The effect of input source (Native vs. Non- native) and EFL learners' perceptions towards it, on their listening performances, across gender

Sadeghi, Karim
Urmia University, Iran (ksadeghi03@gmail.com)

Zeinali, Maryam ✉
Urmia University, Iran (m.zeinali1987@gmail.com)

Abstract

The issue of the "native/non-native speaker's input source" occupies an important place in any EFL listening comprehension test. This study is meant to unveil whether listening to a native or non-native speaker (i.e., input source) and students' perceptions towards it affect the performance of upper-intermediate EFL learners in a listening test. For this purpose, an experimental design was used to compare the performance of two groups of learners on an EFL listening test. A test of 20 multiple choice items was administered to 66 EFL learners (31 male and 35 female), half of whom listened to a native speaker's voice while the other 33 test takers listened to a non-native speaker's voice. Moreover, a perception questionnaire considering students' perceptions towards using native or non-native input source in listening tests was utilized. The results of the study highlighted that the overall performance of the two groups differed significantly. That is, the listeners to the non-native input outperformed those who listened to a native speaker. Also, the results of the questionnaire revealed some noteworthy findings which indicate that students preferred the use of non-native input in listening tests. Moreover, considering gender as a moderator variable, a statistically significant main effect was found for gender; that is, gender does play a significant role as a moderator variable. In other words, female test takers performed better than males in both conditions. Further findings and implications are discussed in the paper.

Keywords: EFL learners; input source; listening comprehension test; native input; non-native input

The effect of input source (Native vs. Non- native) and EFL learners' perceptions towards it, on their listening performances, across gender

1. Introduction

Testing is an important part of every teaching and learning experience. It is one of the ways in which information is gathered about people's language ability. Tests of some kind or another are needed for individuals to be given a statement of what they have achieved in a second or foreign language. They are needed in order to bestow information about the achievement of groups of learners on teachers and without them it is difficult to make logical educational decisions (Hughes, 2003). Considering the language skills, listening comprehension is one of the most fundamental and also the most complicated one (Farhady, Jafarpur, & Birjandi, 2007). As Thrasher (2000) states, listening is a skill that is very difficult to break up into its constituent parts; it is at the bottom, getting meaning from spoken language, and all the sub skills that make up listening are all subservient to this quest for meaning.

Much of the current thinking on the nature of testing listening comprehension is based on earlier work on testing reading comprehension. Because both are receptive skills, the testing of listening parallels in most ways the testing of reading, because of their invisible, cognitive nature, which makes them difficult to describe and to assess (Weir, 1990). But it should be acknowledged that listeners listen to spoken language which is very different from written language since speaking is linear and takes place in real time, in the sense that the text is heard only once, and then it is gone, with no chance of review (Buck, 2007). One important reason to test listening is to encourage teachers to teach it (Buck, 2007).

When a test-developer sets out to design a listening comprehension test, he usually encounters numerous factors that may influence test-takers' performance, such as question format, speech rates, speaker accents, topic familiarity, etc. One of the factors that might affect the test takers' listening performances is the input source; that is, whether the listening text comes from a native speaker or a non-native speaker. While the link between some of the above variables and performance in listening, tests has attracted the attention of testing scholars, investigating the effect of input source on listening performance is relatively new to the field and the relationship does not seem to have been investigated in depth up to the present time. Many studies share the idea that listening to a native speaker can lead an EFL learner to gain higher fluency in speaking since native speakers provide a real context for the learners. But the question here is: 'Do EFL listeners to a native speaker get the same results as the EFL listeners to a non-native speaker do in a test condition?' The answer to such a question is important since one of the main differences in listening comprehension texts is the speaker-related features, such as stress, intonation, accent, pauses, text speed, etc. (Weir, 1990).

By the growing population of non-native users of English in the world which far outnumbers that of native speakers (Crystal, 1997; Seidlhofer, 2001; Yano, 2001), the ownership of native speaker norms in language tests, especially in listening sections, is under question. That is, to limit test input to a native speaker in listening comprehension tests such as TOEFL and IELTS is a misrepresentation of testing English academically. As Backman and Palmer (1996) state, because of the existence of different varieties of English (e.g., English spoken with accents) listening comprehension tests whose input is limited only to native speaker are not considered as authentic ones.

