = 0.9$; range = −6 to −2 dB). All testing was conducted with approval from Vanderbilt University's Institutional Review Board. Participants were compensated for their time.

Stimuli

To measure objective listening effort, a "semantic" dual-task paradigm was used as described by Picou and Ricketts (2014). The primary task in the dual task paradigm was monosyllable word recognition. The monosyllables, spoken by a female talker, were approximately 1700 ms in length and were matched to have approximately equal root mean square (RMS) levels. These stimuli have been used previously in other investigations of listening effort (e.g., Picou & Ricketts 2011; Picou et al. 2011, 2013, 2014), but are not commercially available. The monosyllable words were developed as reported by Picou et al. (2011). Of the 600 original words, the 120 most difficult were eliminated based on pilot data from listeners with normal hearing. The remaining 480 words were rank ordered by difficulty and then arranged into eight word lists semi-randomly based on this rank order. As reported by Picou and Ricketts (2014), this resulted in eight 60-word lists of approximately

equal intelligibility. The speech stimuli were presented at 65 dBA measured at the position of a participant's ear with the participant absent.

The secondary task in the dual task paradigm was a response time measure wherein a participant categorized words heard as either nouns or non-nouns and then pressed a red button on a universal serial bus keypad (Targus ATPKOUS) as quickly as possible if the word heard was judged to be a noun. Longer response times were interpreted as increased listening effort. Within a given word list, approximately 70% of the words could be interpreted as nouns.

This specific dual-task paradigm and associated scoring method is slightly non-traditional for a couple of reasons. First, the secondary task is closely related to the primary task; the same stimulus (monosyllable word) is used for both tasks. This is different from a more traditional secondary task, whose stimulus is often in a different sensory modality (e.g., a light flashing on a computer screen; Picou et al 2013). In the current study, although both tasks are related to the same stimulus, the paradigm is still considered a "dual task" because the participant is performing two simultaneous tasks, word recognition and word categorization.

Second, traditionally, only correct responses are accepted and included in data analyses. However, performance on the secondary task is based on the cognitive resources used to categorize whatever word is heard after that word is identified. Since, by design the secondary task is the act of categorization, the actual word heard is not of great importance as long as approximately the same percentage of words are identified as nouns across conditions and the average amount of time that is required for identification is similar across the lists (indicating list equivalency on the secondary task).

All button presses were included in the analyses, regardless of whether or not the word was a noun, for three reasons. First, the monosyllable words were devoid of context because they were presented in isolation. Even concrete nouns in isolation can be ambiguous. For example, "cat," "goal," and "lung" are concrete nouns that can only be interpreted as nouns. Conversely, "ring," "room", and "sled" could also be interpreted as verbs. For each word list, approximately 84% of the nouns could be interpreted not only as a noun, but also as something else (adverb, verb, or adjective). This percentage is approximately equivalent across the eight word lists. Participants were instructed to respond even if the word could be used as a noun. Because of the high ambiguity, which is a natural consequence of words in isolation, it was not of interest to include only accurate responses. Second, participants naturally varied in their linguistic skill and flexibility. It was not of interest to reward or punish listeners who were skilled or unskilled, only to ensure they were thinking about the language presented. Specifically, accuracy scores for participants who do not consider all possible homonyms before responding would appear to be lower than those with better linguistic mastery. Finally, using this exact scoring technique, the dual-task paradigm used in this study has been validated against other dual-task paradigms and was found to be more sensitive to changes in listening effort than other more traditional paradigms (Picou & Ricketts 2014). For these reasons, all responses were included in response time analyses, unless the data indicated the participant was not engaged in the task (e.g., pressing the

button after every word or every other word). No participants indicated suspicious response patterns.

When present, the background noise was a four-talker babble. Each of the four talkers was a female reading passages from the Connected Speech Test (CST; Cox et al. 1987; Cox et al. 1988). The recordings were edited such that all sentences had the same RMS level. Each talker's voice originated from a distinct loudspeaker during testing, no two loudspeakers played the same talker at a given time, and a single talker was not always presented from the same loudspeaker. The level of the background noise varied by condition and by participant as described below. See Picou et al (2011) for additional details about background noise stimuli development.

Conditions

Three levels of reverberation were tested, low (T30 < 100 ms), moderate (T30 = 475 ms), and high (T30 = 834 ms). These values represent the mean T30 for 1/3 octave average values from 250 – 6300 Hz. See Table 2 for the frequency specific T30 values for the moderate and high reverberation conditions. To obtain these values, an 84 dBA pink noise was presented from sound editing software (Adobe Audition CS6) using a multichannel sound card (Layla Echo 3G) to a multichannel amplifier (Crown) and routed to two loudspeakers (Tannoy System 600). Identical signals were presented to the two loudspeakers which were facing away from each other and positioned in the corner of the reverberant room. Two loudspeakers were used to emulate an omnidirectional loudspeaker. A sound level meter (Larsen Davis 824S) was used to calculate reverberation times using the manufacturer's software.

Table 1 lists the conditions evaluated for a given research question. In each of those three levels of reverberation, listening effort in quiet was measured. In addition, listening effort in noise was evaluated. The level of the background noise, and consequently the SNR, varied as a function of reverberation so that listening effort could be evaluated with comparable word recognition performance (research question A), constant reverberation (research question B), and constant SNR (research questions C and D). For each condition with background noise present, the SNR used was individually chosen to be relative to a participant's BKB-SIN SNR-50. The purpose of individually adjusting the SNR was to target a desired speech recognition performance level for each listener. The relative SNR in low reverberation was BKB-SIN SNR-50 score minus 1 dB, which was −4.3 dB on average. Two SNRs were used in moderate reverberation, BKB-SIN SNR-50 score minus 1 and BKB-SIN SNR-50 score plus 1 dB, which were −4.3 and −2.3 dB on average, respectively. Finally, two SNRs were used in high reverberation, BKB-SIN SNR-50 score plus 1 and BKB-SIN SNR-50 score plus 3, which were −2.3 and −0.3 dB on average, respectively. These SNRs were chosen because they reflect relatively high word recognition performance levels, which are typical in daily communication experiences. In addition, it was of interest to induce a similar change in word recognition performance with changes in background noise level as would occur with changes in reverberation (approximately 8 RAU in both cases).

Test Environment

Testing occurred in two rooms, a sound-attenuating booth (4 × 4.3 × 2.7 m) and a reverberant room with solid walls, floors, and ceilings (5.5 × 6.5 × 2.25 m). During testing in the sound-booth (low reverberation condition, T30 < 100ms) a participant was seated in the center of the room. The speech stimuli were played from a computer (Dell; Round Rock, TX) via custom programming (Neurobehavioral Systems Presentation v. 14.0; San Francisco, CA) to an audiometer for level control (Madsen Orbiter 922 v2; Schaumburg, IL). From the audiometer, the speech stimuli were routed to a loudspeaker (Tannoy System 600; Coatbridge, Scotland) placed 1.25 m directly in front of a participant. The noise stimuli were played from a computer (Dell) via sound editing software (Adobe Audition CSS5.5; San Jose, CA) to a multichannel soundcard (Layla Echo; Santa Barbara, CA), and then to a multichannel amplifier (Crown; Elkhart, IN), and finally to four noise loudspeakers (Definitive BP-2X; Owings Mills, MD). The noise loudspeakers were 1.25 m from the participant and placed at 45, 135, 225, and 315 degrees.

The "moderate" and "high" reverberation conditions were tested in the reverberant room. The walls and ceiling of this room were designed to include semi-random angles and were painted with reflective paint to provide a long natural reverberation time (T30 = 2100 ms unoccupied). This room can be modified by covering the hard surfaces with acoustic blankets and floor carpeting to reduce the reverberation time. Floor carpet and ceiling acoustic blankets (Sound Spotter 124 4×4) were present across both conditions. Acoustic blankets (Sound Spotter 1244×8) were hung on the four walls to create the moderate reverberation condition (T30 = 475 ms). The acoustic blankets were removed from the walls to create the high reverberation condition (T30 = 834 ms).

During testing in the reverberant room, the speech stimuli were played from a computer (Dell) via custom programming (Neurobehavioral Systems Presentation v. 12.0), to an attenuator for level control (TDT System 2 PA5; Alachua, FL), and then to a loudspeaker (Tannoy 600A) placed 1.25 m directly in front of a participant. The noise stimuli were played from the same computer via sound editing software (Adobe Audition v 1.5) to a multichannel soundcard (Layla Echo), and then to a multichannel amplifier (Crown), and to four noise loudspeakers (Tannoy System 600). The noise loudspeakers were 3.5 m from the participant and surrounded the listener (45, 135, 225, and 315 degrees).

