Intelligibility, Comprehensibility, and Accentedness of L2 Speech: The Role of Listener Experience and Semantic Context

Sara Kennedy
Pavel Trofimovich

Abstract: This study investigated how listener experience (extent of previous exposure to non-native speech) and semantic context (degree and type of semantic information available) influence measures of intelligibility, comprehensibility, and accentedness of non-native (L2) speech. Participants were 24 native English-speaking listeners, half experienced and half inexperienced with L2 speech, who transcribed and rated 90 English utterances spoken by six English and six Mandarin speakers. The utterances varied along two dimensions: real-world expectations (true vs. false utterances) and semantic meaningfulness (meaningful vs. meaningless utterances). Listeners with more experience understood more speech from the L1 and L2 speakers than listeners with less experience but did not rate it differently in comprehensibility and accentedness. All listeners understood and rated the utterances from L2 speakers based on the semantic context available: true–false utterances were understood and rated best, meaningless utterances least. These findings have implications for evaluating learner pronunciation and for training learners in successful L2 communication strategies.

Keywords: intelligibility, pronunciation, second language, linguistic experience, semantic context

Résumé : Dans cette étude, nous avons exploré la mesure dans laquelle l'expérience en compréhension orale (l'exposition antérieure à la langue étrangère) et le contexte sémantique (le degré et le type d'informations sémantiques à la disposition de l'auditeur) peuvent influencer l'intelligibilité, la compréhensibilité, et la perception de l'accent dans le discours oral en langue seconde (L2). L'échantillon comptait 24 auditeurs de langue anglaise – la moitié possédant une expérience à l'écoute en L2, l'autre moitié ne possédant pas d'expérience d'écoute en L2 – qui ont transcrit et évalué 90 énoncés en anglais prononcés par six locuteurs d'anglais et six locuteurs de mandarin. Les énoncés étaient variés selon deux dimensions : du point de vue des attentes conformément à la réalité

(déclarations vraies/déclarations fausses) et du point de vue du degré de signification sémantique (énoncés présentant une pleine signification/énoncés dénués de signification). Les auditeurs possédant une plus ample expérience en L2 ont compris une plus grande partie du discours oral des locuteurs de L1 et de L2 que les auditeurs moins expérimentés en L2, mais ils n'ont pas évalué ce discours différemment en ce qui concerne la compréhensibilité ou l'accent. Tous les auditeurs ont compris et évalué le discours des locuteurs de L2 à partir du contexte sémantique à leur disposition : les énoncés vrais/faux sont ceux qui ont été les mieux compris et les mieux évalués, et les énoncés dénués de signification l'ont été le moins. Ces résultats ont des implications pour l'évaluation de la prononciation des apprenants et pour l'apprentissage des stratégies de communication en L2.

Mots clés : intelligibilité, prononciation, langue seconde, expérience linguistique, contexte sémantique

Introduction

The ultimate goal of many second and foreign language (L2 and FL) teachers is to prepare students for successful communication outside the classroom. However, attaining this goal does not necessarily require that students become native-like in an L2. Students whose L2 production is not entirely native-like but who are able to communicate effectively are clearly successful L2 users. This pedagogical goal of successful communication is seen in textbooks on L2 pronunciation methodology. The great majority of textbook authors recommend that teachers strive for their students' oral production to be understandable, not necessarily native-like (Celce-Murcia, Brinton, & Goodwin, 1996; Kenworthy, 1987; Tench, 1981).

Teachers of students struggling to attain native-like L2 pronunciation must be relieved by this standard: non-native characteristics are acceptable if the speaker is understandable. However, if a teacher is able to generally understand an L2 speaker's speech and judges that speaker to be understandable, would a non-teacher do the same? This question is especially pertinent in situations where, unlike L2 teachers, interlocutors outside the classroom may not have frequent exposure to L2 speakers and so may have difficulty understanding their speech. It is also relevant when interlocutors cannot readily predict the meaning of an utterance from the context of communication.

The main objective of this study was, therefore, to investigate how two factors – listener experience and semantic context – influence measures of the intelligibility, comprehensibility, and accentedness of

L2 speech. *Listener experience* is here defined as the extent of previous exposure to L2 speech, while *semantic context* denotes the degree and type of semantic information available to listeners when deciphering an utterance. Such information is independent of the speech signal itself, permitting listeners to extract meaning not only from the segmental (i.e., individual sounds) and suprasegmental (i.e., stress, rhythm, and intonation) aspects of speech but from its syntax- and discourse-level aspects as well. Following Munro and Derwing (1995a), *intelligibility* is here defined as the extent to which a given utterance is understood by a listener. One way in which the construct of intelligibility is often (but not always) assessed is by scoring listeners' transcriptions of an utterance (for discussion of other measures of intelligibility see Munro, Derwing, & Morton, 2006). In turn, *comprehensibility* and *accentedness* are here defined as listeners' *perceptions* of how easily they understand an utterance and how closely the pronunciation of an utterance approaches that of a native speaker, respectively. From this point forward, *intelligibility* refers to listeners' actual understanding of an utterance, while *comprehensibility* refers to their perception of how easy or difficult the utterance is for them to understand.

Linguistic factors in L2 oral production

Numerous studies have targeted the relationship between L2 oral production and listeners' reactions to it. Results consistently show that increased intelligibility and more native-like ratings of comprehensibility and accentedness are linked to several measures of L2 oral production. Significant relationships have been found between grammatical accuracy and ratings of accentedness (Munro & Derwing, 1995a) and comprehensibility (Varonis & Gass, 1982). Phonemic and phonetic production accuracy have been shown to be significantly correlated with accentedness ratings (Anderson-Hsieh, Johnson, & Koehler, 1992; Munro & Derwing, 1995a; Riney & Flege, 1998; Riney, Takada, & Ota, 2000) and related to intelligibility (Jenkins, 2000). Various measures of suprasegmentals, including speech rate and pausing, have also been linked to accentedness (Anderson-Hsieh et al., 1992; Munro & Derwing, 1998, 2001; Trofimovich & Baker, 2006), comprehensibility (Munro & Derwing, 1998, 2001), and intelligibility (Anderson-Hsieh & Koehler, 1988; Tajima, Port, & Dalby, 1997).

