

Roberson, A. (2012). Analyzing comprehensibility among on-native speakers of English: The effect of listener first language background. In J. Levis & K. LeVelle (Eds.). *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2011. (pp. 240-248). Ames, IA: Iowa State University.

ANALYZING COMPREHENSIBILITY AMONG NON-NATIVE SPEAKERS OF ENGLISH: THE EFFECT OF LISTENER FIRST LANGUAGE BACKGROUND

[Audrey Roberson](#), Georgia State University

The evolution of English from a native speaker centered language to a means of communication among speakers of various first language (L1) backgrounds motivates the empirical examination of communication among non-native speakers (NNS) of English. This paper describes a study where NNS of various L1 backgrounds, as well as a native speaker, rated NNS speech samples for comprehensibility. Analysis of variance (ANOVA) and descriptive statistics were used to address the following questions: In comprehensibility ratings of NNS speech, do NNS listeners vary: (1) by L1 background? and (2) when compared with a NS listener trained in assessment? Results showed a significant effect for listener L1 background in comprehensibility ratings of several speakers. This finding contradicts previous research suggesting that the quality of an NNS utterance itself is more influential in determining comprehensibility than are any listener characteristics (Hazan & Markham, 2004; Munro et al., 2006). However, it may be an example of matched interlanguage intelligibility benefit, where NNS listeners of the same L1 background as the speaker find speech easier to understand than their non-matched peers (Bent & Bradlow, 2003). It seems that further investigation is needed into the complicated relationship between L1 background and comprehensibility among NNS of English.

INTRODUCTION

The extent of the global spread of English becomes apparent when one examines statistics about its use. It has been estimated that there are 320 to 380 million speakers of English as a first language, but over one billion users of English overall (Crystal, 2003). These numbers illustrate that second language speakers of English far outnumber those who speak English as a first language. In fact, it is predicted that in the next ten to fifteen years, there will be two billion English language learners worldwide (Graddol, 2006).

In this context, empirical investigations have been made to determine speaker and listener features that make NNS-NNS communications successful or less so. The results of these studies have important implications for university ESL contexts, where the wide reach of English and the predominant use of communicative language teaching methods (see, e.g. Pica et al., 1993) mean that students of various L1 backgrounds are expected to learn English by speaking with each other. However, research on the role of L1 background in comprehensibility among NNS of English has produced mixed results to date, leaving instructors with little guidance on how to facilitate learning in NNS-NNS interactions. While some studies have found that listeners share a response to non-native speech regardless of their L1 (Munro et al., 2006), others have suggested that a listener who shares the speaker's L1 may benefit from this match (Bent & Bradlow, 2003). Still other studies have concluded that such a match is not beneficial, and may even be a detriment

for the listener (Major et al, 2002). Further investigation is needed to unravel the intricate relationship between L1 background and successful oral communication among NNS of English.

RESEARCH QUESTIONS

The previously mentioned empirical studies provide background in effective research methods for assessing speech samples. However, further investigation is needed into the effect of listener first language background on the perception of comprehensibility of non-native speech. In addition, it is important to examine how non-native speaker listeners differ in their ratings of comprehensibility from a native speaker listener who is trained in ESL. Such data will provide direction for ESL teachers and materials creators in developing pronunciation syllabi and communication-based classroom tasks. It may also have implications for oral proficiency assessment, addressing the question of the ability of NNS raters to assess speech in a similar way to that of NS ones. As such, this study addresses the following questions:

1. In comprehensibility ratings of NNS speech, do NNS listeners vary by L1 background?
2. In comprehensibility ratings of NNS speech, do NNS listeners vary when compared with a NS listener trained in assessment?