The System 600 and 600A test loudspeakers have a 90 degree conical distribution pattern (Q = 6). Consequently, calculated critical distance ranged from approximately 1.4 m (high reverberation) to 2.5 m (low reverberation) across the three test environments. That is, the listener was always seated within the critical distance of the speech stimuli loudspeaker. The constant distance of 1.25 m was selected to be representative of a distance for face-to-face communication near the boundary between common distances found for conversation between friends and those greater distances common in social and business situations (Hall 1959).

Procedures

Prior to listening effort testing, informed consent was obtained and hearing screening was completed. Then, each participant was tested using the BKB-SIN to determine the level of the background noise to be used during listening effort testing. Finally, listening effort was evaluated in one of the conditions described above. Prior to data collection, participants practiced the dual task paradigm four times. No feedback was given during the practice sessions. During the first practice, participants performed only the secondary task (press the button if the word was a noun). During the second and third practice conditions, participants performed both the primary and secondary tasks, first in quiet and then in the easiest SNR. Finally, during the last practice, participants performed only the secondary task. This final "practice" was used as a baseline response time for each participant and each room. That is, participants were tested in these preliminary conditions twice, once for each room. In this manner, any potential system timing differences between rooms (sound booth, reverberant room) could be accounted for by subtracting the room-specific baseline response time from the response times measured during listening effort testing. The mean baselines were 1384 ms and 1408 ms for the sound booth and reverberant room, respectively. This difference (24 ms) was not consistent across listeners and was not statistically significant, as measured using a single factor analysis of variance ($F_{1,17} = 1.41$, $p = 0.25$, partial $\eta^2 = 0.08$).

During listening effort testing, monosyllable words were presented from the loudspeaker in front of the listener. A word list consisted of 60 monosyllable words. Participants performed word recognition and also pressed a red button if the word heard was judged to be a noun. Participants were instructed to press the button before repeating the word. No feedback was provided regarding participant accuracy on either the primary or the secondary task. Test order was randomized across participants, but blocked such that participants completed all of the testing within a given reverberation condition before completing the testing within the remaining reverberation conditions. Within a given reverberation condition, test order (quiet, noise) was randomized. For each condition in quiet, participants were tested once (60 words total). For each condition with background noise present, participants were tested twice (120 words total). Two lists were used in noise and only one list in quiet because pilot data suggested that performance in noise was more variable.

Data Analysis

To quantify objective listening effort, the mean response time to indicate that a word was a noun was calculated in each condition for each participant. During calculation, the first response of every test list was ignored. Responses that were ±3 standard deviations from the mean in a condition were also ignored. Each participant's room-specific baseline response time was subtracted from the mean response time for a given condition. For conditions with background noise, the mean of the two repetitions was used in analysis, as each background noise condition was evaluated twice.

To address the four research questions of interest, four separate analyses were conducted. See Table 1 for the conditions used in each analysis. The same analyses were conducted for word recognition and response time data. Prior to analysis, word recognition data were transformed to rationalized arcsine units (RAU) in order to normalize the variance at the

extremes (Studebaker 1985). In addition, the word recognition and response time data were both evaluated for the assumptions underlying analysis of variance, including skewness, kurtosis, and sphericity. All transformed data conformed to the assumptions necessary to perform parametric testing. Finally, the analyses for the response time data were conducted twice, once without adjustment for baseline response time and once with response times that had been adjusted by subtracting room-specific baseline response times. Results revealed an identical pattern of results, so only the corrected response times and analyses are reported.

Results

Word Recognition

The mean word recognition data (RAU) are displayed in Table 3. These results confirmed that the chosen SNRs were generally effective in meeting the goals of the experimental design. Specifically, while SNR and reverberation affected word recognition performance, conditions were also generated for which SNR and reverberation varied, but word recognition performance was comparable across levels of reverberation. For example, word recognition performance in the low reverberation condition (−4.3 dB SNR) was 85.6 RAU ($\sigma = 6.6$) and performance in the moderate reverberation condition (−2.3 dB SNR) was 86.2 RAU ($\sigma = 7.5$). For analysis purposes, the SNRs that led to average speech recognition performance of ~115 RAU (quiet), ~84 RAU (SNR84) or ~77 RAU (SNR77) were grouped together for the specific analyses of the experimental question that had the criteria of comparable word recognition performance (research question A; see Table 1). While not the primary experimental questions, the following analyses of word recognition performance are provided as confirmation of the expected significant effects of SNR and reverberation. Analysis results are displayed in Table 4.

To evaluate the effects of reverberation and background noise when word recognition performance was comparable across levels of reverberation (research question A), a repeated measures analysis of variance (ANOVA) was performed with two within-subject variables, Reverberation (low, moderate, high) and Noise (quiet, SNR84). The SNR77 conditions were not included in this analysis because this performance level was present for only two of the three reverberation levels. Results revealed a significant main effect of Noise ($F_{1, 17} = 631.9$, $p < 0.001$, partial $\eta^2 = 0.97$) and a significant main effect of Reverberation ($F_{2, 16} = 4.7$, $p < 0.05$, partial $\eta^2 = 0.37$). There was no significant Reverberation $\times$ Noise interaction. These results confirmed that the chosen improvement in SNR for the background noise conditions offset decrease in performance resulting from longer reverberation time in moderate (T30 = 475 ms) compared to low (T30 < 100 ms) reverberation; however, in the highest reverberation (T30 = 834 ms), word recognition performance was significantly worse than that in the other two reverberation conditions in quiet and in noise.

One of the goals of this study was to evaluate the effects of reverberation on listening effort when word recognition performance was similar across levels of reverberation in quiet and in noise (research question A). Although the methodology was designed to achieve these goals, the aforementioned results demonstrate that word recognition performance was not actually equivalent across levels of reverberation. Despite the statistically significant effect

of high reverberation on word recognition performance in quiet and in noise (SNR84), the authors moved forward with subsequent analyses because word recognition performance was reasonably comparable across levels of reverberation and effects of reverberation were relatively small. Specifically, the entire range of average performance across the three levels of reverberation was 2.3 RAU in noise and 4.3 RAU in quiet.

To confirm that the three SNRs resulted in different word recognition performance when the degree of reverberation was constant (T30 = 475 ms), a repeated measures ANOVA was conducted with a single within-subject variable, Noise (quiet, SNR84, SNR77). Results revealed a significant main effect of Noise ($F_{2, 16}$ = 152.0, $p < 0.001$, partial η^2 = 0.95). Follow-up testing using multiple pairwise comparisons and controlling for family-wise error rate with Bonferonni adjustments revealed significant performance differences between all three SNRs ($p < 0.01$).

To evaluate the effect of reverberation for a fixed SNR across changes in reverberation, two separate ANOVAs were completed, each with a single with-subject factor, Reverberation (low/moderate or moderate/high). The first ANOVA analyzed the effect increasing reverberation from low to moderate with a fixed SNR (−4.3 dB on average). Results revealed a significant effect of Reverberation ($F_{1, 18}$ = 16.2, $p < 0.001$, partial η^2 = 0.48). The second ANOVA analyzed the effect of increasing reverberation from moderate to high with a fixed SNR (−2.3 dB on average). Results revealed a significant main effect of Reverberation ($F_{1, 18}$ = 22.9, $p < 0.001$, partial η^2 = 0.58). In total, these results suggest that, for a fixed SNR, increasing reverberation decreased performance (~8 RAU for each increase in reverberation at a fixed SNR; see Table 3).

Listening Effort

Table 5 reveals the mean response time data for all conditions tested. Table 6 reveals the results of all statistical analyses on response times. To evaluate the effects of reverberation and background noise when word recognition performance was comparable across levels of reverberation (research question A), a repeated measures ANOVA was performed with two within-subject variables, Reverberation (low, moderate, high) and Noise (quiet, SNR84). Figure 1 displays the mean response time data, minus the room-specific baseline, for each condition. Results revealed a significant main effect of Noise ($F_{1, 17}$ = 47.21, $p < 0.001$, partial η^2 = 0.74). There was no significant main effect of Reverberation and no significant Reverberation × Noise interaction. These results suggest that background noise increased response times (mean 91 ms slower in noise than in quiet). However, degree of reverberation had no effect on listening effort when word recognition performance was comparable (~ 116 RAU in quiet; ~84 RAU in noise) across the three levels of reverberation.