It is clear from the above that aspects of L2 speech internal to the speech itself, such as segmental or suprasegmental accuracy, have

frequently been investigated for their relationship to intelligibility, accentedness, and comprehensibility. Less frequently studied, however, is the link between intelligibility, accentedness, and comprehensibility, on the one hand, and other factors independent of the speech itself, on the other. Such factors include the semantic context in which an utterance is set and individual differences in talkers or listeners. Investigating the contribution of these speech-external factors to intelligibility, comprehensibility, and accentedness may provide researchers, teachers, and speakers (both native and non-native) with additional ways to enhance successful communication. Thus, the particular question examined in this study is how intelligible, comprehensible, and accented L2 speech is to listeners, as a function of their previous experience with L2 speech and their knowledge of the semantic context of an utterance.

Listener experience and semantic context in the L1

Researchers have consistently observed effects of listener experience in native language (L1) communication. L1 listeners remember and understand more words and sentences spoken by a familiar than by an unfamiliar speaker (Bradlow, Nygaard, & Pisoni, 1999; Goldinger, 1996; Nygaard, Sommers, & Pisoni, 1994). Similarly, listeners who are more familiar with a certain type of L1 speech, such as sentences or words produced by hearing-impaired children, are more likely to find this speech intelligible than those who are unfamiliar with it (McGarr, 1981, 1983).

The semantic context available to L1 listeners also affects how accurately they understand L1 speech, particularly in circumstances that place a greater processing load on the listeners. For example, listeners tend to comprehend semantically predictable sentences (e.g., *The actor played the part*) better than semantically unpredictable ones (e.g., *The lawyer named the road*) when listening to sentences in noise (Rosenberg & Jarvella, 1970) or when listening to the speech of hearing-impaired children (McGarr, 1981). Listeners are also aided in their comprehension when they are given appropriate semantic cues prior to hearing an utterance. They are, for example, more accurate in their understanding of dysarthric speech (imperfectly articulated speech due to damage to the central or peripheral nervous system) when presented with thematically related cues than when such cues are not available (Dongilli, 1994).

Taken together, these findings suggest that listeners can understand L1 speech better if they have previous experience with it and if

the semantic context allows for more inferencing in interpreting it. Do listener experience and semantic context play a similar role in listeners' understanding of L2 speech?

Listener experience and semantic context in the L2

Relatively few studies have targeted effects of listener experience in understanding L2 speech (Bradlow & Bent, 2003; Munro et al., 2006).[1] Thus far, the results of these studies have been mixed. Some researchers report no effects of listener experience; Munro et al., for example, found that listeners with more exposure to particular L2 accents were not better at understanding speakers with those accents than were listeners whose exposure was less extensive. By contrast, other researchers reported significant effects of listener experience. In a perceptual training study, Bradlow and Bent presented their listeners with sentences embedded in noise. Those listeners who heard sentences recorded by multiple L2 speakers (i.e., listeners with more extensive experience) were subsequently better at understanding an unfamiliar L2 speaker than listeners who were exposed to sentences spoken by a single speaker (i.e., listeners with more limited experience).

Similarly, very little research has to date focused on semantic context effects in understanding L2 speech. In the only study known to us, Schmid and Yeni-Komshian (1999) showed that listeners are faster and more accurate at detecting a mispronunciation of a word when it is predictable from the preceding semantic context (e.g., *The bomb exploded with a plast*, where *plast* is a mispronunciation of *blast*) than when it is not (e.g., *Tom wants to know about the gake*, where *gake* replaces *cake*). Because detecting a mispronunciation in a word involves accurate identification of the word, it appears that listeners use the available semantic context to identify (and, thus, to understand) L2 speech.

As the preceding discussion suggests, there is a dearth of research on the effects of listener experience and semantic context in understanding L2 speech. The overall goal of the present study was, therefore, to add to and expand on this research, investigating the effects of both listener experience and semantic context within a single study. The point of departure for the investigation was a study by Gass and Varonis (1984), thus far the only researchers to examine how both listener experience and semantic context influence the understanding of L2 speech. We review their study in detail below.

Gass and Varonis (1984)

In their L2 intelligibility study, Gass and Varonis recorded native speakers of Japanese and Arabic reading aloud a short story and two sets of sentences in English. One sentence set was thematically related to the story; the other set was unrelated but was 'readily interpretable as a function of real-world knowledge' (p. 68). Recordings of this story and sentences were presented to native-English-speaking listeners, who transcribed the sentences and wrote a story summary. Semantic context was operationalized as listeners' familiarity with the topic of speech. Those listeners who first heard the story and later heard and transcribed the thematically related sentences were considered to be familiar with the topic. By contrast, those who first heard and transcribed the thematically related sentences and then listened to the story were considered to be unfamiliar with the topic. Although Gass and Varonis measured listener experience in a variety of ways (e.g., exposure to specific L2 accents or to particular speakers), for the purposes of the present study, we restrict our focus to one of their measures of listener experience, namely, the amount of listeners' exposure to L2 speech in the past. Some listeners had extensive exposure of this kind; others had very little.

With respect to semantic context, Gass and Varonis found that it did affect the intelligibility of L2 speech. Listeners who were familiar with the topic transcribed the thematically related sentences more accurately than listeners who were unfamiliar with it, suggesting that the knowledge of the story context aided them in comprehending the sentences. However, this finding may have confounded listeners' familiarity with the topic of an utterance and their familiarity with its individual words: the thematically related sentences transcribed by listeners contained many of the same words as the short story. In other words, listeners may have brought to their understanding of L2 speech not only their previous knowledge of the topic but also their familiarity with individual words in each sentence, suggesting that their significantly more accurate transcription may have been aided by both.

