8-2012

Effects of acoustic environments on speech comprehension by native-English-speaking listeners

Zhao Peng
University of Nebraska - Lincoln, zpeng@huskers.unl.edu

Lily M. Wang
University of Nebraska - Lincoln, lwang4@unl.edu

Siu Kit Lau
University of Nebraska-Lincoln, slau3@unl.edu

Effects of acoustic environments on speech comprehension by native-English-speaking listeners

Zhao Peng [a)]
Lily M. Wang [b)]
Siu-Kit Lau [c)]
Durham School of Architectural Engineering and Construction, University of Nebraska-Lincoln, 1110 S. 67th Street, Omaha, NE 68182-0816

This study investigates the effects of acoustic conditions on speech comprehension, rather than speech intelligibility as often reported in existing literature. Sets of 15-minute-long listening comprehension tests were developed based on the format of the Test of English for International Communication (TOEIC). Each test set includes four types of tasks: matching aural phrases to photographs, selecting appropriate responses to aural questions, and answering questions after listening to conversations (between two talkers) and talks (single talker). Within the Nebraska acoustics test chamber, native-English-speaking participants are asked to perform these tests under 15 acoustic conditions, from combinations of three background noise levels (RC-30, 40 and 50) and five mid-frequency reverberation times (0.4 to 1.2 seconds). The background noise levels are varied via an Armstrong i-Ceiling system, while the reverberation times are simulated from convolving the anechoic test signals with binaural room impulse responses (BRIR), simulated in ODEON for a typical classroom. A two-channel playback system is used to present the convolved audio signals, with loudspeaker-listener configuration embedded in the BRIR auralization output. Pilot testing of three subjects showed no variation of performance scores on overall tasks among all acoustical conditions. However, participants generally scored lowest in tasks to comprehend conversations in the longest RT scenarios.

[a)] email: zpeng@huskers.unl.edu
[b)] email: lwang4@unl.edu
[c)] email: slau3@unl.edu

1 INTRODUCTION

Recent studies regarding the effects of room acoustic conditions in classroom environments, such as noise and reverberation, have advanced beyond looking at only speech intelligibility and have focused more on student learning performance. A study by Shield and Dockrell[1] showed that environmental noise, both internal and external, has negative impact on the academic performance and attainment among primary school children in the UK. However, their results were countered by Xie, Kang, and Tompsett[2] showing that no such significant relationship existed. Another study by Ronsse and Wang[3] found a statistically significant relationship between higher background noise levels due to mechanical equipment and lower scores on standardized achievement tests in the U.S.

Klatte, Lachmann, and Meis[4] investigated language comprehension in a classroom-like setting under four combinations of noise and reverberation time (RT) conditions (2 background sound types X 2 reverberation times). Reverberation was simulated using a virtual room technique through an electroacoustic system in the test lab. Following a between-subjects experimental design, each participant was assigned to one of the four acoustical conditions. Their results suggest that listening comprehension (paper-and-pencil instructional task) is more impaired than mere speech perception (word-to-picture matching task) among children under the presence of background speech and classroom activity noise. This is further supported by Valente et al[5], who found in a more controlled lab setting that listening comprehension tasks are more significantly and negatively affected under reverberation and noise than sentence recognition tasks. Four different room conditions were simulated by superimposing the virtual sound field from computer simulation on the test lab, but since only a few listening comprehension tasks were developed, Valente et al used a nested experimental design in which each subject was only exposed to one acoustic condition and one listening task (classroom-learning or sentence recognition tasks). Based on their results, they conclude that acoustical conditions required for comprehension tasks are more stringent than simple speech intelligibility tasks.

The current study seeks to more comprehensively investigate the effects of room acoustic conditions on speech comprehension. The study utilizes a larger test question database so that a within-subjects experimental design may be used, in which all test participants are subjected to all acoustic conditions. Additionally, a wide range of room acoustic conditions are created in a controlled lab facility by convolving the sets of anechoic listening comprehension materials with simulated binaural room impulse responses (BRIRs).