The mean response time data (minus the baseline) for constant degree of reverberation (research question B) are displayed in Figure 2. To analyze the effect of increasing noise on listening effort (within constant reverberation level [475 ms]), a repeated measures ANOVA was completed with a single within-subject factor, Noise (quiet, SNR84, SNR77). Results revealed a significant main effect of Noise ($F_{2, 17}$ = 10.13, $p < 0.01$, partial η^2 = 0.56). Follow-up testing using paired-sample t-tests and Bonferonni adjustments revealed

significant differences between all three conditions ($p < 0.05$). Response times in quiet were 74 ms faster than in –2.3 dB SNR and 117 ms faster than in –4.3 dB SNR. These results suggest that when reverberation was held constant, adding noise or increasing the noise level both had negative effects on listening effort.

The mean response time data for constant SNR are displayed in Figure 3 (research questions C and D). To evaluate the effect of reverberation when the SNR was constant across levels of reverberation, two separate ANOVAs were completed, each with a single within-subject factor: Reverberation. A single ANOVA was not used to analyze these data because scores differed with regards to both the noise level and degree of reverberation they were obtained in. The first ANOVA analyzed the effect increasing reverberation from low to moderate with a fixed SNR (–4.3 dB on average; left panel Figure 3). Results revealed no significant effect of Reverberation, although there was a trend for reverberation to *decrease* listening effort; response times were 15 ms faster in moderate reverberation than in low reverberation. The second ANOVA analyzed the effect of increasing reverberation from moderate to high with a fixed SNR (–2.3 dB on average; right panel Figure 3). Results revealed no significant effect of reverberation, although there was a trend for reverberation to increase listening effort. Response times were 12 ms slower in high reverberation than in moderate reverberation.

When all the listening effort data are considered together, the results of this study suggest that reverberation did not increase listening effort over the limited range of reverberation times and SNRs evaluated, regardless of whether the SNR was constant or if word recognition performance was comparable. However, background noise did increase listening effort, as evidenced by longer response times in noise than in quiet in all three levels of reverberation and also longer reaction times in the SNR77 compared to the SNR84 condition in moderate reverberation (RT = 475 ms).

Relationship between Variables

Although listening effort and word recognition performance are often conceptualized as separate constructs (e.g., Sarampalis et al. 2009), it is likely that the two co-vary. That is, factors that significantly affect word recognition, like background noise, likely also affect listening effort. Therefore, it was of interest to statistically explore the relationship between word recognition performance (RAU) and listening effort (ms). The Pearson correlations between word recognition and response times were conducted separately for each level of reverberation. Holm-Bonferonni adjustments were made to control for family-wise error rate.

Results revealed that in low reverberation, word recognition performance in quiet was not correlated with performance in noise ($p > 0.05$) and word recognition performance was not correlated with response times in either quiet or for the –4.3 dB SNR ($p > 0.05$). However, response times in quiet and noise were significantly correlated with each other ($p < 0.01$, $r(16) = 0.69$). Similarly, in moderate reverberation, word recognition scores were not correlated with each other ($p > 0.05$) or with response times ($p > 0.05$). However, response times were correlated with each other. Specifically, response times in quiet were significantly correlated with those for the –2.3 dB SNR ($p < 0.01$, $r(16) = 0.67$) and with

those for the −4.3 dB SNR ($p < 0.01$, $r(16) = 0.63$); response times for the −2.3 dB SNR were significantly correlated with those for the −4.3 dB SNR ($p < 0.01$, $r(16) = 0.87$). Similarly, in high reverberation, word recognition scores were not correlated with each other ($p > 0.05$) or with response times ($p < 0.05$). However, response times in high reverberation were significantly correlated with each other Specifically, response times in quiet were significantly correlated with those for the −0.3 dB SNR ($p < 0.01$, $r(16) = 0.87$) and with those for the −2.3 dB SNR ($p < 0.01$, $r(16) = 0.77$); response times for the −0.3 dB SNR were significantly correlated with those for the −2.3 dB SNR ($p < 0.01$, $r(16) = 0.86$). In total, these results support the notion that listening effort and word recognition are distinct constructs. That is, response times were only correlated with each other, and not with word recognition performance. Listeners who were slow responders tended to be slow responders regardless of noise or reverberation; factors that affected word recognition for listeners did not affect response times similarly.

In addition to the relationship between word recognition and response times, it was of interest to evaluate the relationship between age and changes in word recognition performance with noise or reverberation, in addition to the relationship between age and changes in response times with noise or reverberation. Correlation analyses revealed no significant statistical relationship between age and susceptibility to noise or reverberation for word recognition or response times. These results suggest that, despite variability in susceptibility to changes, there was no relationship between these changes in effort and age (within the limited range of ages evaluated).

Discussion

The purpose of this project was to evaluate the effects of background noise and reverberation on listening effort for listeners with normal hearing. Specifically, the effect of reverberation was evaluated either when word recognition performance was comparable or when SNR was held constant and speech recognition was significantly decreased by increasing levels of reverberation. In addition, the effect of background noise was evaluated within a single level of reverberation. The results revealed that, consistent with previous investigations, increases in background noise increased listening effort and decreased word recognition performance (e.g., Murphy et al. 2000; Sarampalis et al. 2009). Furthermore, the approximate 100 ms increase in response times associated with the introduction of background noise is reasonably consistent with changes in listening effort previously shown using similar a methodology (Picou & Ricketts 2014). Also consistent with previous investigations, increasing reverberation decreased word recognition performance if the SNR was held constant across levels of reverberation (see Table 3; e.g., Finitzo-Hieber & Tillman 1978; Harris & Swenson 1990; Helfer & Wilber 1990; Náb lek & Pickett 1974a,b).

Somewhat surprisingly, there was no effect of reverberation on listening effort when word recognition was comparable across levels of reverberation. One of the study aims was to evaluate the effects of reverberation on listening effort when word recognition performance was comparable across levels of reverberation (research question A). Although the methodology was designed to achieve this aim, there was a statistically significant, negative effect of high reverberation on word recognition performance in quiet (~4 RAU) and in

noise (~2 RAU). Despite these small, negative effects, there were no observed changes in response times as a function of changes in reverberation. These data suggest that reverberation did not affect listening effort when word recognition was comparable across levels of reverberation.

Another study aim was to examine the effects of reverberation on listening effort when the SNR was constant across levels of reverberation (research questions C and D). Surprisingly, effort did not increase with increasing reverberation, even for conditions in which word recognition was significantly degraded. Specifically, when the SNR was held constant and reverberation increased, word recognition performance decreased 8 RAU. However, mean response times did not change, suggesting that increasing reverberation did not increase listening effort even when word recognition was degraded.

The non-significant effects of reverberation are in contrast to the expected findings and inconsistent with the ELU model (Rönnberg et al. 2008, 2013). Within the context of the ELU model, signal distortions would be expected to increase listening effort. Furthermore, previous investigations have suggested that cognitive performance can be disrupted by reverberation, even when speech recognition is not affected (Ljung & Kjellberg 2009; Valente et al. 2012).

One possible explanation for the discrepant findings between the current and previous investigations may be word recognition performance levels. Specifically, Ljung and Kjellberg (2009) and Valente et al. (2012) used favorable SNRs (+15 dB and +7/+10 dB, respectively) and speech recognition performance was nearly perfect (~95%). Conversely, in the current investigation, the participants were tested in either quiet or in more difficult listening situations (−2.3 and −4.3 dB SNR on average). Perhaps the effects of reverberation on cognition and listening effort are only evident when noise is present, but performance is still near perfect. Listeners were correctly repeating words ~84% and ~77% of the time in the current study; perhaps this difficulty level overwhelmed the effects of reverberation on listening effort. Conversely, previous investigations suggest that response time paradigms may be most sensitive to changes in effort when speech recognition performance is between 50 and 85% (Gatehouse & Gordon 1990). The effects of reverberation on listening effort with different SNRs and word recognition performance levels warrants further investigation.

Another possible explanation for the lack of significant findings could be that the baseline response times were only room specific; they were not reverberation-time specific. In the context of the ELU model, a change in reverberation time would be expected to slow response times, even in the baseline condition because the baseline condition also requires language processing. Baselines were only collected in the moderate reverberation condition and the effect of more specific baselines is unclear. Conversely, it is not clear that changes in baseline conditions would affect the study conclusions. First, the changes in response times with increases in reverberation were small (e.g., 12 ms from moderate to high) and in some cases were in the opposite of the expected direction (e.g., response times were 27 ms faster in moderate than in low reverberation in quiet). If baseline reaction times were indeed slower in high reverberation, subtracting them would have made the effect of reverberation even more positive (faster reaction times in higher reverberation). In addition, the pattern of

results was the same regardless of whether or not the baseline response times were subtracted prior to analysis.