RESEARCH DESIGN

Participants

The NNS listeners in this study (n=18) were members of an intact English for Academic Purposes (EAP) course taught by the researcher. The course is a required component of an Intensive English Program (IEP) at a large research university in the southeastern United States. This IEP is organized into skill-area courses (i.e., oral communication, academic writing) designed to prepare students for university work. The program is comprised of five levels, ranging from “high beginning” (Level 1) to “advanced” (Level 5). According to in-house proficiency testing and placement procedures, the participants in this study were classified as Level 4, or “high intermediate” English learners. These higher-level students were targeted in order to attempt to control for proficiency level interference in comprehensibility ratings. Any difficulties in comprehensibility can hopefully be attributed either to qualities of the utterance or to the listener’s language background, rather than to the listener’s difficulty with syntactic or lexical items that might have occurred at lower proficiency levels. Table 1 provides further description of NNS rater characteristics.

Table 1
Description of NNS Listeners (n=18)

L1	Age (years)	Gender	Exposure to accented English (1= very little; 2= some; 3= a lot)	Time studying English (years)
Vietnamese (n=4)	26.1 (m)	Male (n=8)	2.56 (m)	3.9 (m)
Spanish (n=3)	7.8 (STD)	Female (n=10)	0.11 (STD)	2.9 (STD)
Chinese (n=3)				
Korean (n=3)				
French (n=2)				
Arabic (n=2)				
Japanese (n=1)				

The native speaker rater (n=1) is a PhD student in Applied Linguistics who has a professional background in assessment of non-native speakers. At the time of the study she was the testing coordinator for the university's ESL program. As such, her duties included administering and rating oral proficiency exams for incoming university students to determine their need for ESL services. Given her background and training, she had been frequently exposed to accented English and was adept at assessing it for comprehensibility.

Speech Samples

The sound files played for NNS raters were collected for the listening and speaking portion of an internal university ESL proficiency test developed by the university's Department of Applied Linguistics. This proficiency test is designed to assess the reading, writing, speaking and listening skills of international students seeking enrollment in university programs, and consists of various sections designed to target these skills. For this study, free speech samples from the oral proficiency interview portion of the exam were played for NNS listeners. Each of the seven sound files was approximately one minute long. In these sound files, respondents were answering the interviewer's question: "Can you tell me a little bit about what you're studying?" The NS rating used for the purposes of this study is the interviewer's rating of the overall oral proficiency interview, which also included warm-up questions and a reading passage that were not played for NNS listeners.

Data collection procedures

Seven 1-minute digital recordings from the interview portion of the oral proficiency exam were obtained with permission from the university. Descriptive statistics for the speakers recorded in these segments are listed in Table 2.

Table 2
Description of NNS Speakers (n=7)

L1	Age (years)	Gender	Academic Major
Chinese(n=5)	24.7 (m)	Male (n=3)	Communications (n=1)
Indian(n=1)		Female (n=4)	Economics (n=1)
Korean (n=1)			Chemistry (n=2)
			Computer Science (n=1)
			Music (n=1)
			Statistics (n=1)

Recordings were played for the 18 NNS raters by the researcher as an in-class listening activity. Before completing the listening activity, informed consent was obtained from each class member. NNS raters were asked to complete a language background / biographical data questionnaire (Appendix A) and then to rate each speech sample on a 7-point qualitative scale (Appendix B) for comprehensibility, or the listener's estimation of his or her difficulty in understanding an utterance (Munro et al., 2006). The native speaker rating had been previously assigned during oral proficiency interviews with the seven interviewees who produced the speech samples.

RESULTS AND DISCUSSION

The Effect of Listener's L1 Background on Comprehensibility Ratings

Statistical analysis.

In order to address research question one, the overall rating means for four different language groups were compared using One-Way Analysis of Variance (ANOVA) calculated in SPSS version 16.0. Although seven L1 backgrounds were represented among the eighteen NNS raters, only four (Chinese, Spanish, Vietnamese, and Korean) were chosen for analysis. This decision was informed by Johnson (2008), which suggests that each group in an ANOVA analysis should have a minimum of three observations to ensure robustness of hypothesis testing. Ratings from French (n=2), Arabic (n=2), and Japanese (n=1) speakers were not included in the ANOVA analysis for research question one because there was not a sufficient number of participants in these language groups.