2 METHODOLOGY

This section describes the screening and listening comprehension tests that have been adapted for use in this study, followed by a description of the testing environment and the acoustic test conditions. A pilot test had been conducted on three native-English-speaking test participants, each of whom has a hearing threshold below 25 dB hearing level across the octave bands from 125 Hz to 8 kHz on both ears.

2.1 English Language Skill Tests

The ability to comprehend speech is affected not only by the acoustic environment a person is in, but also by his/her individual language skills and cognitive abilities. In order to isolate these confounding factors that contribute to listening comprehension, three types of English language skill tests are implemented prior to the main experiment to provide information about each subject's baseline language and cognitive abilities.

The Bilingual Verbal Ability Tests (BVAT)[6] was chosen as the first part of the English proficiency test. The BVAT is intended to measure the verbal ability of a non-native English speaking person (age between 5 to >90 years) in both English and native languages. There are four tasks in the BVAT: 1) name the object shown in a drawing, 2) identify the synonym of a word, 3) identify the antonym of a word, and 4) find verbal analogies between two pairs of words. Under each section of tasks, question items become progressively more difficult. All question items are presented visually on the BVAT test book and aurally prompted by a proctor. A complete BVAT test involves administering the test first in English to individual participants until they reach the limit of their English comprehension, and then switching to test additional questions in their native language. Since all participants in this study are native-English speakers, only the English half of the BVAT is utilized in the current investigation.

The second test of English language skills to be used in this study was adopted from the oral comprehension subtest of the Woodcock-Johnson III Tests of Achievement[7]. In this test, subjects are asked to listen to single sentences and aurally supply the missing word to complete the sentence based on the syntactic and semantic information. This test is suitable for testing subjects of 5 years or older to measure a subject's baseline ability of listening comprehension.

In addition to verbal skills and listening abilities, speech comprehension may also be affected by individual cognitive abilities. Daneman and Carpenter[8] found subjects with larger listening span performed better at answering questions about facts and pronominal references that were orally presented. Their results suggested a positive relationship between larger listening span and better performance on listening comprehension. Hence, a third test of memory for sentences from the Diagnostic Supplement to the Woodcock-Johnson III Tests of Cognitive Abilities[7] is administered to measure each subject's listening span. Subjects are asked to repeat single sentences exactly as they are presented in the audio recordings in this test.

The speech materials for the two Woodcock-Johnson III tests were recorded by a native-English speaking female student at a close-microphone position in the University of Nebraska-Lincoln Indoor Environmental Test Chamber. The test chamber has a natural occurring background noise level of 35 dBA. These audio files were recorded using the Bruel&Kjær PULSE Time Data Recorder program and a 1/2-inch microphone and later formatted to have a sampling frequency of 44.1 kHz and 16-bit digital resolution. The audio recordings are presented to subjects over a set of Sennheiser HD497 headphones at L_{Aeq} of 63 dBA (L_{Amax} of 73 dBA) and played back using the Winamp Media Player on a Dell Precision M2400 laptop computer with internal sound card.

2.2 Noise Sensitivity Questionnaire

Several studies have shown that subjective noise sensitivity is significantly related to noise annoyance and may correlate to one's performance under noisy conditions, as well[9]. In the current study, a reduced version of the recently developed Noise-Sensitivity-Questionnaire (NoiSeQ) by Schutte et al[10-11] is administered to the test subjects during the screening session. This questionnaire has a total of 13 items in three subscales: sleep, work, and residential surroundings, and is available online in English[12].

2.3 Listening Comprehension Materials

The speech intelligibility tests that have been used often in the past are based on word or sentence recognition, rather than comprehension of the spoken verbal information. For this study, 18 sets of equivalent listening comprehension test were developed, based on the format of tests used in the Test of English for International Communication (TOEIC). Each test set takes approximately 15 minutes to complete and is composed of four types of listening comprehension tasks.