It might also be argued that the methodology used was not sensitive enough to accurately detect changes in listening effort resulting from changes in reverberation. For example, the secondary task was to judge quickly whether the word presented was a noun. Although the assumption is that judgements would be slowed under cases of increased cognitive load, other factors might also affect speeded judgements. For example, lexical judgements have previously been found to quicken with word frequency and familiarity (e.g., Connie et al 1990) and to slow with ambiguity (e.g., Klepousniotou 2002). The word lists used in the current study were designed to be approximately equivalent in terms of word recognition performance, but are not inherently equivalent also on word frequency, familiarity, and ambiguity. In addition, the introduction of noise changes the list equivalency on these dimensions in an unpredictable manner based on what a listener perceives. It is possible that the lack of list equivalency renders the dual-task paradigm employed insensitive to changes in listening effort.

However, the list order and test condition were counterbalanced across participants to eliminate effects of any inter-list differences on average listening effort performance. To further examine list equivalency as a potential variable, an ANOVA was completed to examine response times across the eight lists when collapsed across reverberation and noise levels. The findings of this analysis were not significant, ($F(7, 11) = 0.826$, $p > 0.5$, $\eta2 = 0.34$, observed power $= 0.22$) suggesting the lists were indeed equivalent with regards to listening effort. Furthermore, the paradigm used in this study has been shown to be more sensitive than other commonly used dual-task paradigms for adult listeners. Specifically, Picou and Ricketts (2014) compared the paradigm used in the present study to two established dual-task paradigms. They found the task used in this study more sensitive to the effects of background noise than the other tasks, which had been employed previously by Downs (1982) and Sarampalis et al (2009). Although it is possible that the paradigm is accurately sensitive to changes in effort with background noise, but not to changes in effort with reverberation, it would be surprising if this paradigm was differentially sensitive to changes in listening effort. There is no evidence this task would be sensitive to changes in effort with background noise, but insensitive to changes in effort in reverberation within the sample study sample.

Furthermore, the methods in the current study were not insensitive to all changes in listening effort. Instead, response times did demonstrate the expected slowing with changes in background noise level. For example, the change in SNR that resulted in word recognition changes of 8 RAU, also resulted in average increased response time of approximately 44 ms (see Figure 2). Conversely, when word recognition performance changed 8 RAU due to a change in reverberation time, response times did not change consistently or statistically significantly. Specifically, when word recognition performance was negatively impacted by ~8 RAU, response times slowed by only 12 ms (moderate to high reverberation) or decreased by 15 ms (low to moderate reverberation; see Figure 3 and Table 5). Taken together, these results suggest that the paradigm is sensitive to changes in effort due to changes in SNR, but is not equally sensitive to changes in reverberation, despite a similar

change in word recognition performance. In other words, there was a cognitive cost to decreasing word recognition by increasing the noise level, while there was no measured change in listening effort when reverberation increased and word recognition performance decreased.

However, it is not clear how reverberation would affect listening effort using other types of methodology. Clearly other investigators have demonstrated that reverberation has increased perceived listening effort, as indicated by subjective reports (e.g., Reinnes et al. 2014; Sato et al. 2008; 2012). In addition, two other studies have suggested that reverberation affects memory and learning comprehension, in at least some conditions (Ljung et al. 2009; Valente et al. 2012). Valente et al (2012) demonstrated that reverberation affected cognition only for listening situations with multiple sequential talkers whose locations varied. Conversely, in the condition which most closely resembles the current study (talker front), Valente et al. (2012) did not find an effect of reverberation on cognition. These data provide additional support that increased reverberation may not affect cognition in simple listening environments.

Limitations

While the data support that the reverberation times investigated did not affect the magnitude of listening effort in listeners with normal hearing, there are several potential limitations that reduce the generalizability of these results. First, the moderate and high reverberation conditions were conducted in an artificially small environment relative to the reverberation times. The reverberant test room for these conditions was 5.5m × 6.5m × 2.25m, with a room volume of 80m^3. In realistic environments, this size environment is usually associated with very low reverberation times; much larger enclosures are associated with the higher reverberation times (e.g., Knecht et al. 2002). Consequently, in more typical enclosures, later reflections reach the listeners ears after contacting far fewer surfaces. While the relative effect this may have on listening effort is unknown, the results of this study cannot be generalized to larger enclosures until further work is completed.

Another possible limitation relates to loudspeaker location. Specifically, the noise loudspeakers were located closer to the participant in the low reverberation condition (1.25 m) than in the moderate and high reverberation conditions (3.5 m). The distance at which the level of the direct signal is equal to that of the reverberant signal is referred to as the critical distance (Egan 1988; Peutz 1971). While the noise loudspeakers were placed outside the critical distance for all conditions as calculated by the formula proposed by Peutz (1971), the noise loudspeaker distance relative to critical distance differed for each environment. In addition, the participants were always within the critical distance of the speech loudspeaker across all conditions. Based on these results, it is unknown if listening effort would interact with reverberation when listeners are beyond the critical distance. The specific interaction between critical distance, reverberation, and listening effort may be of interest for future studies.

A third possible limitation is related to the range of reverberation times. The range evaluated in this study was relatively small (T30 < 100 ms to T30 = 834 ms). Previous investigations that have evaluated the effects of reverberation on cognitive performance have used larger

variations in reverberation time. For example, Ljun and Kjellberg (2009) used reverberation times of 300 to 1840 ms and Valente et al. (2012) used reverberation times of 600 to 1500 ms. It is possible that longer reverberation times would have had a larger effect on listening effort in the present investigation. Conversely, the range of reverberation times was sufficient to affect word recognition performance (~ 8 RAU). That is, within the context of the ELU model, factors that introduce distortion would be expected to increase listening effort. The degree of reverberation introduced in the current study caused enough distortion to negatively impact word recognition and thus one might also expect a concomitant change in listening effort.

A fourth potential limitation relates to the participants. The participants were primarily female (16 of 18 participants). Although the effects of reverberation and background noise are unlikely to vary by gender, the generalizability of the current findings is limited primarily to female listeners. Also related to participants, the sample size was relatively small. Although the effect sizes of the significant findings were large (partial η^2 from 0.56 to 0.74; see Table 6) and the effect sizes of the non-significant findings were very small (partial η^2 from 0.01 to 0.19; see Table 6), one cannot exclude the possibility that a larger participant pool would have revealed a different pattern of results.

A fifth potential study limitation is that one of the study aims was not completely achieved. That is, one of the goals of the study was to evaluate the effects of reverberation on listening effort when word recognition performance was comparable across the three levels of reverberation. To achieve this aim, the SNR was improved with each increase in reverberation time to compensate for the expected decrease in word recognition performance in noise. Despite this manipulation of the SNR, word recognition was not statistically equivalent across all three levels of reverberation. Specifically, it was approximately 2 RAU worse in high reverberation compared to low and moderate reverberation. Therefore, a definitive statement regarding the effects of reverberation on listening effort with constant word recognition performance cannot be made based on the results of this study.

Finally, it is possible that the unbalanced design may have contributed to the non-significant effects of reverberation on response times. Specifically, there was no single SNR that was tested in all three levels of reverberation. Therefore, it is possible that there would have been an effect of reverberation comparing low and high reverberation at a single SNR. While this may be the case, there are several factors that suggest even this comparison may not have revealed an effect of reverberation on listening effort. First, there was no effect of reverberation (low compared to moderate or high) in quiet. Indeed, responses times were 41 ms faster on average in the high reverberation condition compared to the low reverberation condition (this difference was not statistically significant). Second, there was no effect within a single SNR of increasing reverberation. Response times increased only 12 ms at the −2.3 dB SNR when reverberation increased from moderate to high. Responses times actually decreased 15 ms at the −4.3 dB SNR when reverberation increased from low to moderate. These effects were small and inconsistent across participants, especially considering the large standard deviations. Therefore, it is not clear that testing the three levels of reverberation within a single SNR would change the study conclusions. However,

future investigations should consider using a broader range of reverberation times within a single SNR to fully explore the effects of reverberation on listening effort.