1.1 Background

Research output on the relationship between written production or even the oral production of the learners who work with a native speaker is scant. For example, Zhang and Elder (2011) in their study report that native and non-native English speaking teachers weigh various features of oral proficiency in making their decisions on

the language proficiency of test-takers. Furthermore, in a research study conducted by Bent and Bradlow (2003), perception tests were conducted with some non-native and native speakers. There were Chinese, Korean, and English native speaker and mixed-nationality groups listening to speakers with Chinese, Korean and English native speaker accents reading sentences in English. The results revealed that native listeners found native speakers of English the most intelligible; however, for each of the non-native listener groups, non-native speaker with the same L1 background was considered the most intelligible. This was entitled ''matched interlanguage intelligibility benefit'' by the researchers. This study is accordingly intended to add to the current fuzzy literature on the role of input source on the listening performance of EFL learners.

Munro, Derwing, and Morton (2006) in their study, concluded that a shared-L1 may have little if any effect on the level of intelligibility and comprehensibility of a listening input. The results of their study revealed that interaction effect between the speaker and listeners was significant among listeners of different L1, that is, there is only weak evidence of a shared-L1 effect. Students' perceptions towards native or non-native teachers/input might be affected by several factors (Mahboob, 2004). He asserts that students favor native speaker teachers for their cultural knowledge and oral skills; however, non-native teachers are preferred for their structural grammar lessons and responsiveness for student learning difficulties.

In a detailed study by Butler (2007), the students' performance and their attitudes in two accented-English guises (American-accented & Korean-accented) were examined which revealed that in spite of the popular belief of the students who believed that non-native accented English produced a negative effect on their oral skills, the results failed to find any differences in student performance in terms of comprehension. Furthermore, Liaw (2012) found that students had different preferences regarding the aim of their English learning and when preparing for standardized English tests, such as TOEFL, they also preferred non-native speaking English teachers. Abeywickrama (2013) scrutinized the effects of non-native input on the test-takers' performances in a listening test. He concluded that using the variety of accented English did not have any impact on test-takers' performance. And the test takers felt non-native or accented English should not be used.

This study has also taken another factor into account, and this factor is 'gender'. Up to the present time, gender has been an ongoing if sometimes peripheral area of interest for researchers and practitioners in language teaching and testing. Early researchers in listening (e.g., Caffrey, 1953; Goldhaber & Weaver, 1968) suggest that men are better listeners than women. Weaver suggests that men and women listen differently. He believes that women are easily distracted by details. So these claims led the researchers to consider gender as a separate variable in this study. Apart from considering the effect of gender on listening comprehension ability of the test takers, this study also seeks to find out whether the two genders show the same tendency towards different input source (native vs. non-native speaker) or not.

1.2 Research questions

The variables involved in the present study generate the following research questions:

➢ Q1: Does input type (native speaker input vs. non-native speaker input) affect the performance of Iranian upper intermediate EFL test-takers in a listening test?

➢ Q2: Is there any significant difference between male and female listeners in a listening test, with different input types (native vs. non-native input)?

➢ Q3: What are test- takers' perceptions towards using non-native varieties of speech in listening comprehension tests?

2. Method

2.1 Participants

A total number of 66 upper-intermediate EFL learners from Shukuh-e- Iran language school (31 males and 35 females) within the age range of 18 to 25 took a proficiency test, among whom no outlier was excluded. The participants were randomly selected and four groups of 15 to 17 were made; and then the listening test was administered to them with the help of computers.

2.2 Instruments

Proficiency test - In order to have a controlled level of listening proficiency as well as homogenized participants, as far as their listening ability was concerned, a format of institutional Test of English as a Foreign Language (TOEFL) was used. For the purpose of the current study, the results of the listening section of the test were considered relevant only, which consisted of 20 multiple choice items. The reliability of the listening section of this test was also checked using KR-21 method and the result was 0.86, which reveals a high level of reliability of the test. Also, the reliability of the whole test was 0.9.

Listening test - A listening test consisting 20 multiple choice items (12 mini talks and two longer conversations each followed by four questions) taken from TOEFL Test section of the book "Longman Preparation Course for the TOEFL" Test by Deborah Phillips (2003) published by Pearson ESL was administered to the participants to seek the answer to the research question. In order to meet this objective, two groups listened to the recordings of native speakers while the other two groups listened to the talks between two non-native English speakers. For this purpose, the researcher asked two non-native English instructors (1 male and 1 female) to record their voice reading the same texts as those spoken by native English speakers. The test enjoyed a high level of reliability, that is, 0.85 employing K-R 21 Formula.