1) Photo recognition: Subjects listen to four sentences and are asked to identify the sentence that best describes the photo displayed.
2) Question and response: Subjects are presented with a question aurally and are asked to select the response that best answers that question from four possible responses. The asker and respondent are of opposite gender and alternate with every other item presented.
3) Conversations between two talkers: Subjects listen to a conversation between two talkers and must answer three questions pertinent to the content of the conversation after it is finished. The questions are presented aurally while the response options are displayed visually.
4) Talks by a single talker: This task is similar to that in part (3) but there is only a single talker.

All listening comprehension materials were recorded in the anechoic chamber at Armstrong World Industries by native-English speakers. The anechoic recordings were formatted at a sampling frequency of 44.1 kHz and 16-bit digital resolution. Each listening comprehension test is convolved with a number of simulated BRIRs as discussed in sub-section 2.5.

2.4 Testing Environment

All tests are conducted in the University of Nebraska's new Acoustics Test chamber. The chamber has an interior volume of 25.5 m^3 with two slightly slanted walls and is fitted using typical materials – carpet on floor and gypsum board wall construction, with an average mid-frequency RT of 0.49 seconds. Additional foams were added at three corners to dampen modes and reduce the baseline RT to 0.35 seconds. A tablet chair is placed near the center of the test chamber. The three steady-state background noise conditions are produced in the test chamber via an Armstrong i-ceiling panel loudspeaker and a supplemental JBL Northridge E250P subwoofer placed in the corner of the room. Two JBL Type HS50M monitor speakers, connected to a PreSonus AudioBox 44VSL external sound card, are set up on adjustable loudspeaker stands at approximately 1.52 m away symmetrically in front of the subject. The anechoic listening comprehension speech materials are convolved with the BRIRs from the five RT scenarios and played back through the two-channel loudspeaker system. The playback level was calibrated at 59 dBA at the listener position for all acoustical conditions, as discussed in the following sub-section.

A 23-inch LCD monitor mounted on tripod is located between the speakers at approximately 1.37 m in front of subject. See Fig. 1 for test chamber layout. A computer-based testing program interface was developed for this study for display on the LCD monitor. The laptop computer hosting the testing program interface, amplifiers, and other equipment are located in a control room outside of the test chamber.

2.5 Acoustical Conditions

Combinations of three background noise levels (BNL) and five RTs are being tested, totaling 15 acoustical conditions. The three BNL conditions match room criteria curves RC 30 hissy, 40 neutral and 50 vibrational. The resulting A-weighted BNL are 38, 48 and 58 dBA, respectively. BRIRs of the five RT scenarios were simulated in ODEON (Version 11.00 Combined) using a typical classroom with an interior volume of 260 m^3. The permanent surface materials of the classroom included thin carpet on floor, drywall with glazed marble blackboard on the front wall, a glass window on one side wall, and a glass door on the back wall. Combinations of three ceiling types and back and two side walls, with scaled absorption coefficients of 25mm acoustical wall panel, were utilized to achieve the five simulated RTs. The scaling factors are uniformly applied on the acoustical wall panel absorption coefficients across octave band frequencies. The simulated RTs and the corresponding surface material combinations are shown in Table 1. The simulations provided a well-sample range of RTs, given as T30 averaged across mid-frequencies (500, 1000 and 2000 Hz): 0.39, 0.60, 0.81, 1.01 and 1.18 seconds (Fig. 2). One additional RT condition of 0.30 seconds was selected for use as the first session of each participant's main testing sessions, to provide a scenario for training purposes and not to be included in the final analysis. The presentation order of all test sets and acoustic conditions to the test participants are randomized and balanced.

Because of the reverberation in test chamber, the actual RTs perceived by participants seated at the listener position were not identical to those of the simulated scenarios. The perceived RTs were identified using interrupted noise method by convolving the simulated BRIRs with pink noise. A Larson Davis 824 sound level meter was utilized to record the perceived RTs at the receiver position. Measurements were repeated ten times for each simulated RT scenario. The mid-frequency T20s averaged across the ten trials, with error bars indicating one standard deviation, are shown in Fig. 3.