The reason that increasing reverberation did not lead to increased listening effort in listeners with normal hearing, even when word recognition was degraded, remains unclear and is inconsistent with the ELU model. It seems apparent, however, that how reverberation affects speech recognition (e.g. Bradley et al. 2003; Houtgast & Steeneken 1985) may not apply to listening effort. These results also suggest that, within the construct of the ELU model, the specific type of distortion of the speech signal may affect the efficiency with which it is matched to representations within the long term memory stores. If the results of the current study and the investigation by Valente et al (2012) are confirmed in future investigations, taking into consideration the aforementioned limitations, the ELU model may need to be modified to reflect the differential effect of noise and reverberation on cognitive load. The observed pattern of results suggests that some distortions of the signal that reduce speech recognition performance may in fact, not increase cognitive load.

One possibility is that perhaps the representation of language in the long term memory store is not exclusively clean. Instead a representation might reflect a listener's typical experience; a listener may have long term memory stores that reflect experience listening to reverberant speech. Certainly previous investigators have suggested that the quality of listening experiences can influence long-term memory stores and can affect subsequent changes in effort (e.g., Rudner et al. 2008). The concept of a distorted template representation is also consistent with research demonstrating that low reverberation levels (primarily early reflections) can lead to improved speech recognition compared to speech presented with no reverberation. Of course in the case of this study, there was a decrease in word recognition associated with increased reverberation, but perhaps the reverberation distortion of speech that resulted in reduced word recognition was off-set in the cognitive realm because the signal, once recognized, closely matches the representation in long term memory. Although listeners may also have experience listening in noise, they did not have experience listening to the specific four-talker babble used in this study. Perhaps reverberation is more likely to be generalized into long-term memory stores than environmental noise. This is purely conjecture, but future studies may focus on accounting for reverberation in the ELU model, specifically as it relates to the creation of long-term memory stores rather than as an inherent distortion factor.

Future studies

Although these results suggest that reverberation does not increase listening effort for adults with normal hearing over the range of SNRs and reverberation times evaluated, the results may be different for other listeners. Specifically, with regard to speech recognition, some populations have been shown to be more vulnerable to the effects of reverberation than young adults with normal hearing. These populations include children (Johnson 2000; Klatte et al. 2010; Neuman & Hochberg 1983; Neuman et al. 2010; Valente et al. 2012; Wróblewski et al. 2012), older adults (Gordon-Salant & Fitzgibbons 1995; Harris & Reitz 1985; Helfer & Wilber 1990), and listeners with hearing loss (Duquesnoy & Plomp 1980; Finitzo-Hieber & Tillman 1978; Harris & Reitz 1985; Harris & Swenson 1990; Helfer &

Wilber 1990; Náb lek & Pickett 1974b). Future studies should examine the potential interactions between age, hearing loss, reverberation, and listening effort. If indeed these populations prove to be more susceptible to the effects of reverberation, it would provide additional support to the mounting evidence that children, older adults, and listeners with hearing loss are more sensitive to the effects of reverberation and are thus candidates for remediation to counteract the effects of reverberation, such as acoustically engineered classrooms or microphone based techniques for bypassing reverberation in hearing aids. In turn, the effects of these interventions on listening effort will also warrant further investigation.

Acknowledgments

References

American National Standards Institute. Acoustical performance criteria, design requirements and guidelines for schools. Standard S12.60. 2002

Beranek, LL. Acoustics. McGraw-Hill Electrical and Electronic Engineering Series. New York: McGraw-Hill; 1954.

Bradley JS. Speech intelligibility studies in classrooms. J Acoust Soc Am. 1986; 80:846–854. [PubMed: 3760338]

Bradley JS, Reich R, Norcross S. A just noticeable difference in C50 for speech. Applied Acoustics. 1999; 58:99–108.

Bradley JS, Sato H, Picard M. On the importance of early reflections for speech rooms. J Acoust Soc Am. 2003; 113:3233–3244. [PubMed: 12822796]

Connine CM, Mullennix J, Shernoff E, Yelen J. Word familiarity and frequency in visual and auditory word recognition. J Exp Psychol Learn Mem Cogn. 1990; 16:1084. [PubMed: 2148581]

Cox R, Alexander G, Gilmore C. Development of the connected speech test (CST). Ear Hear. 1987; 8:119s–126s. [PubMed: 3678650]

Cox R, Alexander G, Gilmore C, et al. Use of the Connected Speech Test (CST) with hearing-impaired listeners. Ear Hear. 1988; 9:198–207. [PubMed: 3169400]

Culling JF, Hodder KI, Toh CY. Effects of reverberation on perceptual segregation of competing voices. J Acoust Soc Am. 2003; 114:2871–2876. [PubMed: 14650021]

Culling JF, Summerfield Q, Marshall DH. Effects of simulated reverberation on the use of binaural cues and fundamental-frequency differences for separating concurrent vowels. Speech Commun. 1994; 14:71–95.

Darwin C, Hukin R. Effects of reverberation on spatial, prosodic, and vocal-tract size cues to selective attention. J Acoust Soc Am. 2000; 108:335–342. [PubMed: 10923896]

Desjardins JL, Doherty KA. The Effect of Hearing Aid Noise Reduction on Listening Effort in Hearing-Impaired Adults. Ear Hear. 2014 e-publish ahead of print.

Downs D. Effects of hearing aid use on speech discrimination and listening effort. J Speech Hear Disord. 1982; 47:189–193. [PubMed: 7176597]

Duquesnoy A, Plomp R. Effect of reverberation and noise on the intelligibility of sentences in cases of presbyacusis. J Acoust Soc Am. 1980; 68:537–544. [PubMed: 7419810]

Egan, MD. Architectual Acoustics. New York: McGraw-Hill; 1988.

Etymotic Research. BKB-SIN[TM] speech in noise test version 1.03 (compact disc). Elk Grove Village, IL: Etymotic Research; 2005.

Finitzo-Hieber T, Tillman TW. Room acoustics effects on monosyllabic word discrimination ability for normal and hearing-impaired children. J Speech Hear Res. 1978; 21:440–458. [PubMed: 713515]

Fraser S, Gagne J, Alepins M, et al. Evaluating the effort expended to understand speech in noise using a dual-task paradigm: The effects of providing visual speech cues. J Speech Lang Hear Res. 2010; 53:18–33. [PubMed: 19635945]

Gatehouse S, Gordon J. Response times to speech stimuli as measures of benefit from amplification. Br J Audiol. 1990; 24:63–68. [PubMed: 2317602]

Gatehouse S, Noble W. The Speech, Spatial and Qualities of Hearing Scale (SSQ). Int J Audiol. 2004; 43:85–99. [PubMed: 15035561]

George ELJ, Festen JM, Goverts ST. Effects of reverberation and masker fluctuations on binaural unmasking of speech. J Acoust Soc Am. 2012; 132:1581–1591. [PubMed: 22978887]

Gordon-Salant S, Fitzgibbons PJ. Recognition of multiply degraded speech by young and elderly listeners. J Speech Hear Res. 1995; 38(5):1150–1156. [PubMed: 8558883]

Haas H. The influence of a single echo on the audibility of speech. J Audio Eng Soc. 1972; 20:146–159.

Hall, ET. The Silent Language. New York, NY: Anchor; 1959.

Harris RW, Reitz ML. Effects of room reverberation and noise on speech discrimination by the elderly. Int J Audiol. 1985; 24(5):319–324.

Harris RW, Swenson DW. Effects of reverberation and noise on speech recognition by adults with various amounts of sensorineural hearing impairment. Audiology. 1990; 29(6):314–321. [PubMed: 2275646]

Helfer KS, Wilber LA. Hearing loss, aging, and speech perception in reverberation and noise. J Speech Lang Hear Res. 1990; 33:149–155.

Hétu R, Jones L, Getty L. The impact of acquired hearing impairment on intimate relationships: Implications for rehabilitation. Int J Audiol. 1993; 32(6):363–380.

Hétu R, Riverin L, Lalande N, et al. Qualitative analysis of the handicap associated with occupational hearing loss. Br J Audiol. 1988; 22:251–264. [PubMed: 3242715]

Hicks C, Tharpe A. Listening effort and fatigue in school-age children with and without hearing loss. J Speech Lang Hear Res. 2002; 45:573. [PubMed: 12069009]

Hornsby BW. The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. Ear Hear. 2013; 34:523–534. [PubMed: 23426091]

Houtgast T, Steeneken HJ. A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. J Acoust Soc Am. 1985; 77:1069–1077.