Native vs. non-native perception questionnaire - A perception questionnaire with the main purpose of finding out how test takers felt about the use of non-native input in listening comprehension tests was given to them. The questionnaire consisted of four main questions, each followed by options to be chosen. Also, there were two open-ended questions at the end, for which the test-takers were asked to respond if they thought something was not included in the questionnaire (See Appendix A)

2.3 Procedure and data analysis

In order to conduct this research the following steps were taken:

At the first step, a proficiency test was administered to a cohort of upper-intermediate EFL learners (i.e., 66) at Shukuh-e-Iran Language School. Having used the SPSS software, the researchers checked the normality of the distributions and they were assured of the homogeneity of the participants. After the participants were homogenized, they were given a cover letter the session before the test, in which the objectives of the study were clarified and participants were also assured that the results of the tests would not affect their final grades, and would not be disclosed to anyone else (even to their teacher). Participants were randomly assigned a participant number. The computer-based tests were set up so that odd-numbered participants automatically received native speaker input and even-numbered participants automatically received non-native speaker input. In order to have the non-native input, the researchers asked four EFL instructors (2 males and 2 females) to record their voice reading the required texts. Having done this, they checked the recordings and selected two of them whose voices sounded more natural: one male and female instructor. Also, test-takers were asked to fill out the Perception questionnaire which aimed at discovering their attitudes about non-native input in a listening test.

The data were entered into SPSS (Statistical Package for Social Sciences) software, PASW Statistics 18. In

order to find the answer to questions 1 and 2 independent samples *t*-test and two-way ANOVA were utilized, respectively. However, the last research question was analyzed by the researchers.

3. Results

3.1 Findings of the study

To ensure the homogeneity of the participants, an adapted version of TOEFL test was utilized. The scores of the performances of the participants in this test were entered into the SPSS software to check their homogeneity. Table 1 presents the results of the normality test.

Table 1

Tests of normality

	Kolmogorov-Smirnov Shapiro-Wilk					
	Statistic	df	Sig	Statistic	df	Sig
Exam scores	.14	66	.06	.93	66	.07

As it can be seen in the table above, the non-significant result (i.e., .06 which is more than .05) indicates normality. Furthermore, Figure 1 presents the related box plot which shows that there was no outlier among the participants.
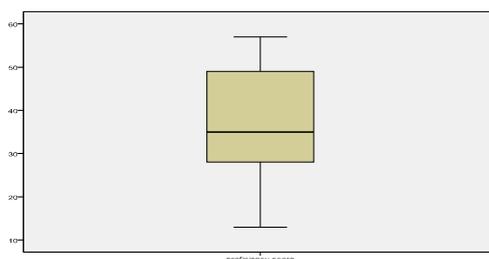

Figure 1. Box plot for homogeneity of participants.

After ensuring the homogeneity of the participants, an independent samples *t*-test was used to find the answers to the first research question. Table 2 provides the independent samples *t*-test statistics.

Table 2

Independent samples t-test for listening test considering input source variable

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Lower
Test core	Equal variances assumed	.56	.45	-2.63	64	.01	-2.09	.79	-3.67	-.50
	Equal variances not assumed			-2.63	63.79	.01	-2.09	.79	-3.67	-.50

As it can be seen from Table 2, the significance level shown by Levene's Test is .45 which is more than .05 and this suggests that equality of variances assumption is violated. It is shown in the table that the significance level is .01 which is less than the cutoff level of .05 and this suggests that there is a significant difference

between the mean scores of the groups listening to two different input sources, that is, native input and non-native input.

Comparing the mean scores of the groups, it is obvious that the test takers who have listened to non-native input (M= 13.85) had better performance compared to those who had listened to native input (M= 11.76). Also, the effect size is .29 which is somehow a big size. In other words 29% of the time the performance of the test takers is dependent on the input source.

Moreover, in order to find the answer to research question 2, in which the researchers tried to investigate whether different input sources had any significant effect on the listening test scores of the test takers considering their gender as the moderate variable or not, another set of analysis, that is, a two- way ANOVA was run. Table 3 presents descriptive statistics of the two- way between groups ANOVA. The Levene's test of homogeneity is also presented in Table 4.