Gass and Varonis also found that listener experience, unlike semantic context, did not affect listeners' understanding of L2 speech, a finding similar to that reported by Munro et al. (2006). Listeners who had extensive exposure to L2 speech did not transcribe the unrelated sentences significantly more accurately than listeners who had little exposure. This finding is surprising, given that listener experience plays an important role in understanding L1 speech (e.g., Bradlow et al., 1999). There are several possible methodological

reasons for this finding. The stimulus set used by Gass and Varonis was small, containing only five sentences. It may be that a larger stimulus set heard by the listeners would have brought out differences due to exposure to L2 speech. Likewise, Gass and Varonis tested only a handful of experienced listeners. A larger participant sample might have accentuated potential effects of listener experience.

The present study

The present study was designed to extend the scope of Gass and Varonis' (1984) study and to address its methodological concerns in order to investigate the contribution of semantic context and listener experience to the intelligibility of L2 speech. First, a larger stimulus set was used (90 utterances), with a wider sample of experienced and inexperienced listeners (24 listeners in all). Second, L2 speakers of a different language background (Mandarin) were tested. Third, the degree of semantic context available to listeners was operationalized along two separate dimensions: real-world expectations (true vs. false utterances) and semantic meaningfulness (meaningful vs. meaningless utterances). Fourth, listeners' word familiarity was controlled in measuring semantic context effects. Finally, effects of semantic context and listener experience were investigated in relation to two other dimensions important in listeners' reactions to L2 speech: accentedness and comprehensibility. The study was guided by the following research questions:

1 How do intelligibility and ratings of accentedness and comprehensibility of L2 speech differ for sentences that vary in the degree of semantic context available?
2 How do intelligibility and ratings of accentedness and comprehensibility of L2 speech differ as a function of listeners' experience with L2 speech?

Method

Participants

Speakers

Twelve speakers recorded sentences in English. Six (3 female, 3 male) were native speakers of Mandarin (mean age: 30 years; range: 18–37).

All were born in China, where they had finished high school or completed a university degree. They had arrived in Canada at a mean age of 29 years (range: 18–35) and had resided there for an average of 1.5 years (range: 1–3). At the time of the study, they were students at the same English-medium university in Montreal. All had passed a university entrance exam that tested reading, writing, grammar, and vocabulary and had completed or were currently taking a required intermediate-level ESL writing course. The remaining six speakers (3 female, 3 male) were native speakers of North American English (mean age: 33.7 years; range: 26–42) born and raised either in Canada (4) or in the United States (2).

Listeners

The listeners were 24 native English speakers. Half the listeners (6 female, 6 male) were ESL teachers in Montreal (mean age: 39 years; range: 31–54) born and raised in Canada. All had at least three years of classroom teaching experience (mean: 8 years; range: 3–17). These listeners were classified as experienced in their exposure to L2 speech. The other half (10 female, 2 male) were residents of a small city in Ontario in which English is the only language used in public life. All were speakers of North American English (mean age: 37 years; range: 20–78), with the exception of one who had grown up in the United Kingdom but had moved to Canada as a young adult and had resided there for more than 30 years. These 12 listeners reported having little to no contact with L2 speakers of English. They were therefore classified as inexperienced in their exposure to L2 speech.

Materials

The materials included three lists of English sentences (see Appendix A). The first list, adapted from Munro and Derwing's (1995b) bank of single-clause true–false (T-F) sentences, contained 42 sentences that were meaningful and had a known (real-world) context. Half of the sentences could be labelled as true (e.g., *Most children like to eat cookies*), the other half as false (e.g., *Ice cream is very hot*), based on real-world knowledge. Each sentence consisted of four or five content words, all frequent lexical items. The second and third lists (containing 24 sentences each) were adapted from or modelled after Mack's lists of single-clause semantically meaningful and semantically anomalous sentences (Mack, Tierney, & Boyle, 1990;

TABLE 1
Characteristics of the three sentence lists used in this study

List characteristics		Sentence type	
	True–False	Semantically meaningful	Semantically anomalous
Mean sentence length (words)	5.8	6.4	6.4
Mean number of keywords per sentence	4.3	5.0	5.0
Proportion of 2,000 most frequent words	0.86	0.83	0.77

see also Mack, 1988, 1992). Each sentence consisted of five content words, and each had an adjective or numeral and a noun before the verb and an adjective or numeral and noun after the verb. Semantically meaningful sentences (e.g., *A new plan makes John nervous*) were meaningful but, unlike the T-F sentences, were from contexts unknown to the listener. Semantically anomalous sentences (e.g., *A dark nail zaps a ready reason*), although grammatical, were meaningless and so were interpretable only syntactically. The categorization of each sentence as either meaningful or anomalous was verified by a native speaker of English. The three lists were matched for sentence length, number of keywords (nouns, verbs, adverbs, adjectives), and word frequency. Relative word frequency for each list was estimated by determining the total number of words in each list that appear in West's (1953) list of the 2,000 most frequent English words. The characteristics of the sentence lists appear in Table 1.

The three sentence lists were designed to provide listeners with varying degrees of contextual support. The T-F sentences offered the most contextual support, as listeners were told that these sentences had to be either true or false in a real-world sense. Thus, for the T-F sentences, the listeners knew beforehand that all the sentences not only would be meaningful but would also refer to situations that were either true or false based on their common knowledge of the world. The meaningful sentences were semantically possible yet set in contexts unknown to listeners. These sentences, therefore, provided less contextual support than the T-F sentences. For these sentences, the listeners did not have any information about the context beyond the sentences; they could, however, decode each sentence using its semantic content, as the meaning of every sentence was semantically possible. The anomalous sentences, although syntactically acceptable, were not semantically possible at all and thus provided the least contextual support. In this case, then, the listeners could not use semantic content of the sentences to decode them; they had recourse

only to the meaning conveyed by the syntactic relationships among words.