In the ODEON model, source and receiver were positioned at 4 m apart, simulating a typical teacher-student configuration in classrooms. The recommended reduction in sound pressure level (SPL) per doubling distance is 3.5 dB for reverberant-to-diffuse field, as specified in the ODEON user manual. The resulting SPL at the receiver position is 59 dBA if the talker is speaking with a raised voice of 66 dBA. Therefore, the playback level at the listener position in the test chamber was calibrated and fixed at 59 dBA for all acoustical conditions to mimic the receiver position in the ODEON model.

3 RESULTS

Three male college students, with ages from 24 to 25 years old, participated in the pilot study. No statistical analyses were performed for the pilot group on the initial screening tests or the listening comprehension tests due to the small sample size.

The average performance scores in percent correct, summed across all four listening comprehension tasks, are shown in Fig. 4 for all 15 acoustical conditions. As indicated on this figure, the total average performance scores do not show pronounced variations among conditions. The average performance score for the conversation portion only of the listening comprehension tasks is shown in Fig. 5. With the exception of the 0.57 sec RT/RC-50 BNL condition, there exhibits a general trend that participants score lowest in the longest RT scenario in the respective BNL conditions. The 0.57 sec RT/RC-50 BNL condition resulted in the lowest performance score in the listening comprehension task of conversations. We suspect this may be an outlier data point among the participants in this condition. More statistical analyses will be performed once testing on 30 native-English-speaking participants has been completed.

4 ACKNOWLEDGEMENTS

This project has been supported by a UNL Durham School of Architectural Engineering and Construction Faculty Seed Grant and by the Paul S. Veneklasen Research Foundation. The authors would like to thank these sponsors for their support.

The authors are also grateful to our collaborators at Armstrong World Industries (particularly Dr. Kenneth P. Roy and Sean Browne) for their assistance in recording the anechoic listening comprehension materials and lending monitor speakers for the subjective testing.

5 REFERENCES

1. B. Shield and J. Dockrell, "The effects of environmental and classroom noise on the academic attainments of primary school children," *J. Acoust. Soc. Am.*, **123**(1), 133-144, (2008).

2. H. Xie, J. Kang and R. Tompsett, "The impacts of environmental noise on the academic achievements of secondary school students in Greater London," *Appl. Acs*., **72**(8), 551-555, (2011).

3. L. M. Ronsse and L. M. Wang, "AB-10-C037: Effects of noise from building mechanical systems on elementary school student achievement," *ASHRAE Transactions,* **116**(2), 347-354, (2010).

4. M. Klatte, T. Lachmann and M. Meis, "Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting," *Noise and Health*, **12**(49), 270-82, (2010).

5. D. Valente, H. Plevinsky, J. Franco, E. Heinrichs-Graham and D. Lewis, "Experimental investigation of the effects of the acoustical conditions in a simulated classroom on speech recognition and learning in children," *J. Acoust. Soc. Am*., **131**(1), 232-246, (2012).

6. A. F. Munoz-Sandoval, J. Cummins, G. Alvarado and M. L. Ruef, *Bilingual Verbal Ability Test*, Riverside Publishing, (2005).

7. R. W. Woodcock, K. S. McGrew and N. Mather, *Woodcock-Johnson III Normative Update (NU) Complete*, Rolling Meadows: Riverside Publishing, (2007).

8. M. Daneman and P. A. Carpenter, "Individual differences in working memory and reading," *J. of Verbal Learning and Verbal Behavior*, **19**(4), 450-466, (1980).

9. W. Ellermeier, M. Eigenstetter and K. Zimmer, "Psychoacoustic correlates of individual noise sensitivity," *J. Acoust. Soc. Am*., **109**(4), 1464-73, (2001).

10. M. Schutte, A. Marks, E. Wenning and B. Griefahn, "The development of the noise sensitivity questionnaire," *Noise and Health*, **9**(34), 15-24, (2007).

11. M. Schutte, S. Sandrock and B. Griefahn, "Factorial validity of the noise sensitivity questionnaire," *Noise and Health*, **9**(37), 96-100, (2007).