Hua H, Karlsson J, Widén S, et al. Quality of life, effort and disturbance perceived in noise: A comparison between employees with aided hearing impairment and normal hearing. Int J Audiol. 2013; 52:642–649. [PubMed: 23808681]

Humes L. Dimensions of hearing aid outcome. J Am Acad Audiol. 1999; 10:26–39. [PubMed: 9921723]

International Standards Organization. Acoustics—Measurement of Room Acoustic Parameters—Part 1: Performance Spaces. Standard 3382-1. 2009 2009.

Johnson CE. Children's phoneme identification in reverberation and noise. J Speech Lang Hear Res. 2000; 43:144–157. [PubMed: 10668658]

Kahneman, D. Attention and Effort. Englewood Cliffs, N.J.: Prentice-Hall; 1973.

Kidd G, Mason CR, Brughera A, et al. The role of reverberation in release from masking due to spatial separation of sources for speech identification. Acustica. 2005; 91:526–536.

Klepousniotou E. The processing of lexical ambiguity: Homonymy and polysemy in the mental lexicon. Brain Lang. 2002; 81:205–223. [PubMed: 12081393]

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; 12:270–282. [PubMed: 20871182]

Knecht HA, Nelson PB, Whitelaw GM, et al. Background Noise Levels and Reverberation Times in Unoccupied Classrooms: Predictions and Measurements. Am J Audiol. 2002; 11:65–71. [PubMed: 12691216]

Kramer S, Kapteyn T, Houtgast T. Occupational performance: Comparing normally-hearing and hearing-impaired employees using the Amsterdam Checklist for Hearing and Work. Int J Audiol. 2006; 45:503–512. [PubMed: 17005493]

Koelewijn T, Zekveld AA, Festen JM, Kramer SE. The influence of informational masking on speech perception and pupil response in adults with hearing impairment. J Acoust Soc Am. 2014; 135:1596–1606. [PubMed: 24606294]

Ljung R, Kjellberg A. Long reverberation time decreases recall of spoken information. Building Acoustics. 2009; 16:301–311.

Ljung R, Sörqvist P, Kjellberg A, Green A-M. Poor listening conditions impair memory for intelligible lectures: Implications for acoustic classroom standards. Building Acoustics. 2009; 16:257–265.

Lochner J, Burger J. The influence of reflections on auditorium acoustics. J Sound Vib. 1964; 1:426–454.

MacKeith N, Coles R. Binaural advantages in hearing of speech. Journal of Laryngology and Otology. 1971; 85:213–232. [PubMed: 5550604]

Mackersie CL, Cones H. Subjective and psychophysiological indices of listening effort in a competing-talker task. J Am Acad Audiol. 2011; 22(2):113–122. [PubMed: 21463566]

McGarrigle R, Munro KJ, Dawes P, Stewart AJ, Moore DR, Barry JG, Amitay S. Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'white paper'. Int J Audiol. 2014; 53:433–445. [PubMed: 24673660]

McCoy S, Tun P, Cox C, et al. Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. Q J Exp Psychol A. 2005; 58:22–33. [PubMed: 15881289]

Moncur JP, Dirks D. Binaural and monaural speech intelligibility in reverberation. J Speech Lang Hear Res. 1967; 10:186–195.

Morimoto M, Sato H, Kobayashi M. Listening difficulty as a subjective measure for evaluation of speech transmission performance in public spaces. J Acoust Soc Am. 2004; 116:1607–1613. [PubMed: 15478426]

Murphy D, Craik F, Li K, Schneider B. Comparing the effects of aging and background noise on short-term memory performance. Psychol Aging. 2000; 15:323–334. [PubMed: 10879586]

Nabelek AK, Mason D. Effect of noise and reverberation on binaural and monaural word identification by subjects with various audiograms. J Speech Hear Res. 1981; 24:375–383. [PubMed: 7300279]

Náb lek AK, Pickett J. Reception of consonants in a classroom as affected by monaural and binaural listening, noise, reverberation, and hearing aids. J Acoust Soc Am. 1974a; 56:628–639. [PubMed: 4415491]

Náb lek AK, Pickett JM. Monaural and binaural speech perception through hearing aids under noise and reverberation with normal and hearing-Impaired listeners. J Speech Hear Res. 1974b; 17:724–739. [PubMed: 4444292]

Nábelek AK, Robinette L. Influence of the precedence effect on word identification by normally hearing and hearing-impaired subjects. J Acoust Soc Am. 1978; 63:187–194. [PubMed: 632410]

Nachtegaal J, Kuik DJ, Anema JR, et al. Hearing status, need for recovery after work, and psychosocial work characteristics: Results from an internet-based national survey on hearing. Int J Audiol. 2009; 48:684–691. [PubMed: 19863354]

Neuman AC, Hochberg I. Children's perception of speech in reverberation. J Acoust Soc Am. 1983; 73:2145–2149. [PubMed: 6875100]

Neuman AC, Wróblewski M, Hajicek J, et al. Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. Ear Hear. 2010; 31:336–334. [PubMed: 20215967]

Ng EHN, Rudner M, Lunner T, et al. Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. Int J Audiol. 2013; 52:433–441. [PubMed: 23550584]

Peutz V. Articulation loss of consonants as a criterion for speech transmission in a room. J Audio Eng Soc. 1971; 19:915–919.

Pichora-Fuller M, Schneider B, Daneman M. How young and old adults listen to and remember speech in noise. J Acoust Soc Am. 1995; 97:593–608. [PubMed: 7860836]

Picou EM, Aspell E, Ricketts TA. Potential benefits and limitations of three types of directional processing in hearing aids. Ear Hear. 2014; 35:339–352. [PubMed: 24518429]

Picou EM, Ricketts TA. The effect of changing the secondary task in dual-task paradigms for measuring listening effort. Ear Hear. 2014; 35:611–622. [PubMed: 24992491]

Picou EM, Ricketts TA, Hornsby BW. How hearing aids, background noise, and visual cues influence objective listening effort. Ear Hear. 2013; 34:e52–e64. [PubMed: 23416751]

Picou EM, Ricketts TA, Hornsby BWY. Visual cues and listening effort: Individual variability. J Speech Lang Hear Res. 2011; 54:1416–1430. [PubMed: 21498576]

Plomp R. Binaural and monaural speech intelligibility of connected discourse in reverberation as a function of azimuth of a single competing sound source (speech or noise). Acustica. 1976; 34:200–211.

Rabbitt P. Channel-capacity, intelligibility and immediate memory. Q J Exp Psychol. 1968; 20:241–248. [PubMed: 5683763]

Rabbitt P. Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. Acta Oto-Laryngologica. 1991; 111:167–176.

Rennies J, Schepker H, Holube I, Kollmeier B. Listening effort and speech intelligibility in listening situations affected by noise and reverberation. J Acoust Soc Am. 2014; 136:2642–2653. [PubMed: 25373965]

Rönnberg J, Lunner T, Zekveld A, et al. The Ease of Language Understanding (ELU) model: Theory, data, and clinical implications. Front Syst Neurosci. 2013; 7:1–17. [PubMed: 23420631]

Rönnberg J, Rudner M, Foo C, Lunner T. Cognition counts: A working memory system for ease of language understanding (ELU). Int J Audiol. 2008; 47:S99–S105. [PubMed: 19012117]

Rudner M, Foo C, Sundewall-Thorén E, et al. Phonological mismatch and explicit cognitive processing in a sample of 102 hearing-aid users. Int J Audiol. 2008; 47:S91–S98. [PubMed: 19012116]

Sarampalis A, Kalluri S, Edwards B, et al. Objective measures of listening effort: Effects of background noise and noise reduction. J Speech Lang Hear Res. 2009; 52:1230–1240. [PubMed: 19380604]

Sato H, Morimoto M, Sato H, Wada M. Relationship between listening difficulty and acoustical objective measures in reverberant sound fields. J Acoust Soc Am. 2008; 123:2087–2093. [PubMed: 18397016]

Sato H, Morimoto M, Wada M. Relationship between listening difficulty rating and objective measures in reverberant and noisy sound fields for young adults and elderly persons. J Acoust Soc Am. 2012; 131:4596–4605. [PubMed: 22712933]

Soulodre G, Popplewell N, Bradley JS. Combined effects of early reflections and background noise on speech intelligibility. J Sound Vib. 1989; 135:123–133.

Studebaker GA. A "rationalized" arcsine transform. J Speech Hear Res. 1985; 28:455–462. [PubMed: 4046587]

Surprenant A. The effect of noise on memory for spoken syllables. Int J Psychol. 1999; 34:328–333.