A non-parametric equivalency test for differences in means, the Kruskal-Wallis test, was initially considered given the small number of participants in each group. However, the four groups chosen for analysis were also analyzed for equality of variances, which is another of the required assumptions of an ANOVA analysis (Johnson, 2008). This analysis showed that the equality of variances assumption was met ($Levene > .05$). Given this assumption of normal data distribution, ANOVA was used despite the limited number of participants in each group. Table 3 shows results of the ANOVA calculation.

Table 3
ANOVA Summary Table for Comprehensibility Scores by L1 Background (Chinese, Spanish, Vietnamese, Korean)

Source	SS	Df	MS	F
Between	37.29	3	12.43	6.211*
Within	174.10	87	2.00	
Total	211.39	90		

* $p = .001$

Because the initial ANOVA analysis showed significant differences among the four groups for two of the seven speech samples that were rated, a post-hoc Tukey HSD analysis was conducted to further examine these differences. This analysis revealed that overall Spanish speaker ratings ($m=4.14$) were significantly lower (the speech was interpreted as less comprehensible) at a significance level of $p<.01$ than were Vietnamese ($m=5.72$), Korean ($m=5.67$), and Chinese ($m=5.57$) speaker ratings. There were no significant differences, however, among the speakers of the Asian languages. Furthermore, Spanish listeners rated speaker #2 ($m=3.67$) and speaker #5 ($m=3.0$) as significantly less comprehensible than did Chinese or Vietnamese listeners ($p<.01$). Both of these low-rated speakers were Chinese.

Discussion.

It appears that for certain speech samples, listener first language background was influential in determining the comprehensibility rating. These results seem contradictory to those of Munro et al. (2006), which found that listeners of various language backgrounds generally shared a response to L2 speech when they evaluated it for intelligibility, accentedness, and comprehensibility. In the current study, in both cases where differences among the four rating groups were significant, Spanish listeners rated Chinese speakers as less comprehensible than did Chinese or Vietnamese listeners, respectively. It would appear, then, that Chinese and Vietnamese listeners experienced a comprehensibility benefit when listening to the accented English of Chinese speakers. This finding may be further support for the matched interlanguage intelligibility benefit posited by Bent and Bradlow (2003).

Although it was not possible given the scope of this study, it would have been helpful to gather qualitative data from the Spanish-speaking listeners about their perception of English spoken with an accent other than their own L1. Interestingly, based on the researcher's observations from teaching this class, the Spanish speakers were among the most highly proficient speakers in the class. Although they had little trouble speaking fluently, this data suggests they had more trouble understanding Chinese speakers' English than did their Chinese or Vietnamese speaking classmates. Thus, the Spanish speakers' own insights about their ability to understand speakers of various language backgrounds would be helpful. This additional information about the listener's attitude toward the speaker as a non-native user of the language (see, e.g., Hu & Lindemann, 2009; Lindemann 2002; Lindemann, 2005; Lippi-Green, 2001) might paint a more comprehensive picture of why Spanish speakers assigned lower ratings than their classmates.

If further research continues to suggest that L1 background does have an effect on the comprehensibility ratings of NNS speech, more investigation will be needed into exactly how L1 background interacts with comprehensibility ratings of different types of L2 accented speech. More data about the listeners and speakers and more analysis of the speech itself might further reveal the interaction between speaker and listener. For example, studies might be conducted with listeners who vary in their L1 background, length of stay in target language community, academic background, and attitudes toward the speaker. The speech itself might be analyzed for lexical, syntactic, and phonological similarity to the target language. This growing body of research analyzing different

aspects of listener and speaker characteristics might result in suggestions about how best to support students for successful mutual comprehension in communicative language activities.

Native Speaker Versus Non-native Speaker Comprehensibility Ratings

Statistical analysis.

