Effects of room acoustics on comprehension of foreign-accented speech by native and non-native English-speaking listeners

Zhao PENG[1]; Kristin E. HANNA[2]; Brenna N. BOYD[3]; Lily M. WANG[4]

[1,2,3,4] University of Nebraska-Lincoln

ABSTRACT

In a previous study by the authors, reverberation time (RT) and background noise level (BNL) were both found to have negative effects on native and non-native English-speaking listeners in comprehending English speech produced by native American-English-speaking talkers. Comprehension scores were adjusted for listeners' baseline English proficiency levels. In the present study, instead of native English-speaking talkers, two native Mandarin Chinese talkers (one male, one female) with similar English spoken proficiency were recruited to produce the same speech materials used in the previous study. A similar methodology was adopted to conduct speech comprehension tests on three groups of listeners: 1) native American-English, 2) native Mandarin Chinese and 3) other non-native English. The listeners' performance on foreign-accented speech comprehension is investigated under 15 acoustic conditions, created from five levels of RT (0.4 to 1.2 seconds) and three levels of BNL (RC-30, 40 and 50). Does a talker having a foreign accent change the results amongst the different listener groups in comprehending English speech under adverse acoustic conditions? Comparisons are made with the previous study, which showed that the negative impacts of RT and BNL varied between native and non-native listener groups.

Keywords: Background Noise, Reverberation, Classroom Acoustics, Speech Comprehension, Non-native, Foreign-accent

I-INCE Classification of Subjects Number(s): 63.3

1. INTRODUCTION

Design considerations for classroom acoustics may be complicated by the diverse linguistic backgrounds of the occupants. In the United States, the National Science Foundation (NSF) reported that in 2008 the percentage of full-time postsecondary teachers who were foreign-born with terminal doctoral degrees ranged from 19% in psychology to 54% in engineering (1). A considerable portion of this population are non-native English speakers, who regularly give lectures to an even more linguistically diverse student audience, as the U.S. remained the most popular destination for international students (2).

It has previously been found that non-native English speakers with later on-set of immersion in an English dominant environment are more likely to produce more heavily accented speech throughout their lifetime (3). Research had also shown that native English-speaking listeners do not perceive foreign-accented speech as accurately as speech produced by native speakers, particularly in more adverse listening conditions (4, 5). The perception of foreign-accented speech by non-native listeners had been less investigated. Bent and Bradlow (6) identified an interlanguage speech intelligibility benefit whereby it was easier for non-native listeners to perceive English sentences spoken by highly proficient non-native speakers, rather than by native English speakers.

In this study, 59 listeners belonging to one of three groups (native American English, native Mandarin Chinese or other non-native English speakers) were individually immersed in realistic acoustic environments to perform a dual-task involving speech comprehension. The background noise level and reverberation time were systematically varied to achieve a wide range of acoustic conditions.

[1] zpeng@huskers.unl.edu
[2] khanna@unomaha.edu
[3] bnboyd@unomaha.edu
[4] lwang4@unl.edu

Speech comprehension materials in English were recorded by two native Mandarin Chinese talkers with similar degree of accent.

Analyses of the resulting data aim at answering the following research questions.

1) What are the effects of background noise level and reverberation time in comprehending foreign-accented speech?
2) Does matched accent (talker and listener sharing the same accent) improve the comprehension of English as a foreign language?
3) How does foreign accent affect the impact of acoustics on speech comprehension?

2. METHODOLOGY

2.1 Participants

Recruitment of participants was completed through bulletin board postings on the University of Nebraska at Omaha campus. After giving consent to participate, each participant went through an initial screening session of two hours to determine their eligibility to proceed to the main experiment. During the initial screen, participants were given an audiometric screen, a demographic survey including the Language Experience and Proficiency Questionnaire (LEAP-Q; (7)), as well as three sets of English proficiency tests including listening span, oral comprehension and English verbal abilities.

All participants eligible to participate in the main experiment had normal hearing on both ears (≤25 dB hearing level from 125 Hz to 8000 Hz), and were identified based on the LEAP-Q as being in one of the three groups as listed below.