Table 3

Descriptive statistics for two- way ANOVA

item modality	gender	Mean	Std. Deviation	N
	male	16.31	2.02	16
	female	16.94	1.98	17
written modality	Total	16.64	1.99	33
oral modality	male	11.73	2.73	15
	female	12.11	2.58	18
	Total	11.94	2.62	33
Total	male	14.10	3.31	31
	female	14.46	3.34	35
	Total	14.29	3.30	66

Table 4

Levene's test of equality of error variances

F	df1	df2	Sig.
2.52	3	62	.06

Table 4 presents the significance level for the Levene's test of variances which is .06 and since this is greater than .05, the assumption of the equality of variances is not violated. Having checked Levene's test of homogeneity, it is safe to report the data in Table 5 which shows the results of the two- way ANOVA conducted.

Table 5

Test of between groups two- way ANOVA

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	204.53[a]	3	68.17	7.91	.00	.27
Intercept	10610.16	1	10610.16	1232.10	.00	.95
input	63.15	1	63.15	7.33	.00	.10
gender	122.72	1	122.72	14.25	.00	.18
input * gender	9.93	1	9.93	1.15	.28	.01
Error	533.90	62	8.61			
Total	11557.00	66				
Corrected Total	738.43	65				

As it is clear in the table above, the interaction level between the two variables gender and input is not significant (gender* input: Sig. = .28). The data in this table reveal that input source affects the listening performances of the test takers (discussed in research question 1); and also, there is a statistically significant main effect for gender, that is gender does play a significant role as a moderate variable(sig. = .00). However, the effect size for the gender variable which is provided in the column Partial Eta Squared (i.e., .18) is classified as

small. Also the mean differences for the two genders considering input source is 1.96 and 3.51, a result that indicates that female test-takers have performed better than male test-takers in both conditions (native and non-native input source).

4. Discussion

The current study was designed to determine the effect of input source on the performance of EFL test takers in a listening test. It was hypothesized that listening to a native speaker or to a non- native speaker (i.e., input source) would not have any significant effect on the listening test scores. However, the results obtained through running an independent samples *t*-test showed that the null hypothesis was rejected. In fact, very little was found in the literature on the effect of this variable. As it was presented in the results section, the analysis of the elicited data suggests that the test takers who had listened to non-native input outperformed in their listening test.

The results of the current study corroborated those of Moinzedeh, Rezaei, and Dezhara (2012) who also concluded that the performance of the students in the group which listened to non-native accent (Persian) was significantly better compared to those students who listened to American accent English. Furthermore, the present findings seem to be consistent with the study conducted by Bent and Bradlow (2003). The results of their study showed that the listeners did significantly perform better when they listened to a speaker with the same L1. Major, Fitzmaurice, Bunta, and Balasubramanian (2002) also set up a study to investigate the relationship between the L1 background of the listener and the speaker. Their results were interesting since two sets of contradictory findings were reported in the results section. The first set of analysis reported findings which are in line with the findings of the current research. The results of Major et al.'s study showed a significant interaction effect between speaker and listener L1 background, and post hoc analyses showed that Spanish learners did much better when they listened to a speaker whose native language was Spanish, but the Chinese listener group was disadvantaged by listening to a speaker with the native language of Chinese. Therefore, the researchers concluded that listeners 'sometimes' performed better on a listening test when listening to a speaker sharing the same L1 with the listeners.

Moreover, the findings of this research support those of Ekong (1982). The results of his study indicated that Nigerian listeners understood English better when the speaker was a native speaker of Nigerian language. Also, the findings of the present study supported those of Wilcox (1978) in which he found that nonnative accent has superiority to native accent in the case of comprehensibility on the part of EFL learners. Furthermore, the results discussed above show that the findings of the present research are in line with claims made by Brown (1968). Brown suggested that accent familiarity (in our study Persian accent) can lead to the benefit of the listener.

As it was mentioned in the literature contradictory results were also abundant. As a prime instance, in Abeywickrama's study (2013) test takers from three different countries performed similarly even when the speakers shared the same native languages as the test takers, and students preferred native variety of English as the input in listening tests. Also, Smith and Bisazza (1982) found that Native American and British are more intelligible than native Chinese speakers. In another study, Stibbard and Lee (2006) aimed at finding the potential of using non-native English in listening tests. They concluded that there were no significant differences in intelligibility and comprehension scores for high proficiency non-native speakers and native speakers. Similarly, Munro et al. (2006) reached the conclusion that the effect of shared-L1 is so weak which is considered non-noticeable.