In order to examine research question two, mean rating scores for all non-native raters (n=18) and the native speaker rater (n=1) were compared. Because of the disparate group sizes, inferential statistics were not used. Descriptive statistics for average NNS/NS ratings for all speakers are presented in table 4.

Table 4

Means and Standard Deviations of Comprehensibility Scores by NS/NNS

L1 Background	N	Mean	Standard Deviation
English	1	6.0	0.52
All NNS	18	5.36	0.99
Spanish	3	4.14	0.71
Chinese	3	5.57	0.29
Vietnamese	4	5.71	0.61
Korean	3	5.67	1.14

Discussion.

This second research question provides insight into how the ability of language learners to understand other learners compares to the ability of native speakers to perform the same task. In addition, there is another factor in this analysis: the NS rater is an expert in this type of assessment, while the NS raters were not. Previous research in this area has suggested that non-native speakers rate speech samples as more comprehensible than do native speakers (Bent & Bradlow, 2003), which may mean that there are traits of NNS interlanguage that the non-native listeners share but that the NS rater may not understand. However, the analysis for this study suggests opposite trends: the NS rater found non-native speech to be more comprehensible than did the NNS ones.

Although caution should be used in interpreting the results of this exploratory study, higher NS ratings may signal an area that merits further investigation. In this study, the NS rater may be a more sympathetic listener than are the NNS raters. That is to say, she was aware that the students she was speaking with were nervous given the context of an oral proficiency interview, and she is accustomed to interpreting non-native speech. However, her overall goal in assigning a rating was to determine whether or not the interviewee was in need of ESL services at the university. For a few speakers, the NS rating meant that the interviewee was deemed comprehensible enough so as not to need ESL course work. However, these same speakers were rated much lower by NNS listeners, who assigned a score that would have placed the student in a pronunciation course. This possible disparity between assessment outcomes and classroom demands warrants further attention in comprehensibility research. Such investigations may also be of interest to the assessment community, in which there has been debate about whether or

not non-native speakers are appropriate candidates for assessing English oral proficiency examinations (see, e.g., Kim, 2009).

LIMITATIONS AND CONCLUSION

The one minute clips used in this study may not have been ideal. The rating assigned by the NS listener also took into account additional portions of the oral interview: warm-up questions and a reading passage. Scholars in pronunciation teaching and assessment have suggested that in order to obtain a global picture of the speaking and listening skills of test takers, both a free speech sample and a diagnostic passage should be obtained (see, e.g., Celce-Murcia, 1996). NNS raters, however, assigned their comprehensibility score based only on the one-minute free speech clip. Therefore, caution should be used when comparing ratings across these two groups of listeners. Also, playing sound files from an interview in an experimental setting may detach the interview from what researchers have called its context of use (Field, 2003; Jenkins, 2000). In future studies, the social context of an interview may be more closely replicated for raters if they are able to see a video of the interview in addition to listening to it.

The findings of this study suggest that there may be an L1 effect for listener ratings of comprehensibility of L2 speech. Spanish listeners rated speech from Chinese speakers as significantly less comprehensible than did Chinese, Vietnamese, and Korean raters. These results imply that there may be features of the interlanguage of the speech samples that the raters who speak Asian languages share, but that the Spanish speaking raters do not. However, more raters and speakers of these languages and others should be analyzed to confirm that this difference is actually related to L1 influence. In addition, studies of authentic contexts may be appropriate for addressing comprehensibility, given the relative nature of this construct (Pickering, 2006). Such studies would shed light on the complicated relationship among speakers, listeners, and the unique environments in which they communicate.