12. B. Griefahn, “Silence,” [Online]. Available at: *http://www.ifado.de/silence/* [Last accessed 31 May 2012].

Table 1 – Summary of added absorptive materials for each reverberation time scenario

Simulated RT Scenario	Location of added absorptive materials in ODEON simulation model
1) T_{30} = 0.39 sec	Full NRC 0.7 ceiling tile, back and two side walls with absorption coefficients at 60% of 25mm acoustical wall panel
2) T_{30} = 0.60 sec	Full NRC 0.7 ceiling tile, back and two side walls with absorption coefficients at 30% of 25mm acoustical wall panel
3) T_{30} = 0.81 sec	Full NRC 0.55 ceiling tile, back and two side walls with absorption coefficients at 15% of 25mm acoustical wall panel
4) T_{30} = 1.01 sec	Full NRC 0.55 ceiling tile, back and two side walls with absorption coefficients at 5% of 25mm acoustical wall panel
5) T_{30} = 1.18 sec	Full gypsum board ceiling, back and two side walls with absorption coefficients at 9% of 25mm acoustical wall panel

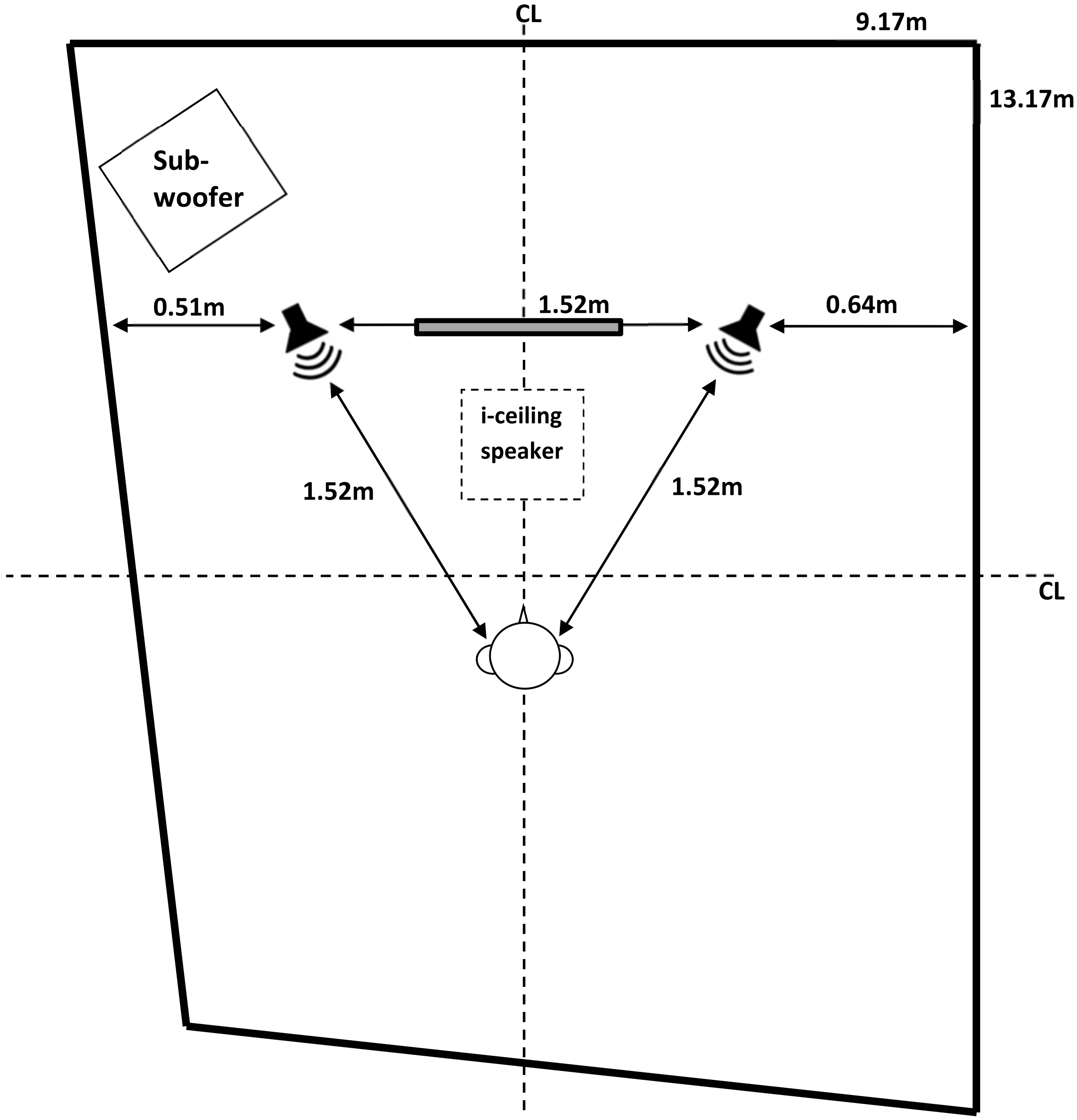

Fig. 1 – Plan view of monitor speakers, LCD monitor, and test subject position in Indoor Environmental Test Chamber

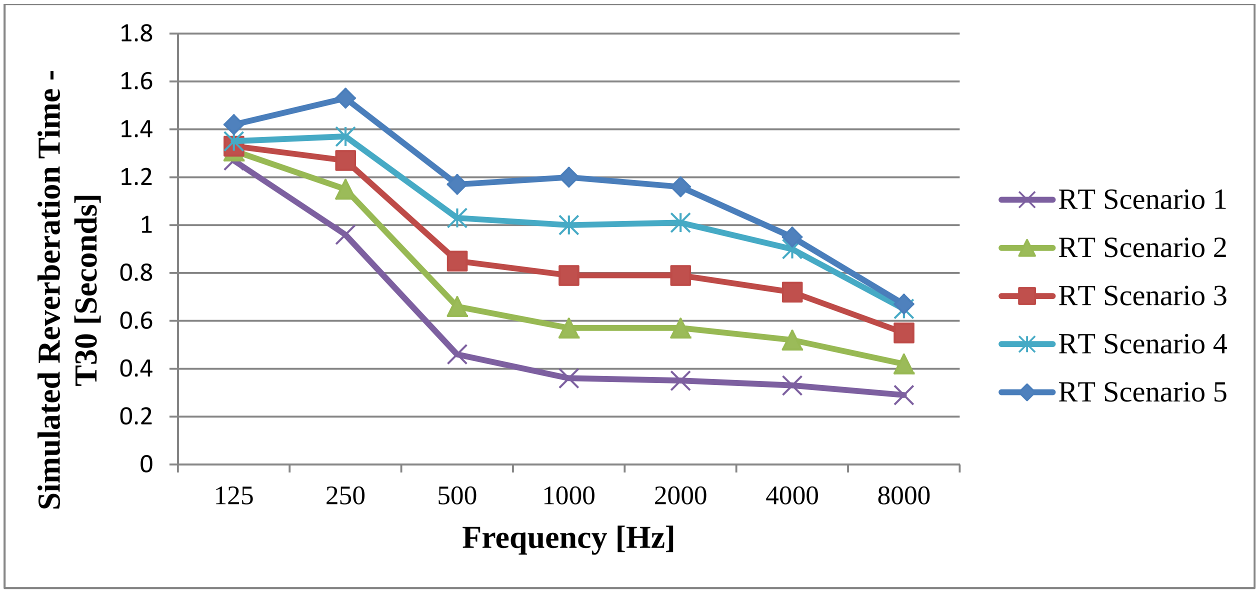

Fig. 2 – Simulated reverberation times (T_{30}) from the binaural room impulse responses in five reverberation time scenarios