1) Native American English speaker (n = 20): English was identified as the first learned and currently dominant language
2) Native Mandarin Chinese speaker (n = 19): Mandarin Chinese was identified as the first learned and currently dominant language
3) Other non-native English speaker (n = 20): Neither English nor Mandarin Chinese was identified as the first learned or currently dominant language

A total of 59 participants completed the study. All non-native English-speaking participants had lived in the US for at least one month prior to their participation in the study.

2.2 Production of Foreign-accented Speech

Two native Mandarin Chinese speakers (one male and one female) were recruited to record speech comprehension materials developed from practice materials for the Test of English for International Communications (TOEIC). They both had normal hearing and no known speech impairment. The two talkers were chosen based on their similar scores from the Versant spoken English test as shown in Table 1. Speech comprehension materials were recorded individually with these two talkers in a sound attenuated booth using a close microphone recording technique.

Table 1 – Versant spoken English test score and estimated percentile ranking among typical non-native English speakers

	Fluency		Pronunciation	
	Versant Score (T-Score)	Percentile Ranking	Versant Score (T-Score)	Percentile Ranking
Male Chinese Talker	58	81%	53	63%
Female Chinese Talker	55	75%	52	55%

2.3 Acoustic Conditions

The current ANSI S12.60 standard (8) specifies design criteria for two acoustic metrics: reverberation time (RT) and background noise level (BNL). For this study, a total of 15 acoustic conditions were created by combining five levels of RT (0.37, 0.62, 0.84, 1.05, and 1.19 seconds; averaged from 500 Hz to 2000 Hz) and three BNLs (RC-30, 40, and 50).

The RT conditions were achieved by simulating a typical classroom of 260 m^3 in the auralization program ODEON, and systematically changing the absorption coefficients on the ceiling and side and back walls. Binaural room impulse responses (BRIR) were exported which incorporated the loudspeaker-listener orientation for playback in a low-reverberant listening chamber. The BRIRs were then digitally convolved in MATLAB with the accented speech to embed the RT conditions.

Background noise was supplied via an overhead loudspeaker mounted in a ceiling tile and a corner subwoofer, independent from the playback of speech materials. The BNLs were created by digitally adjusting the steady-state sound pressure levels to match that of the RC-30, 40 and 50 contours at ear height at the fixed listener position.

2.4 Procedure

After passing the initial screen, participants were invited back for six one-hour sessions to conduct the main experiment individually for all 15 acoustic conditions (3 BNL X 5 RT) in the listening chamber. During each hourly session, the BNL was kept consistent. The BNLs were nested in six blocks, which were counterbalanced for presentation across six main experiment sessions. The RT conditions were counterbalanced within each BNL. For each hourly session, participants completed three 15-minute tests that corresponded to three RT conditions under the same BNL. Fifteen sets of speech comprehension materials recorded by the Mandarin Chinese talkers were randomly assigned to the acoustic combinations, with three additional practice tests given at the beginning of each BNL.

While being immersed in each acoustic condition over a 15-minute period, participants were asked to simultaneously attend to two tasks: speech comprehension and adaptive pursuit rotor (APR). An illustration of the testing sequence is provided in Figure 1. For the speech comprehension test, participants completed four tasks based on their understanding of the accented speech. The four tasks are 1) identifying a photograph, 2) responding to an oral question, 3) understanding a simple conversation, and 4) understanding a short monologue or paragraph. All questions were in multiple choice format with visual prompts provided on a computer monitor screen in front of the participant. While attending to the speech comprehension tasks, participants simultaneously performed the APR task by tracing a dot that continuously rotated around a fixed circle on a second monitor screen using a stylus and pad.