Procedure

Recording

The sentences were recorded in individual sessions (20–45 minutes long) in a quiet room. Prior to the recording, each speaker responded orally to a demographic and language-background questionnaire. This provided the speakers, especially non-native ones, with practice in speaking English before the actual recording. Sentences were recorded list by list, and each speaker was given as much time as he or she needed to read through each list and to ask questions before recording. At that time, meanings and pronunciations of unfamiliar words (a range of one to five words per list for the L2 speakers) were explained or modelled. The order of lists was identical for all speakers: first true, then false, followed by semantically anomalous and, finally, semantically meaningful sentences. All speakers were told that the semantically anomalous sentences were 'weird' and were not supposed to make sense.

The recording procedure was identical for each list. Speakers were given a set of cards, each printed with one sentence. They were asked to read the sentences at a normal speed, pausing at the end of each. The sentences were digitally recorded directly onto a computer using a Plantronics (DSP-300) microphone and speech-editing software (CoolEdit 2000) at a 44.1-kHz sampling rate and a 16-bit resolution. If speakers misspoke (e.g., by removing or adding words or morphemes) or hesitated while reading a sentence, they were asked to read the sentence again. After recording each list, speakers took a short break if they so desired. Seven speakers returned to re-record some sentences (15 in total) because of audio distortions in their original recordings. Possible practice effects were of no concern, as the focus here was listeners' reactions to sentences, not the particular properties of speakers' production.

Stimulus preparation

Each recorded sentence was extracted from the audio file in which it was located and saved as a separate file. Each of the resulting files contained a 300- to 400-millisecond silence at the beginning and end. The 1,080 sound files (90 sentences × 12 speakers) were later

normalized for peak intensity to reduce differences in perceived loudness. The sentences were organized in four sets, with three randomized lists of sentences in each set: the T-F list, the semantically meaningful list, and the semantically anomalous list. In each of the four sets there were 24 T-F sentences, 12 semantically meaningful sentences, and 12 semantically anomalous sentences. These 48 sentences used in each set were drawn semi-randomly from the original pool of 1,080 recorded sound files to satisfy two conditions: (a) every speaker – male, female, native, and non-native – should appear equally frequently in each list and (b) every sentence should be equally represented across all sets, with half of the sentences appearing in two sets and the other half in the remaining two. Because there were fewer T-F sentences (42 sentences × 12 speakers) than semantically meaningful or anomalous sentences (48 sentences × 12 speakers), six T-F sentences were common to all four sentence sets.

Listening

Listening sessions were conducted with each participant individually in a quiet room. Each listener performed three tasks, always in the same order: (a) sentence transcription, (b) comprehensibility and accentedness rating, and (c) vocabulary knowledge test. In the first two tasks, each listener heard the same 48 sentences in three lists presented in the same fixed sequence (T-F, meaningful, anomalous), with sentences presented in a different random order for each list. The sentences were played using speech presentation software (Smith, 1997) over stereo headphones (Plantronics DSP-300).

For the transcription task, listeners were instructed first orally, then in writing, to write down what they heard as accurately as possible. Listeners were told to guess if they were unsure and to leave a space if they were unable to guess a word or sequence of words. For the T-F sentences (the first list presented), listeners were asked to label the sentences as true or false based on their general knowledge of the world. This was done only to ensure that listeners approached the T-F sentences explicitly within the context of their real-world knowledge; these T-F labels were therefore not analyzed further. For the other two sentence lists, listeners were given no additional instructions other than to transcribe what they heard. Prior to starting each list, listeners transcribed two practice items of the same sentence type, recorded by unfamiliar native and L2 speakers. Listeners proceeded at their own pace, writing each sentence they heard on the response sheet before them, then initiating the presentation of the next sentence

by pressing a button on the computer screen. After completing each list, listeners were allowed a break if needed.

For the comprehensibility and accentedness rating task, listeners were instructed to make two ratings for every sentence using two nine-point Likert scales. The comprehensibility scale (1 = very easy to understand, 9 = very hard to understand) and the accentedness scale (1 = no non-native accent, 9 = strong non-native accent) were based on similar scales previously used by Munro and Derwing (1995a). Listeners again completed two practice items before proceeding to rate each list. As in the transcription task, listeners were allowed to take a break after each list.

At the end of the listening session, listeners completed a written vocabulary knowledge test, modelled after Meara (1992), that contained all the words and names previously presented in sentences for transcription and rating. Nonce words (e.g., *purdy*) and nonce names (e.g., *Vark*) were included in the test as distracters. Listeners were asked to indicate, by circling 'yes,' 'no,' or 'not sure,' whether they knew, did not know, or were not sure of knowing the meaning of each word on the test. No listener marked any real words or names as being unknown or unfamiliar.

Data analysis

In the present study, intelligibility was operationalized as the accuracy of listeners' transcription (i.e., recognition and reproduction) of words in sentences. Each sentence was analyzed for transcription accuracy by computing an intelligibility score, defined as the ratio of accurately transcribed keywords (nouns, verbs, adverbs, adjectives) in each sentence to the total number of keywords in each sentence. Homophonous words (e.g., *seen* for *scene*) and misspellings (e.g., *artic* for *arctic*) were accepted as long as they were recognizable; words missing morphemes were coded as inaccurate. For example, if a listener transcribed the sentence 'Shy Jacob plays a boring game' as 'Side check plays a boring game,' the transcription score for this sentence was 3/5, or 0.6. The resulting sentence-based scores were then averaged across all sentences for each list, yielding three separate scores for each listener: one for T-F sentences, one for semantically meaningful sentences, and one for semantically anomalous sentences.

Comprehensibility and accentedness ratings were calculated similarly, by averaging for each listener the ratings across all sentences for each list. Again, there were three separate comprehensibility scores and three separate accentedness scores, one for each

sentence list. Inter-rater reliability analyses comparing comprehensibility and accentedness ratings for the listener groups yielded moderate to high indexes (range: 0.72–0.84), which suggests that the listeners were fairly consistent in their judgements. Data points that were more than three standard deviations above or below the mean for a given list were removed as outliers. Outlier data accounted for about 1% (1 sentence) of the comprehensibility ratings and 6% (4 sentences) of the accentedness ratings by inexperienced listeners. For all analyses discussed below, the alpha level for significance was set at 0.05; the reported effect sizes are partial eta squared (η_p^2), calculated by dividing the effect sum of squares by the effect sum of squares plus the error sum of squares. A Bonferroni procedure was applied to adjust the level of significance for all tests of simple main effects and all correlation analyses.