Valente DL, Plevinsky HM, Franco JM, et al. Experimental investigation of the effects of the acoustical conditions in a simulated classroom on speech recognition and learning in children. J Acoust Soc Am. 2012; 131:232. [PubMed: 22280587]

Winckel FW. Optimum acoustic criteria of concert halls for the performance of classical music. J Acous Soc Am. 1962; 34:81–86.

Wróblewski M, Lewis DE, Valente DL, et al. Effects of reverberation on speech recognition in stationary and modulated noise by school-aged children and young adults. Ear Hear. 2012; 33:731–744. [PubMed: 22732772]

Zekveld A, Kramer S, Festen J. Pupil response as an indication of effortful listening: The influence of sentence intelligibility. Ear Hear. 2010; 31:480–490. [PubMed: 20588118]
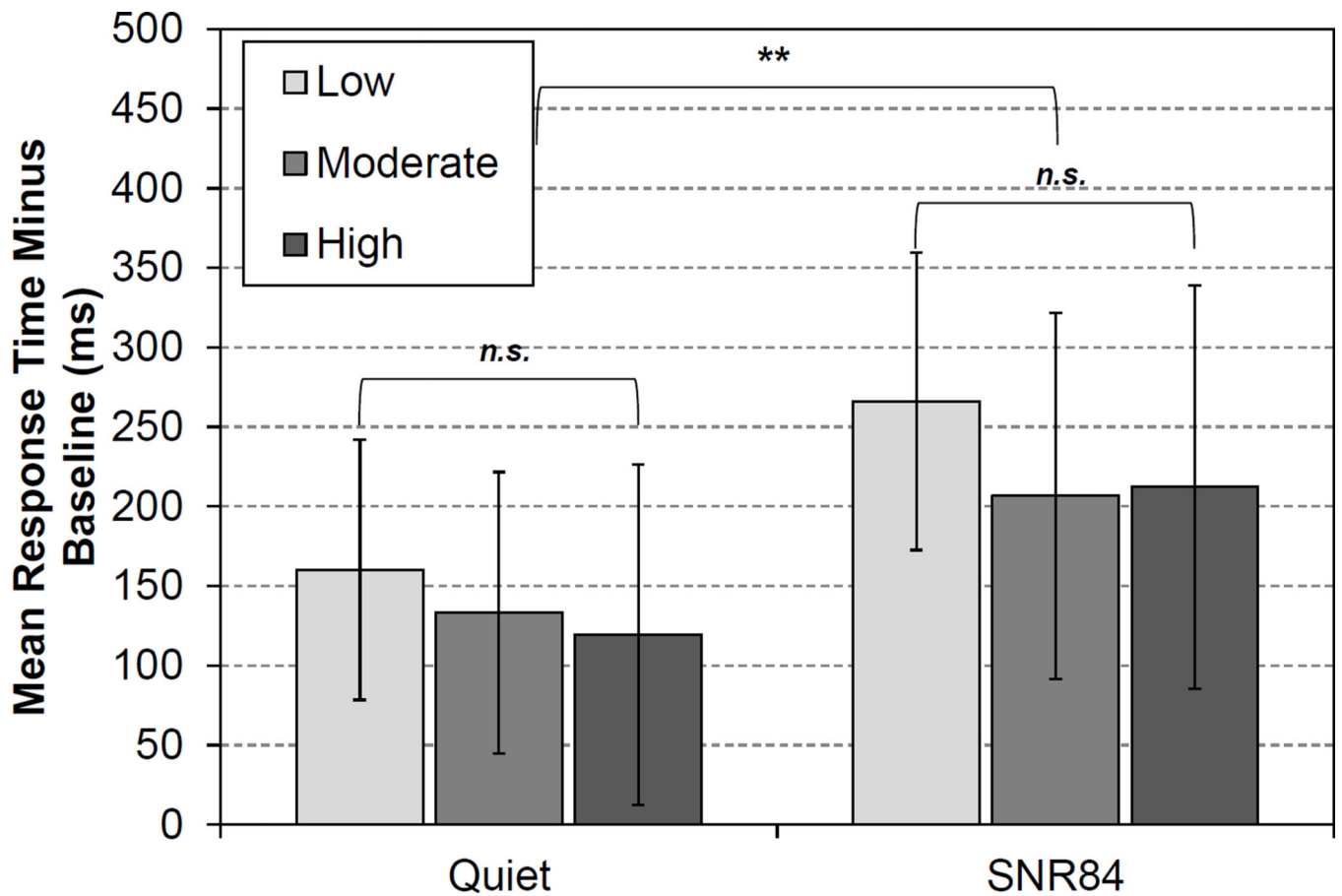
Figure 1.
Mean response times minus baseline (ms) for quiet and noise with comparable word recognition performance (research question A). In noise, the SNR was varied to approximate 84% word recognition (SNR84). Light grey bars, medium gray bars, and dark gray bars indicate performance in low, moderate, and high reverberation environments, respectively. Error bars indicate ±1 standard deviation from the mean. Significant effects are indicated (p < 0.01); non-significant effects are marked with "n.s"
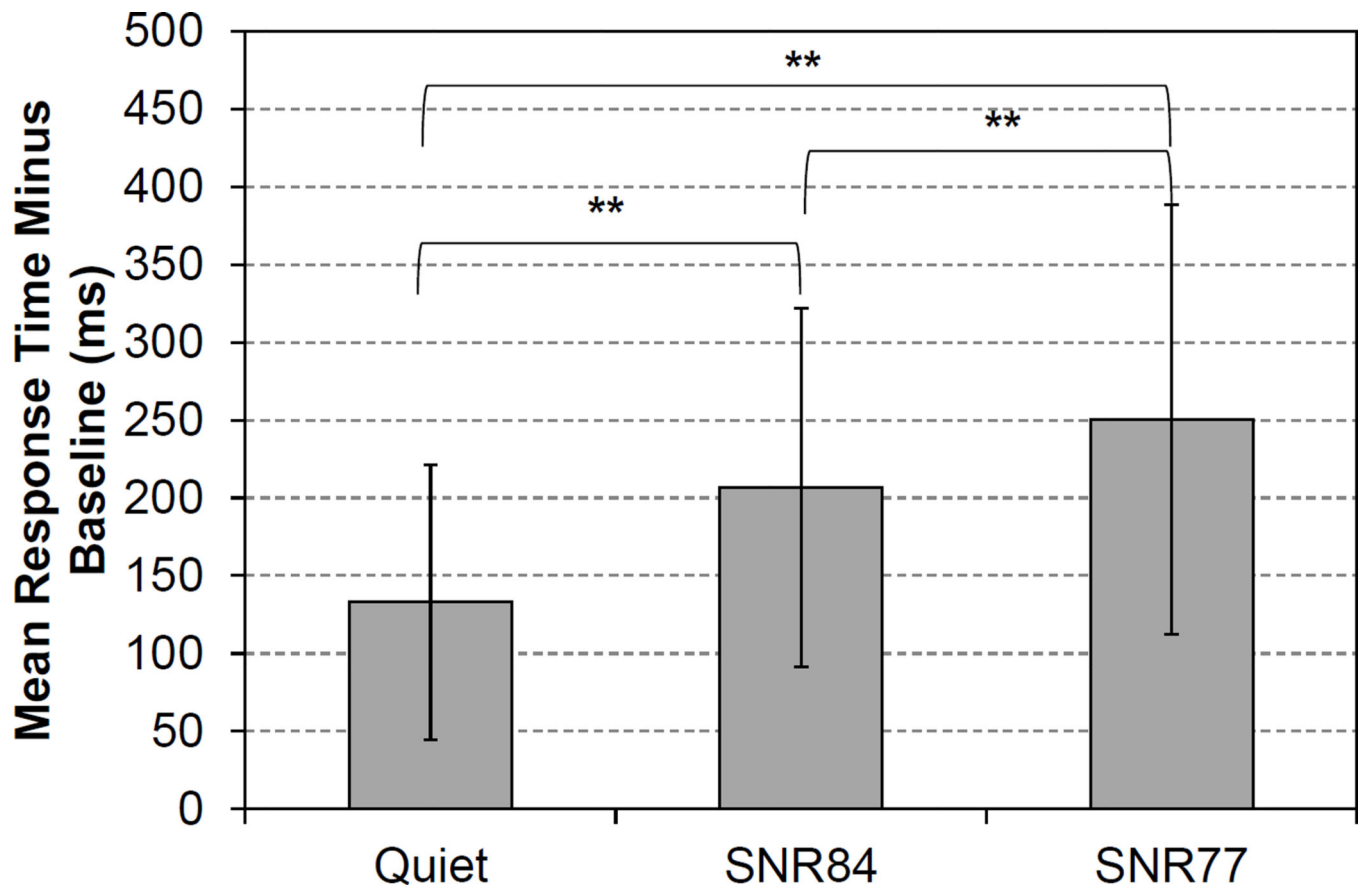
Figure 2.
Mean response times (ms) for the constant reverberation (T30 = 475 ms) condition in quiet, noise (SNR84), and noise (SNR77; research question B). Error bars indicate ±1 standard deviation from the mean. Significant effects are indicated.
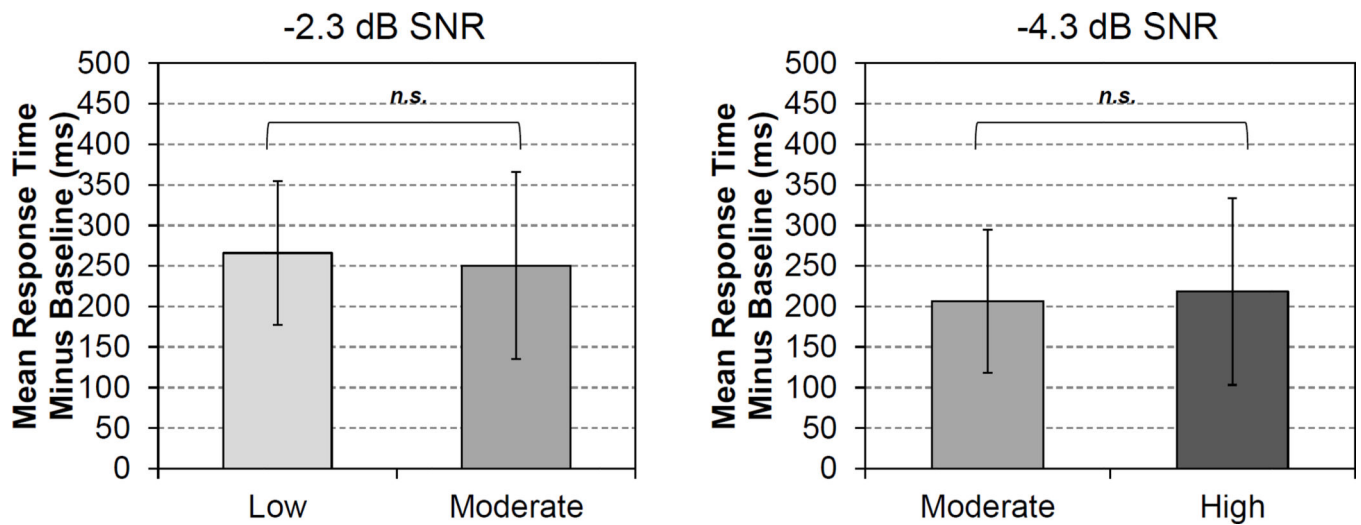
Figure 3.
Mean response times (ms) for the constant signal-to-noise ratio (SNR) conditions when reverberation was changed from low to moderate (left panel; research question C) and from moderate to high (right panel; research question D). Error bars indicated ±1 standard deviation from the mean. Non-significant effects are marked with "n.s"