As a justification for the findings of the current study, the researcher would like to draw on the concept of cross- language speech perception. The cross-language speech perception considers the effects of L1 on L2 accents (i.e., transfer). So the listeners who have the same L1 as the speaker may have a kind of familiarity with the phonological patterns, intonation of saying an utterance, etc. Consistent with this justification, Bent and Bradlow (2003) refer to the listener and speaker who have the same native language as sharing a knowledge base which includes, "the system of consonants and vowel categories, phonotactics, stress patterns, and intonation as

well as other features of the sound system" (p. 1607). In addition, the speaking pace in which the text is read by the non-native speaker might have provided more processing time for the listeners to comprehend the input. Furthermore, the researcher of the present study draws on the EFL learners' common scare of not being able to understand everything their teachers, other speakers, or the speakers on a tape say. Chastain (1988) suggests that this kind of stress can be illuminated by reducing the speed of speech. Also, he asserts that speaking in phrases and lengthening the pauses between phrases can help accomplish this objective. The researcher believes that non-native input provides this chance for the learners, so that they can overcame their stress and this may lead to their better performance in listening tests.

Finally, test-takers' perceptions and attitudes towards using non-native input in listening comprehension tests were overall found to favor non-native input. More than half of the test-takers had a clear perception and preference towards using non-native input in the listening sections of the tests. In other words, they thought when the input was produced by a non-native speaker they had better understanding of the input and this would lead to better performances in tests. Some test takers also identified a native input as being difficult to understand.

The results of the perception questionnaire were mostly consistent with the findings of test-takers' performance data in listening test. Test-takers preferred non-native input as the input source in the listening tests and believed that their performance would be better in that case. Analyzing the performance data it was found that this claim was true for 88% of the test takers. It is worth mentioning that although some test takers (28 %) acknowledged the presence and importance of native input in their language use contexts and said that they would perform better in such tests (23 %) and also believed that they should be included in listening tests, overwhelmingly, they did not perform better on the tests. Since the majority of test takers had an advantage in listening comprehension and comprehensibility of non-native input, the researcher can refer to the concept of "Test Fairness Framework" mentioned by Kunnan (cited in Abeywickrama, 2013) who believes that factors such as non-equal access of test-takers to input sources might lead to considering the use of that input as unfair.

To sum up, the overall findings from the attitude questionnaire was that test takers felt native input is better not to be used and that by using a non-native variety a test taker who shares the same language as the speaker input will perform better on the listening tests.

5. Conclusion and implications

All in all, the findings of the study reveal that input source affects the performances of the EFL test takers and gender, as a moderate variable, plays a role in participants' performance as well. Furthermore, the results of the questionnaire reveal that more than 50 % of the test-takers preferred the use of non-native input in listening tests. The results of the present study might have direct and indirect implications for EFL teachers, curriculum designers, etc. An impetus to the EFL teachers is that since they are considered as a kind of model for EFL learners, it is very important for them to familiarize the learners with the native-like speech. This can be done through introducing or presenting materials such as movies, news programs, and whatever that can help the learners get used to native-like speech. Besides that, teachers should try to observe the elements that would help a speaker seem more native-like (e.g., stress, intonation, etc.).In addition, curriculum designers play an important role in helping the students improve their listening comprehension which leads to better performance in tests. As an example, they can specify a teaching-time for teaching note-taking strategies such as using symbols and abbreviations, using concept maps and diagrams, etc. Or they can allocate some sessions of the course for watching films or listening to real life conversations and for devising some activities upon them.

As with most research studies, there are some limitations with this study as well. In the present study, the sample only included upper intermediate EFL learners. Also, the numbers of participants per testing group was insufficient to permit meaningful comparisons; therefore, the results of this study may not generalize to other populations. In addition, the participants were recruited from courses held in one language school in September

2012. It is possible that students, who were in the first days of their attending schools or colleges, may show different traits, or even study habits. Therefore, this sample may not be representative of typical upper intermediate EFL learners. Moreover, this study is, of course, limited in the number of test items, test formats and features investigated as well.

Although several directions for future research have been touched upon so far, it is necessary to discuss some directions for further research as well. Other studies can apply some questionnaires in which the ideas of test takers on and which listening input source are solicited and compared. This information can be so useful for test designers as well as for test administers to consider some affective factors that might affect the results of the tests. Also, in view of the fact that listening tests include different types of questions (e.g., asking about the main idea, going into details, considering supra-segmental elements, etc.) analyzing the items one by one may reveal more interesting results. In other words, researchers can consider the effects that each of the pre-mentioned variables may have on different types of items in particular an overall test in general.