Results

Intelligibility

The intelligibility scores were subjected to a three-way analysis of variance (ANOVA) with listener experience (experienced, inexperienced) as a between-subjects factor and speaker background (native, L2) and semantic context (T-F, meaningful, anomalous) as within-subjects factors. This analysis yielded a significant main effect of listener experience ($F(1, 22) = 13.53$, $p = 0.001$, $\eta_p^2 = 0.38$). No other interaction involving the listener experience factor (including the three-way interaction) reached statistical significance ($F < 0.88$, $p > 0.38$, $\eta_p^2 < 0.04$). The obtained significant main effect of listener experience suggests that overall, regardless of speaker background and the semantic context of sentences, experienced listeners were more accurate than inexperienced listeners at transcribing sentences. This analysis also yielded a significant main effect of speaker background ($F(1, 22) = 363.03$, $p = 0.0001$, $\eta_p^2 = 0.94$), a significant main effect of semantic context ($F(2, 44) = 85.29$, $p = 0.0001$, $\eta_p^2 = 0.80$), and a significant interaction between these two factors ($F(2, 44) = 95.46$, $p = 0.0001$, $\eta_p^2 = 0.81$). Tests of simple main effects carried out to explore this significant interaction revealed two additional findings. First, sentences spoken by native speakers were always transcribed more accurately than sentences spoken by L2 speakers ($p < 0.0001$). Second, semantic context affected transcription accuracy only for sentences spoken by L2 speakers; transcription was most accurate for the T-F sentences, less accurate for the semantically

FIGURE 1

Mean proportion of keywords correctly transcribed by inexperienced and experienced listeners as a function of speaker background and semantic context of sentences (± 2 SE). The invisible error bars indicate very small variances.
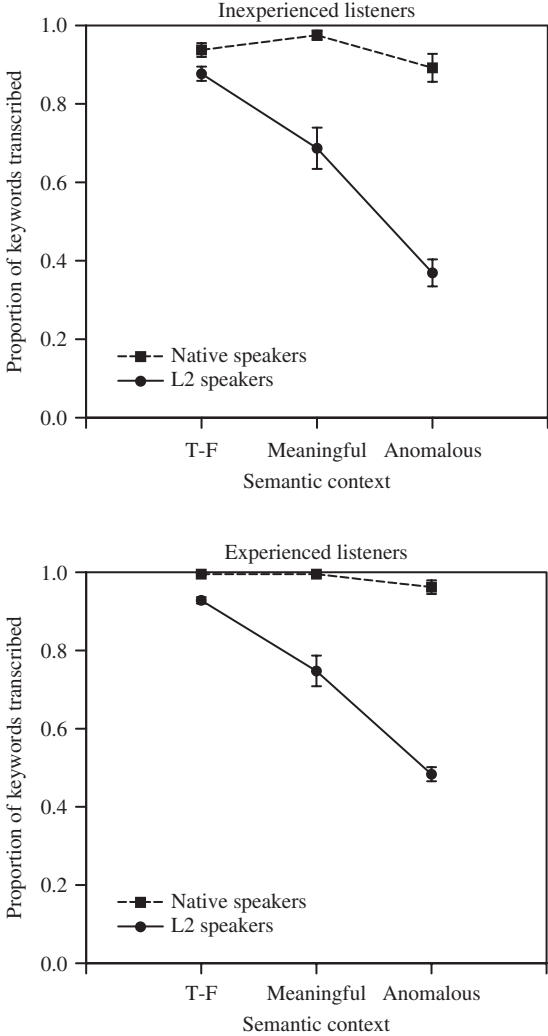
meaningful sentences, and least accurate for the semantically anomalous sentences ($p < 0.004$). Mean intelligibility accuracy scores are plotted separately for experienced and inexperienced listeners in Figure 1. Higher values represent more intelligible sentences.

Comprehensibility

The comprehensibility ratings were subjected to a similar three-way ANOVA. This analysis yielded significant main effects of speaker background ($F(1, 22) = 656.02$, $p = 0.0001$, $\eta_p^2 = 0.97$) and semantic context ($F(2, 44) = 90.47$, $p = 0.0001$, $\eta_p^2 = 0.80$), as well as a significant interaction between these two factors ($F(2, 44) = 27.90$, $p = 0.0001$, $\eta_p^2 = 0.56$). Neither the main effect of listener experience nor any interaction involving this factor (including the three-way interaction) reached statistical significance ($F < 1.95$, $p > 0.15$, $\eta_p^2 < 0.08$). Tests of simple main effects carried out to explore the significant speaker background × semantic context interaction revealed two findings, for inexperienced and experienced listeners alike. First, sentences spoken by native speakers were always rated more comprehensible than sentences spoken by L2 speakers ($p < 0.0001$). Second, semantic context affected comprehensibility ratings only for sentences spoken by L2 speakers. Both groups of listeners rated the T-F sentences significantly the most comprehensible, the semantically meaningful sentences less comprehensible, and semantically anomalous sentences the least comprehensible ($p < 0.004$). Mean comprehensibility ratings are plotted separately for experienced and inexperienced listeners in Figure 2. Lower values represent more comprehensible sentences.