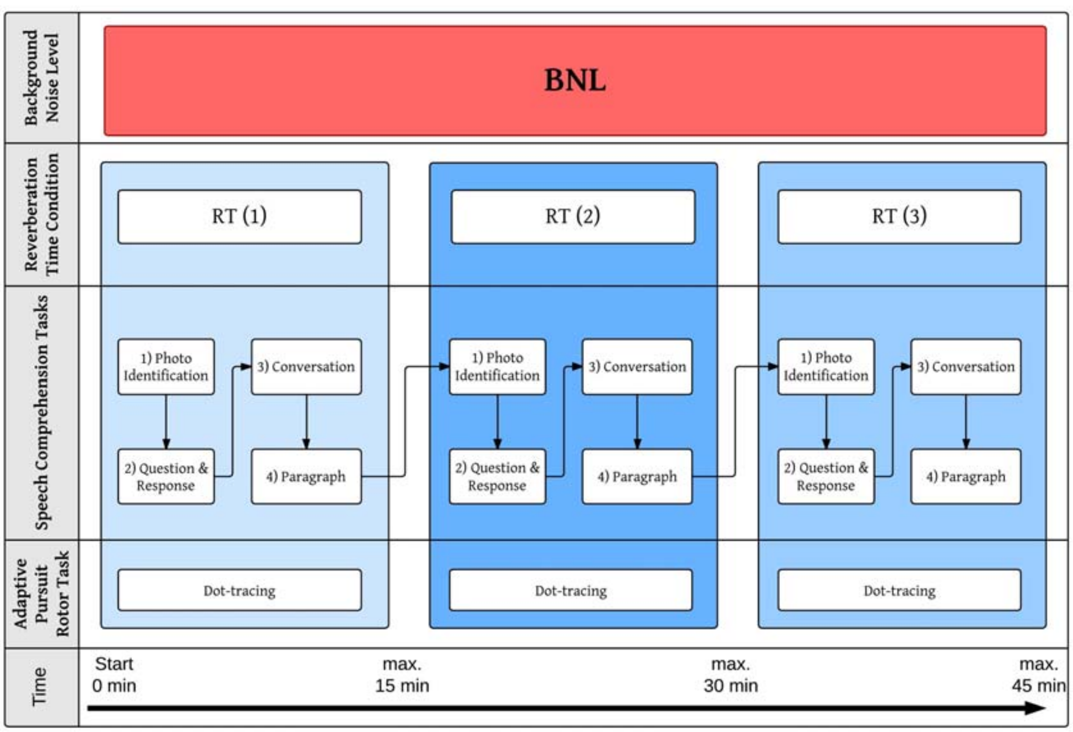

Figure 1 – Testing process of acoustic conditions and simultaneous competing tasks of speech comprehension and adaptive pursuit rotor during each one-hour main experiment session.

3. RESULTS

As gathered from the initial screen, participants in this study exhibited a wide range of English proficiency levels with clear clustering as identified in Figure 2. As expected, the native American English (NAE) listeners who are highly proficient in their native language were concentrated in the upper 16th percentile (one standard deviation). The native Mandarin Chinese (NNC) listeners participated in this study seemed to be mostly lower proficient English speakers, centered on the lower 16th percentile among all participants. The other non-native English-speaking (NNO) listeners as a group scored slightly higher on English proficiency level than the NNC listener group.

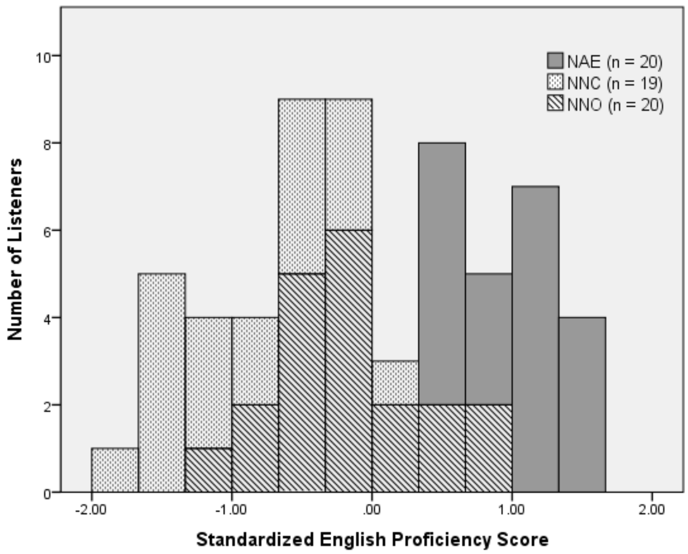

Figure 2 – Histogram of English proficiency levels for three groups of listeners: native American English (NAE), native Mandarin Chinese (NNC) and other non-native English (NNO).