ABOUT THE AUTHOR

Audrey Roberson is a PhD student in Applied Linguistics at Georgia State University in Atlanta, Georgia, where she teaches ESL and linguistics courses. Her research interests include comprehensibility in oral communication, as well as connections between oracy and literacy in second language learners. **Contact Information:** Audrey Roberson PhD Student, Department of Applied Linguistics and ESL, Georgia State University, 34 Peachtree St. Suite 1200, One Park Tower Building, Atlanta, GA 30303, (404) 413-5197, aroberson10@gsu.edu

REFERENCES

- Bent, T. & Bradlow, A.R. (2003). The interlanguage and speech intelligibility benefit. *Journal of the Acoustical Society of America*, 114, 1600-1610.
- Celce-Murcia, M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.

- Crystal, D. (2003). *English as a global language*. Cambridge: Cambridge University Press.
- Field, J. (2003). The fuzzy notion of ‘intelligibility’: A headache for pronunciation teachers and oral testers. *IATEFL Special Interest Groups Newsletter*, 34-38.
- Graddol, D. (2006). *English next*. Plymouth, UK: British Council.
- Hazan, V., & Markham, D. (2004). Acoustic-phonetic correlates of talker intelligibility for adults and children. *Journal of the Acoustical Society of America*, 116, 3108-3118.
- Hu, L. & Lindemann, S. (2009). Stereotypes of Cantonese English, apparent native/non-native status, and their effect on non-native English speakers’ perception. *Journal of Multilingual and Multicultural Development*, 30 (3), 253-269.
- Jenkins, J. (2002). *The phonology of English as an International Language*. Oxford: Oxford University Press.
- Johnson, K. (2008). *Quantitative methods in linguistics*. Malden, MA: Blackwell Publishing.
- Kim, Y. (2009). A G-theory analysis of rater effect in ESL speaking assessment. *Applied Linguistics*, 30(3), 435-440.
- Lindemann, S. (2002). Listening with an attitude: A model of native speaker comprehension of non-native speakers in the United States. *Language in Society*, 31(1): 419-441.
- Lindemann, S. (2005). Who speaks “broken English”? US undergraduates’ perceptions of non-native English. *International Journal of Applied Linguistics*, 15(2), 187-212.
- Lippi-Green, R. (2001). *English with an accent: Language, ideology, and discrimination in the United States*. London: Routledge.
- Major, R., Fitzmaurice, S., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly* 36, 173-190.
- Munro, M, Derwing, T. & Morton, S. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, 28, 111-131.
- Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language instruction and research. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 9-34), Clevedon, UK: Multilingual Matters.
- Pickering, L. (2006). Current research on intelligibility in English as a lingua franca. *Annual Review of Applied Linguistics*, 26, 219-233.

Appendix A: Questionnaire for NNS Raters

Please answer the following questions. Please do not write your name.

Participant # _____ (to be completed by researcher)

1. What is your age? _____ years
2. What is your first language? _____
3. How long have you been studying English? _____ years _____ months
4. Please rate your amount of exposure to English spoken with a non-native accent:
 no exposure a little exposure a lot of exposure

Appendix B: NNS Listener Rating Sheet

Participant # _____ (to be completed by researcher)

You will be rating the following speech samples for their **comprehensibility: how easy or difficult is it for you to understand the speaker?**

Listen to each speech sample (1-7). Assign each one a number (1-7) based on **how comprehensible it is to you**. A score of 1 is not comprehensible at all. A score of 7 is completely comprehensible.

1	2	3	4	5	6	7
not at						completely
all comprehensible						comprehensible

- | | |
|---------------------------------|---------------|
| Speaker # 1 score (circle one): | 1 2 3 4 5 6 7 |
| Speaker # 2 score (circle one): | 1 2 3 4 5 6 7 |
| Speaker # 3 score (circle one): | 1 2 3 4 5 6 7 |
| Speaker # 4 score (circle one): | 1 2 3 4 5 6 7 |
| Speaker # 5 score (circle one): | 1 2 3 4 5 6 7 |
| Speaker # 6 score (circle one): | 1 2 3 4 5 6 7 |
| Speaker # 7 score (circle one): | 1 2 3 4 5 6 7 |