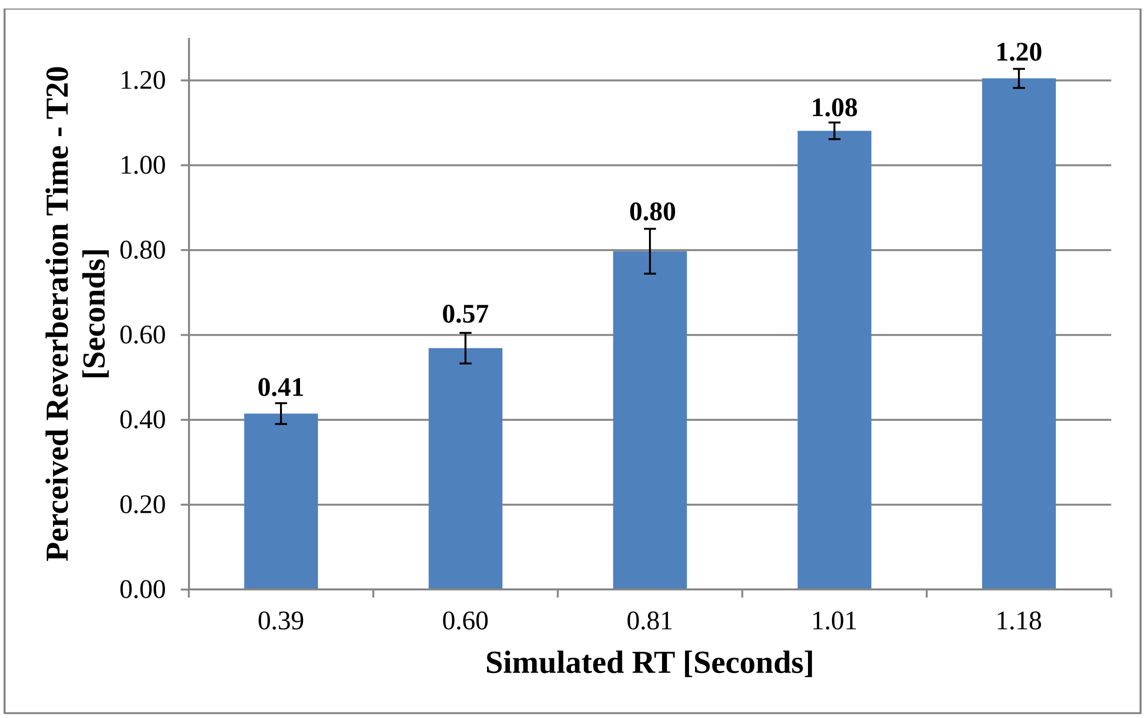

Fig. 3 – Perceived reverberation times (T20) at the listener position in the test chamber for the five reverberation time scenarios