3.1 Effects of Acoustics in Foreign-Accented Speech Comprehension

Prior to analyses, the speech comprehension scores were converted from percent correct to rationalized arcsine unit (RAU) using the transformation proposed by Studebaker (9, 10). The speech comprehension RAU scores were then subjected to a mixed-design analysis of covariance (ANCOVA); with BNL and RT as the within-subject variables, listener groups as between subject variable, and English proficiency level as the covariate.

Results of ANCOVA revealed only significant main effects for BNL ($F[2, 110] = 121.47$, $p < .001$, $\eta^2 = .69$), RT ($F[4, 220] = 451.42$, $p < .001$, $\eta^2 = .10$), listener group ($F[2, 55] = 16.47$, $p < .001$, $\eta^2 = .38$), and English proficiency level ($F[1, 55] = 28.32$, $p < .001$, $\eta^2 = .34$). No significant interactions were found. The marginal means of speech comprehension performance are provided in Figures 3 and 4. Specifically, significant linear trends were observed through planned comparisons for both BNL ($F[1, 55] = 192.87$, $p < .001$, $\eta^2 = .78$) and RT ($F[1, 55] = 23.33$, $p < .001$, $\eta^2 = .30$).

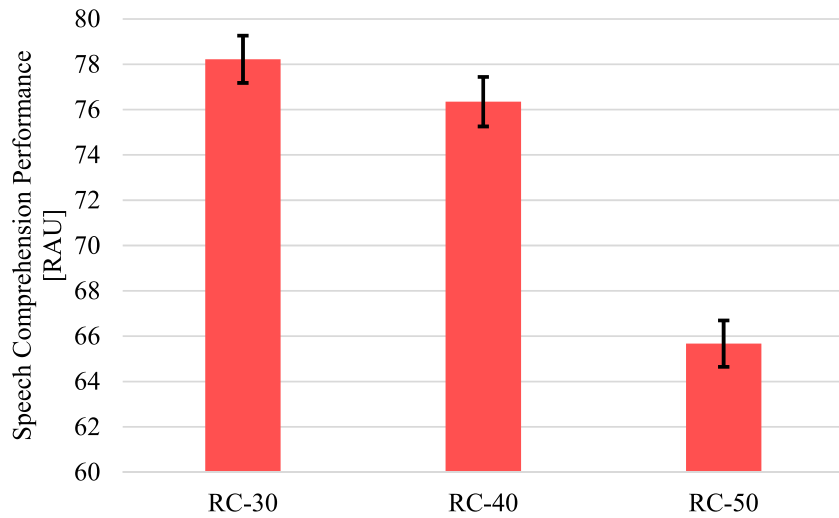

Figure 3 – Marginal means (averaged across five RT conditions) of speech comprehension performance in RAU for three BNLs, adjusted for English proficiency level at 0. Error bars indicate one standard deviation.

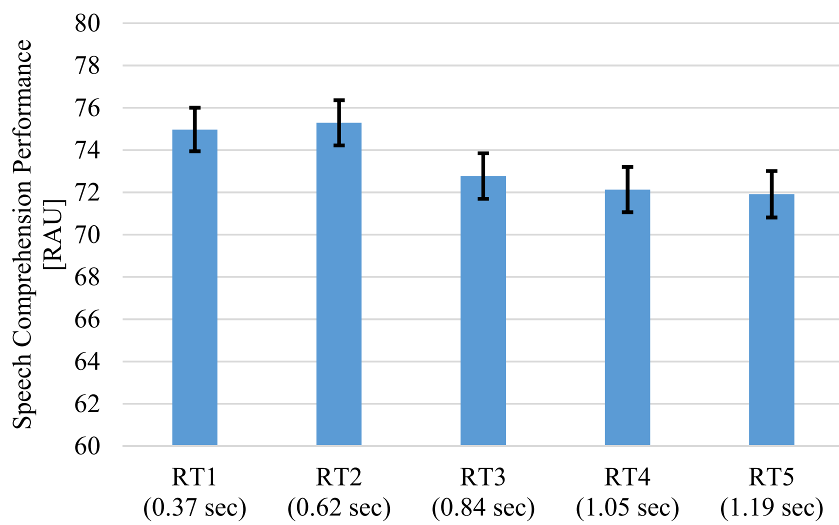

Figure 4 – Marginal means (averaged across three BNLs) of speech comprehension performance in RAU for five RTs, adjusted for English proficiency level at 0. Error bars indicate one standard deviation.