Table 1

Condition description and analysis. Letters indicate which conditions were used to answer each of the four specific research questions. These questions were A) With comparable word recognition performance across levels of reverberation, what are the effects of noise and reverberation on listening effort? (B) Within a constant reverberation level T30 = 475ms), what is the effect of noise on listening effort? (C) For a given SNR, what is the effect on listening effort of increasing reverberation from low (T30 < 100 ms) to moderate (T30 = 475 ms) and word recognition performance is degraded? (D) For a given SNR, what is the effect on listening effort of increasing reverberation from moderate (T30 = 475 ms) to high (T30 = 834 ms) and word recognition performance is degraded?

	Low	Moderate	High
Quiet	A	A, B	A
Avg SNR = −0.3 dB (σ = 0.9)			A
Avg SNR = −2.3 dB (σ = 0.9)		A, B, D	D
Avg SNR = −4.3 dB (σ = 0.9)	A, C	B, C	

Table 2

Reverberation time (T30; ms) as a function of 1/3 octave band (kHz) for moderate and high reverberation conditions tested.

Frequency	0.25	0.315	0.4	0.5	0.63	0.8	1.0	1.25	1.6	2.0	2.5	3.15	4.0	5.0	6.3	Mean
Moderate	648	421	522	572	423	503	392	407	466	382	512	460	474	499	438	**475**
High	1160	1196	935	856	839	793	812	874	637	728	802	793	671	716	794	**834**

Table 3

Mean word recognition performance (RAU) for each of the conditions tested. Numbers in parentheses represent standard deviations. Filled circles reflect worse performance in the high reverberation condition ($p < 0.05$) for research question A (comparable word recognition performance).

	Low	Moderate	High
Quiet	117.4 (2.9)	116.6 (4.2)*	113.1 (5.0)•
Avg SNR = −0.3 dB (σ = 0.9)			83.9 (6.1)•
Avg SNR = −2.3 dB (σ = 0.9)		86.2 (7.5)*‡	77.7 (7.7)‡
Avg SNR = −4.3 dB (σ = 0.9)	85.6 (6.6)†	78.1 (7.4)*†	

Asterisks indicate significant main effects ($p < 0.001$) of comparisons for research question B (constant reverberation).

Crosses (†) indicate a significant main effect of reverberation (<100 to 475ms) within a fixed SNR ($p < 0.001$) for research question C.

Double crosses (‡) indicate a significant main effect of reverberation (475 to 834ms) within a fixed SNR ($p < 0.001$) for research question D.

Table 4

Results of the repeated measures analysis of variance tests conducted for each research question of interest regarding word recognition performance (RAU). Significant main effects and interactions are indicated by bold typeface.

	Noise	Reverberation	Interaction
Question A (comparable word recognition)	**$F_{1,17} = 631.9$, $p < 0.001$, partial $\eta^2 = 0.97$**	**$F_{2,16} = 4.7$, $p < 0.05$, partial $\eta^2 = 0.37$**	$F_{1,17} = 1.0$, $p = 0.40$, partial $\eta^2 = 0.11$
Question B (constant reverberation)	**$F_{1,17} = 152.04$, $p < 0.001$, partial $\eta^2 = 0.95$**		
Question C (constant SNR – low to moderate reverberation)		**$F_{1,17} = 16.2$, $p < 0.01$, partial $\eta^2 = 0.49$**	
Question D (constant SNR – moderate to high reverberation)		**$F_{1,17} = 22.9$, $p < 0.001$, partial $\eta^2 = 0.56$**	

Table 5

Mean response times minus baseline (ms) for each of the conditions tested. Numbers in parentheses represent standard deviations. Filled circles indicate significant effects of background noise ($p < 0.01$) for research question A (comparable word recognition).

	Low	Moderate	High
Quiet	160.1 (81.8)•	133.1 (88.4)•*	119.4 (107.1)•
Avg SNR = −0.3 dB (σ = 0.9)			212.3 (126.8)•
Avg SNR = −2.3 dB (σ = 0.9)		206.6 (115.2)•*	218.3 (127.1)
Avg SNR = −4.3 dB (σ = 0.9)	265.9 (93.3)•	250.5 (138.1)*	

Asterisks indicate significant main effects of noise ($p< 0.001$) of comparisons for research question B (constant reverberation). There were no significant effects of reverberation on response times for research questions A, C, or D.

Table 6

Results of the analysis of variance tests conducted for each research question of interest regarding listening effort (reaction times in ms). Significant main effects and interactions are indicated with bold typeface.

	Noise	Reverberation	Interaction
Question A (comparable word recognition)	**$F_{1,17} = 47.21$, $p < 0.001$, partial $\eta^2 = 0.74$**	$F_{2,16} = 1.92$, $p = 0.18$, partial $\eta^2 = 0.19$	$F_{1,17} = 1.66$, $p = 0.22$, partial $\eta^2 = 0.17$
Question B (constant reverberation)	**$F_{1,17} = 10.13$, $p < 0.01$, partial $\eta^2 = 0.56$**		
Question C (constant SNR – low to moderate reverberation)		$F_{1,17} = 0.37$, $p = 0.55$, partial $\eta^2 = 0.02$	
Question D (constant SNR – moderate to high reverberation)		$F_{1,17} = 0.20$, $p = 0.66$, partial $\eta^2 = 0.01$	