Accentedness

The accentedness ratings were subjected to a similar three-way ANOVA. This analysis yielded significant main effects of speaker background ($F(1, 22) = 1634.62$, $p = 0.0001$, $\eta_p^2 = 0.99$) and semantic context ($F(2, 44) = 27.96$, $p = 0.0001$, $\eta_p^2 = 0.56$) and a significant interaction between these two factors ($F(2, 44) = 25.87$, $p = 0.0001$, $\eta_p^2 = 0.54$). As in the previous analysis, neither the main effect of listener experience nor any interaction involving this factor (including the three-way interaction) reached statistical significance ($F < 4.18$, $p > 0.06$, $\eta_p^2 < 0.17$). Tests of simple main effects carried out to explore the significant speaker background × semantic context interaction revealed two findings, again identical for inexperienced and experienced listeners. First, sentences spoken by native speakers were always rated less accented than sentences spoken by L2 speakers ($p < 0.0001$). Second, semantic context affected accentedness ratings only for sentences spoken by L2 speakers. Both groups of listeners rated the T-F sentences less accented than either semantically

FIGURE 2
Mean comprehensibility ratings by inexperienced and experienced listeners as a function of speaker background and semantic context of sentences (± 2 SE). The invisible error bars indicate very small variances.
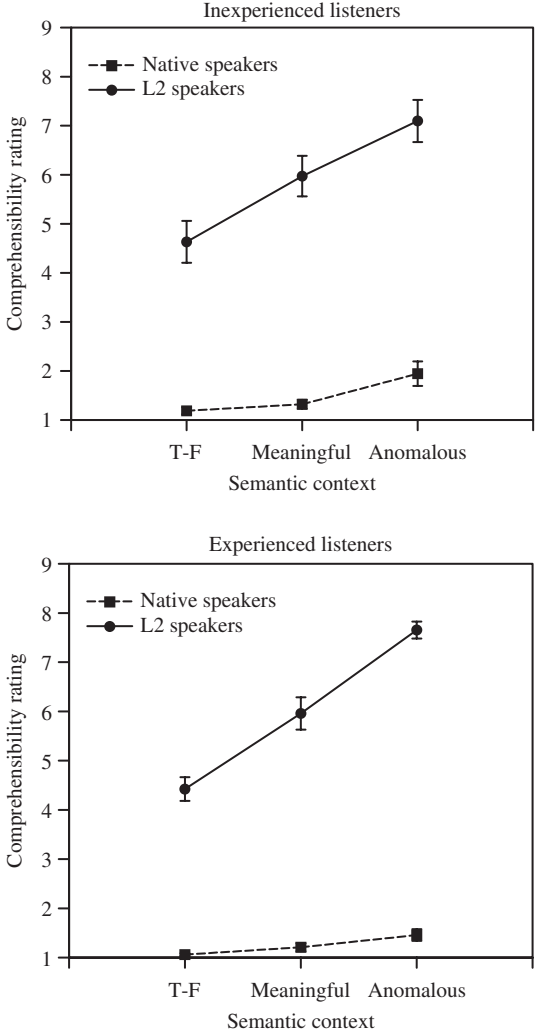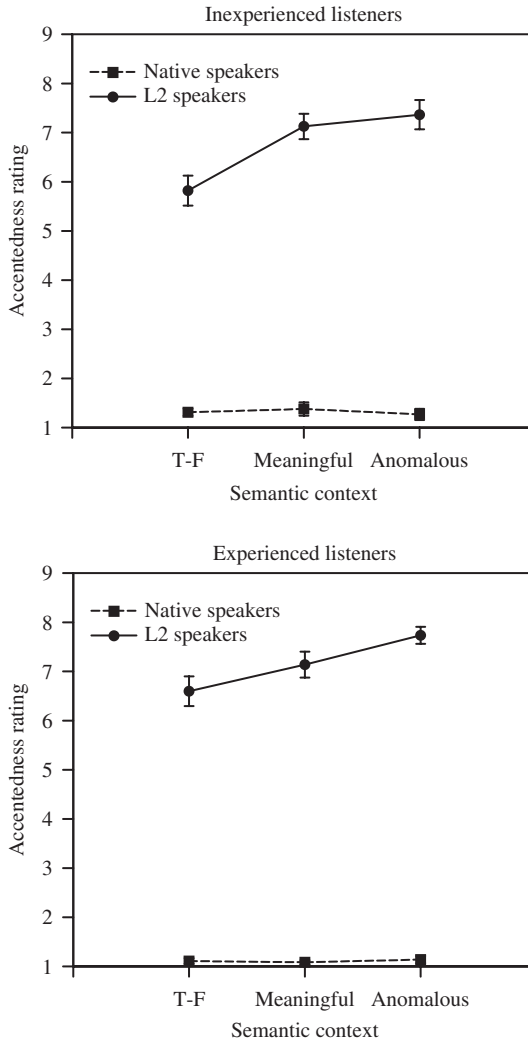
FIGURE 3
Mean accentedness ratings by inexperienced and experienced listeners as a function of speaker background and semantic context of sentences (± 2 SE). The invisible error bars indicate very small variances.

meaningful or semantically anomalous sentences ($p < 0.002$). Mean accentedness ratings are plotted separately for experienced and inexperienced listeners in Figure 3. Lower values represent less accented sentences.

Listener age

One possible concern about the results reported above is that listeners' age might have influenced their understanding and rating of L2 speech. Indeed, several listeners, particularly those in the inexperienced group, were over 50 (the oldest being 78). Although all participants reported no problems with hearing, it is still possible that older participants may have experienced some hearing loss or decline in cognitive ability. In fact, previous research has shown that listeners' perception of L2 speech can be adversely affected by aging (Burda, Casey, Foster, Pilkington, & Reppe, 2006; Burda, Scherz, Hageman, & Edwards, 2003). Before discussing the results of the present study, therefore, it was important to determine whether the participants' understanding and rating of L2 speech depended on their age. To investigate this, we carried out three analyses.