6. References

Abeywickrama, P. (2013). Why not non-native varieties of English as listening comprehension test input? *RELC Journal, 44* (1), 59–74. http://dx.doi.org/10.1177/0033688212473270

Bachman, L. F., & Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America, 114*(3), 1600–1610. http://dx.doi.org/10.1121/1.1603234

Brown, K. (1968). Intelligibility. In A. Davies (Ed.) *Language testing symposium: A psycholinguistic approach*, (pp. 180-191). London: O.P.U.

Buck, G. (2007). *Assessing listening*. Cambridge: Cambridge University Press.

Butler, Y. G. (2007). How are non-native English speaking teachers perceived by young learners? *TESOL, 41*(4), 731-755.

Caffrey, J. (1955). Auding ability at the secondary level. *Educations, 75,* 303-310.

Chastain, K. (1988). Developing second language skills: Theory and practice. NY: Harcourt Brace Jovanovich.

Crystal, D. (1997). *English as a global language*. Cambridge: Cambridge University Press.

Ekong, Pamela. (1982). On the use of an indigenous model for teaching English in Nigeria. *World Language English, 1*(2). 87-92. http://dx.doi.org/10.1111/j.1467-971X.1982.tb00474.x

Farhady, H., Jafarpur, A., & Birjandi, P. (1994). *Testing language skills from theory to practice*. Iran: SAMT.

Goldhaber, G., & Weaver, C. (1968). Listener comprehension of compressed speech with difficulty, rate of presentation and sex of listener varied. *Speech Monographs*, *35*, 20-25. http://dx.doi.org/10.1080/03637756809375562

Hughes, A. (2003). *Testing for Language Teachers*. New York: Cambridge University Press.

Kunnan, A. J. (2004). Test fairness. In M. Milanovic, & C. Weir (Eds.) *Studies in language testing 18: European language testing in global context* (pp. 27–48). Cambridge: Cambridge University Press.

Liaw, E. (2012). Examining student perspectives on the differences between native and non-native language teachers. *The Journal of Asia TEFL, 9* (3), 27-50.

Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, *36*(2), 173–190. http://dx.doi.org/10.2307/3588329

Mahboob, A. (2004). Native or non-native: What do students enrolled in an intensive English program think? In K. Stein (Ed.), *Learning and teaching from experience: Perspective on non-native English-speaking professionals* (pp. 121-149). Ann Arbor, MI: The University of Michigan Press.

Moinzadeh, A., Rezaei, O., & Dezhara, S. (2012). The effect of non-native accent on Iranian EFL learners' listening comprehension, focusing on Persian accent of English. *Journal of Language Teaching and Research, 3*(5), 967-972. http://dx.doi.org/10.4304/jltr.3.5.967-972

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second*

Language Acquisition, 28(1), 111–131. http://dx.doi.org/10.1017/S0272263106060049

Seidlhofer,   B. (2001). Closing a conceptual gap: The case for a description of English as a lingua franca. *International Journal of Applied Linguistics, 11*(2), 133–158. http://dx.doi.org/10.1111/1473-4192.00011

Smith, L.E., & Bisazza, J. A. (1982). The comprehensibility of three varieties of English for college students in seven countries. *Language Learning 32* (2), 129–269. http://dx.doi.org/10.1111/j.1467-1770.1982.tb00971.x

Stibbard, R. M., & Lee, J. (2006). Evidence against the mismatched interlanguage speech intelligibility benefit hypothesis. *Journal of the Acoustical Society of America, 120*(1), 433–442. http://dx.doi.org/10.1121/1.2203595

Thrasher, R. (2000). *Language testing*. International Christian University.

Weir, C. (1990). *Communicative language testing*. UK: Prentice Hall International.

Wilcox, G. K. (1978). The effect of accent on listening comprehension: A Singapore study. *English Language Teaching Journal, 32*,118-127. http://dx.doi.org/10.1093/elt/XXXII.2.118

Yano, Y. (2001). World Englishes in 2000 and beyond. *World English, 20* (2), 119–31. http://dx.doi.org/10.1111/1467-971X.00204

Zhang, Y., & Elder, C. (2011). Judgments of oral proficiency by non-native and native English speaking teacher raters: Competing or complementary constructs? *Language Testing*, *28*(1), 31–50. http://dx.doi.org/10.1177/0265532209360671