3.2 Benefit of Matched Accent in Speech Comprehension

With English proficiency adjusted to the same level, pairwise comparisons using Bonferroni correction were conducted among listener groups to reveal the benefit of matched accent in speech comprehension by listeners. When averaged across all acoustic conditions, as seen in Figure 5, native Mandarin Chinese listeners' performance on speech comprehension tasks was significantly higher than both native American English ($p = .009$) and other non-native English speakers ($p < .001$). However, native American English and other non-native English listeners did not differ significantly on their speech comprehension performance ($p > .05$).

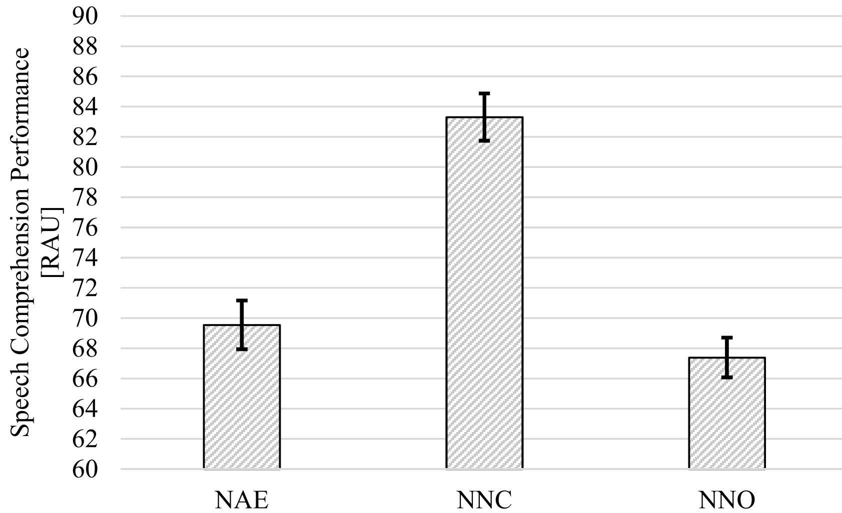

Figure 5 – Speech comprehension performance, averaged across all acoustic conditions, among three listener groups. Performance score is adjusted for English proficiency level at 0. Error bars indicate one standard deviation.

3.3 Impact of Foreign Accent on the Acoustics Effects on Speech Comprehension

In a previous study by the authors (11), similar methodology was utilized but with speech comprehension materials recorded by native American English-speaking talkers instead. (Listeners participated in the previous study were not eligible to participate in the main experiment of the present study.) The effect sizes in η^2 for the main effects and interaction (only between background noise level and English proficiency level) from two studies are compared in Table 2.

Table 2. Effect sizes of the acoustic main effects and interaction between two experiments using native American English (NAE) versus native Mandarin Chinese (NNC) talkers

	American English Accent N = 56		Mandarin Chinese Accent N = 59	
	Significance Level	Effect Size η^2	Significance Level	Effect Size η^2
Main Effect				
English Proficiency Level	p < .001	0.56	p < .001	0.26
Background Noise Level	p < .001	0.40	p < .001	0.68
Reverberation Time	p = .006	0.07	p < .001	0.10
Interaction				
BNL X English Proficiency	p = .004	0.10	p > .93	0.001

The general trend of significance levels calculated for the main effects were similar between two studies (NAE talkers vs. NNC talkers). It was previously reported that the effects of both BNL and RT were more pronounced (larger effect size) for non-native listeners (NNC and NNO, n = 29) than for NAE listeners (n = 27), if the speech was produced by native American English talkers. With more non-native English-speaking listeners (NNC and NNO, n = 39) than native listener (n = 20) in the present study using native Mandarin Chinese talkers, the main effect size increased slightly for RT but drastically for BNL from η^2 = 0.40 to 0.68. However, there is a large reduction in the effect size for the main effect of English proficiency level and the interaction of BNL X English proficiency level.