In the first analysis, we examined the original data (i.e., data containing all original responses, including outliers) in order to check for inexplicable or inconsistent scores and ratings (particularly those from older listeners) within each sentence type and for each speaker group. This analysis revealed no individual listeners with patterns of responses atypical of or highly inconsistent with those of the other listeners in their group. In the second analysis, we computed several sets of Pearson correlations (two-tailed) between the participants' age, on the one hand, and all their intelligibility scores and comprehensibility and accentedness ratings, on the other ($N = 24$). This analysis yielded no significant associations between the three performance measures and the participants' age ($r(22) < 0.37$, $p > 0.07$), suggesting that older participants did not tend to understand less or to give different ratings than younger participants. In the final analysis, we recomputed all ANOVAs reported earlier with the data for the two oldest participants removed. These ANOVAs yielded the same pattern of statistically significant main effects and interactions as the analyses that included the data from all participants. Taken together, these results helped rule out the possibility that differences in participants' age affected the findings obtained.

Discussion

Semantic context

The first research question targets the contribution of semantic context to intelligibility and to ratings of comprehensibility and accentedness

of native and L2 speech. Semantic context did affect the intelligibility of sentences spoken by L2 speakers. This finding both adds to and extends previously published results (Schmid & Yeni-Komshian, 1999). For all listeners, both experienced and inexperienced, less semantic context was associated with less intelligible L2 speech. In making sense of L2 speech, at least to the extent revealed through transcription of individual utterances, listeners appeared to use the semantic context available to them at sentence and discourse levels. They did so by relying on their real-world expectations (required for T-F sentences) and on the semantic possibility of an utterance (available in semantically meaningful sentences). In essence, for these listeners, understanding L2 speech went beyond merely decoding the speech signal.

Semantic context was also a factor in comprehensibility and accentedness ratings, again only for L2 speech. Not surprisingly, listeners' comprehensibility ratings were affected by semantic context in much the same way as their understanding of L2 utterances (T-F sentences were rated easiest to understand, and semantically anomalous sentences rated hardest to understand). The difficulties listeners had in understanding L2 speech were thus reflected in their ratings: what was in fact more difficult to understand was rated more difficult to understand. However, what was more difficult to understand was also rated as more accented. This finding was puzzling, at least at first glance. Because all sentences were spoken by the same speakers, accentedness ratings should not have varied significantly from list to list. One possibility is that it was not only the speaker's degree of accent that counted in the rating but also the listener's degree of understanding (Munro & Derwing, 1995a). Those sentences for which a real-world context was available were easily understood, leading listeners to rate them as less accented. Those sentences not situated in real-world contexts (semantically meaningful and anomalous) proved difficult to understand, leading listeners to rate these sentences as more accented.[2]

For native speech, however, the results were quite different. Neither sentence intelligibility nor comprehensibility and accentedness ratings were affected by semantic context. Regardless of the degree of semantic context available, both experienced and inexperienced listeners understood native speech, rating it very easy to understand and unaccented. Semantic context thus appears to have been much less important when speakers' pronunciation was native-like. Native-like pronunciation conveyed the consistent, unambiguous phonological detail that listeners could use to recognize and reproduce words from sentences.

Listener experience

The second research question deals with the role of experience in listeners' understanding and their ratings of accentedness and comprehensibility of native and L2 speech. Listener experience affected the intelligibility of sentences spoken by both native and L2 speakers. The experienced listeners in this study (ESL teachers with at least three years of classroom experience) were, overall, significantly more accurate than the inexperienced listeners (native speakers of English with little reported exposure to L2 speech) in their transcription of sentences. This finding clarifies the unexpected result previously reported by Gass and Varonis (1984), who found no significant effects of experience on listeners' understanding of L2 speech (see also Munro et al., 2006). With a larger stimulus set and a greater number of experienced listeners, the present study allowed for a more powerful test of experience effects, showing that, in this case, listeners with greater experience with L2 speech find it more intelligible (see Bradlow et al., 1999, and McGarr, 1981, for similar findings in L1 research).

Furthermore, this study goes beyond previously published research by demonstrating that listener experience does not mediate the effect of semantic context on listeners' understanding of L2 speech. Indeed, the lack of interaction between listener experience and semantic context suggests that the experienced listeners' advantage in understanding L2 speech was not because they were more successful at using the available semantic context. Rather, the experienced listeners' advantage likely stemmed from their greater knowledge of how L2 speakers' pronunciation differs from that of native speakers. This knowledge enabled the experienced listeners to be more accurate and adept at decoding L2 speech. More generally, it would appear that experienced listeners' exposure to the large range of variability found in L2 speech (e.g., variability in speakers' accents, voice quality, speech rates) enhances their skill in understanding L2 speech and possibly boosts their already-native ability to understand *native* (L1) speech as well. However, the latter finding should be investigated further in future research.[3]

Although experienced listeners understood more of L2 speech, they did not rate this speech as more comprehensible than inexperienced listeners did. This finding points to an important distinction between intelligibility measurements, on the one hand, and ratings of comprehensibility, on the other (see also Munro & Derwing, 1995a). It suggests that comprehensibility ratings, indicating listeners' ease of

understanding, do not necessarily reflect the extent to which speech is actually understood (e.g., Munro et al., 2006). One possible reason that the experienced and inexperienced listeners' comprehensibility ratings did not differ is that the experienced listeners, although more skilled than the inexperienced listeners at understanding L2 speech, overstated their difficulty in understanding it. A more likely explanation, however, is that the inexperienced listeners understated their difficulty in understanding L2 speech. Being less skilled at understanding L2 speech, the inexperienced listeners may not have been aware of how much they did not understand. The latter interpretation is supported by research showing that, unlike individuals who are more skilled in a particular domain (e.g., logical thinking, writing), individuals who are less skilled in that domain tend to underestimate their difficulties (Dunning, Johnson, Ehrlinger, & Kruger, 2003; Kruger & Dunning, 1999).

Like comprehensibility ratings, listeners' accentedness ratings do not seem to have been influenced by listener experience. Neither experienced nor inexperienced listeners differed in how they rated the accentedness of sentences spoken by native and L2 speakers (although, as discussed earlier, the ratings for L2 speech were affected by semantic context). This finding suggests that the experienced and inexperienced listeners shared a common scale for judging accentedness. Both sets of listeners also seem to have applied this scale in similar ways in a variety of situations – both when a real-world semantic context was available and when such context was lacking.