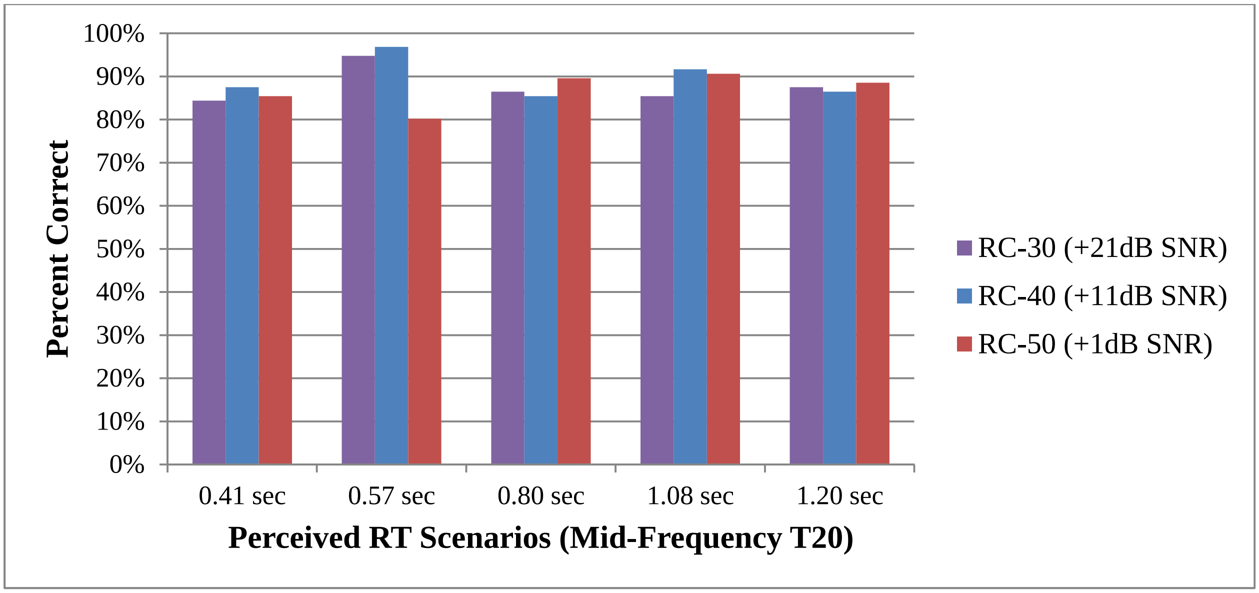

Fig. 4 – Average total performance scores over all four listening comprehension tasks across the 15 acoustical conditions

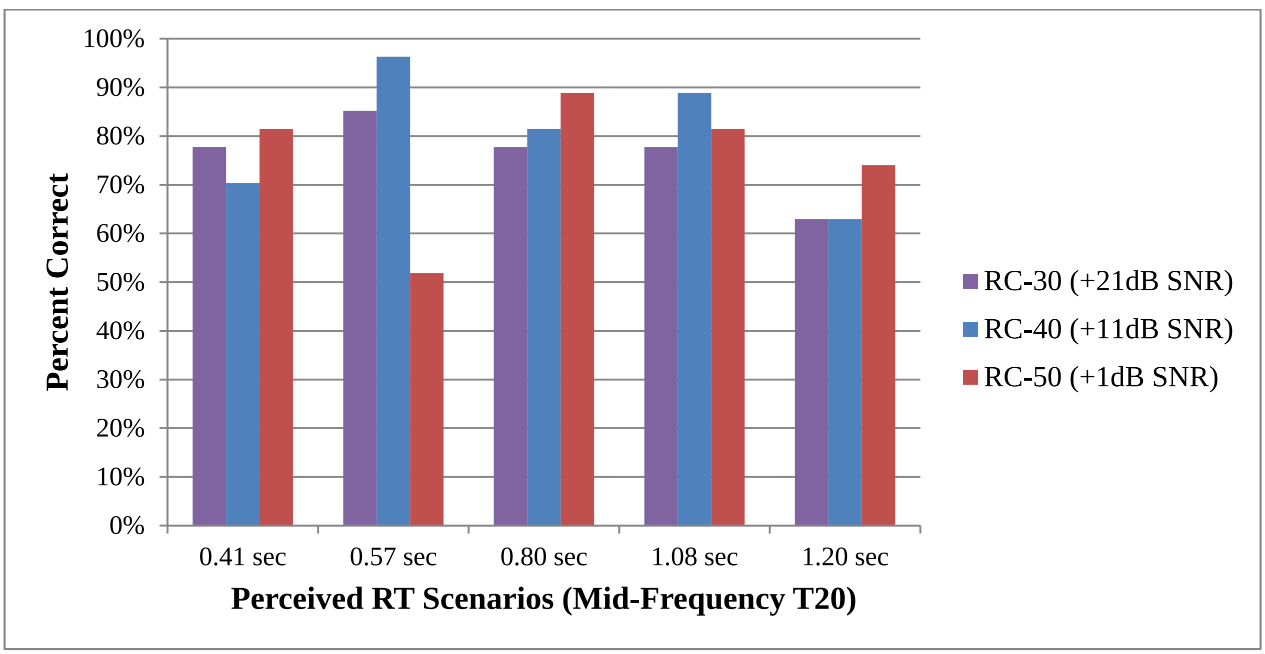

Fig. 5 – Average performance scores on the conversation portion only of the listening comprehension tasks across the 15 acoustical conditions