4. CONCLUSION & DISCUSSION

The present study investigated the effects of background noise level and reverberation time on the comprehension of English speech from native Mandarin Chinese speakers by native and non-native English-speaking listeners. Three groups of listeners were examined, including native American English, native Mandarin Chinese and other non-native English speakers whose native language was neither English nor Chinese. The results suggest that in comprehending foreign-accented speech, similar to perceiving non-foreign-accented speech, listeners' performance deteriorates with increasing background noise level or reverberation time.

The two native Mandarin Chinese talkers shared a similar degree of accent, which was considered moderate among typical non-native English speakers determined through the Versant spoken English test. Despite only being moderately accented, the effect of Mandarin Chinese accent was prominent. Surprisingly, when the accented speech was presented under an assortment of acoustic conditions, native American English listeners (NAE to NNC) did not comprehend significantly better than non-native English listeners (NNO to NNC), who were perceiving the foreign language with mismatched accent. The projected benefit of matched accent was tremendous in English foreign speech comprehension, particularly for those with lower English proficiency level. The results predicted that the native Mandarin Chinese listeners can surpass both native American English and other non-native English-speaking listeners on comprehension performance, if they can slightly improve English proficiency.

By comparing the effect sizes, the negative effects of background noise level and reverberation time became more severe when the English speech was produced by native Mandarin Chinese talkers than by native American English talkers. The increase in effect size was much more pronounced for background noise level, promoting it from medium ($\eta^2 = 0.40$) to large ($\eta^2 = 0.68$). However, the strength of English proficiency as a confounding factor in speech comprehension performance under various acoustic conditions weakened drastically. This was perhaps due to the improved performance in speech comprehension among native Mandarin Chinese listeners, who benefited from matched accent even though they were of lower English proficiency levels.

ACKNOWLEDGEMENT

The Paul S. Veneklasen Research Foundation is acknowledged for providing generous financial support for this project. The authors would also like to thank Dr. Thomas Carrell and Dr. Nirmal Srinivasan for their input on recreating the adaptive pursuit rotor task.

REFERENCES

1. National Science Board. Science and Engineering Indicators 2012. Arlington VA: National Science Foundation (NSB 12-01). 2012.
2. Institute for International Education. Open doors 2012 fast facts. 2012
3. Flege JE, Yeni-Komshian GH, Liu S. Age constraints on second-language acquisition. Journal of memory and language. 1999;41(1):78-104.
4. Munro MJ. The effects of noise on the intelligibility of foreign-accented speech. Studies in Second Language Acquisition. 1998;20(2):139-54.
5. Rogers CL, Dalby J, Nishi K. Effects of noise and proficiency on intelligibility of Chinese-accented English. Language and speech. 2004;47(2):139-54.
6. Bent T, Bradlow AR. The interlanguage speech intelligibility benefit. J Acoust Soc Am. 2003;114(3):1600.
7. Marian V, Blumenfeld HK, Kaushanskaya M. The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. Journal of Speech, Language and Hearing Research. 2007;50(4):940.
8. American National Standard Institute. ANSI/ASA S12.60 Performance Criteria, Design Requirements, and Guidelines for Schools, Part 1: Permanent Schools. 2010.
9. Sherbecoe RL, Studebaker GA. Supplementary formulas and tables for calculating and interconverting speech recognition scores in transformed arcsine units. International Journal of Audiology. 2004;43(8):442-448.
10. Studebaker GA. A "rationalized" arcsine transform. Journal of Speech, Language and Hearing Research. 1985;28(3):455.
11. Peng Z, Steinbach AM, Hanna KE, Wang LM. How room acoustics impact speech comprehension by listeners with varying English proficiency levels. J Acoust Soc Am. 2013;134(5):3970.