Learning to understand L2 speech

The results of this study suggest that experience with L2 speech may enable listeners to find it more intelligible. How, then, could listeners gain the experience needed to increase their understanding of L2 speech? In previous research, listener experience has typically been characterized in terms of listeners' exposure to a number of sentences spoken by speakers of the same language background. For example, in some studies, listeners were exposed to sentences produced by several speakers over one training session and were later tested on their understanding of other sentences from those same speakers (Burda, Overhake, & Thompson, 2005; Clarke, 2000; Wingstedt & Schulman, 1984). As a result, listeners did tend to improve in their ability to understand L2 speech. What was not tested, however, was their understanding of sentences spoken by unfamiliar speakers or

by speakers of an unfamiliar language background. This ability to transfer knowledge about L2 speech gained through training is clearly an important benchmark that can help determine both the effectiveness of listener training in a laboratory setting and the benefits of listener experience in the real world.

This benchmark has been applied in one study focusing on particular types of listener experience needed to understand L2 speech. Bradlow and Bent (2003) exposed listeners, over two days, to five repetitions of two sets of simple English sentences embedded in noise (for a total of 160 sentences) spoken by a single Mandarin speaker or by multiple Mandarin speakers. They then tested listeners with new sentences from familiar and unfamiliar Mandarin speakers (same language) and an unfamiliar Slovakian speaker (different language). Bradlow and Bent found that the training with multiple Mandarin speakers gave listeners the most advantage in understanding an unfamiliar Mandarin speaker. Becoming better at understanding L2 speech may thus involve exposure to multiple speakers (i.e., exposure to L2 speech that is highly variable). Bradlow and Bent also found that the training with multiple Mandarin speakers did not confer any advantage to listeners in their understanding of an unfamiliar Slovakian speaker. Thus, it appears that understanding L2 speakers of a particular language background may require prior exposure to speakers of this language background.

These findings suggest that it is important to expose inexperienced listeners to L2 speech produced by multiple speakers of multiple language backgrounds. To date, there have been several practical suggestions as to how to put this plan into action, most of them specific to educational settings. One suggestion, for example, includes running orientation sessions for first-year university students (Krech Thomas, 2004). Through the involvement of L2 speakers in these orientations, students could be exposed to a variety of L2 accents. Another suggestion involves designing individual pedagogical activities or, at a larger scale, entire curricula that increase native speakers' awareness of different cultures and highlight ways in which L2 speech may differ from their own (Derwing, Rossiter, & Munro, 2002). Overall, the successful implementation of all these suggestions clearly depends, at a minimum, on listeners' openness to interacting with L2 speakers.

Implications for teaching

The findings of this study reveal effects of listener experience and semantic context on experienced and inexperienced listeners'

understanding and ratings of L2 speech. Although these findings come from a laboratory-based experimental study employing single sentences as stimuli, they may nevertheless have implications for classroom L2 teaching, particularly for teacher evaluations of learner intelligibility. First, teachers may typically evaluate learner intelligibility based on how well they themselves understand those learners; such evaluation could be problematic if teachers, as experienced listeners, understand their learners' speech better than many other listeners, both inside and outside the classroom. Voicing this concern, Kenworthy (1987) encourages teachers to invite non-teachers or, at the very least, unfamiliar teachers to their classrooms to judge learners' intelligibility. In another approach, Elson (1992), citing Klyhn (1986), emphasizes the importance of classroom speaking activities between learners of different language backgrounds. By observing such interactions, he suggests, teachers can identify and subsequently address those aspects of learner speech that cause failures in communication (see also Jenkins, 2000). A teacher who has become expert at understanding L2 speech should thus bear in mind that the intelligibility of her learners may depend not only on their own skills but on their teacher's as well.

If intelligibility is not just a matter of a speaker's pronunciation quality, what else can teachers do, besides targeting their learners' pronunciation, to help them communicate effectively in an L2 environment? Because listeners use more than just the speech signal to interpret an utterance, teachers could teach their learners how to provide more contextual information to their listeners in order to be more easily understood (Dörnyei, 1995; Rossiter, 2003) and assess them using tasks that involve little shared contextual knowledge between interlocutors (e.g., explaining a little-known aspect of a student's native culture). Littlewood (1984) mentions several communication strategies that might provide the listener with more contextual support, including paraphrasing (e.g., saying *the thing you boil water in* for *kettle*), approximation or substitution (e.g., saying *curtains* for *blinds* when the word *blinds* is unknown), or rephrasing (i.e., repeating the same idea in a more explicit, simplified form). These and similar communication strategies are often explicitly taught when pragmatic and pragmalinguistic features of language are the focus of attention (e.g., making requests more explicit when an earlier, indirect request seems to have been misinterpreted). These strategies, then, can also be used in conversation management, either to prevent or to repair communication breakdowns. All of these strategies can enhance the semantic context available to

a listener when a speaker's pronunciation may be a source of misunderstanding.

Future directions

The study reported here provides some evidence that both listener experience and semantic context can play a role in L1 listeners' understanding of L1 and L2 speech and their ratings of its accentedness and comprehensibility. In future research, it is important to investigate the importance of these two factors in L2 listeners' understanding of L2 speech. As Jenkins (2000) puts it, 'English is now learnt and spoken most frequently to serve international functions among L2 speakers in international contexts' (p. 16). Another target for future research should be listeners' understanding of L2 speech in the context of authentic interactions with interlocutors in a real-world setting. The roles of listener experience and semantic context in these types of interactions have not often been explored in depth (see Varonis & Gass, 1985, and Jenkins, for rare exceptions). The ultimate goal of these (and other) future studies would be to investigate how native and non-native interlocutors can work, in the context of real-time interaction, to avoid and repair communication breakdowns.

Sara Kennedy is completing her PhD studies in Second Language Education at McGill University in Montreal. Her research focuses on intelligibility of L2 speech; effects of classroom instruction, particularly the teaching of L2 oral skills; and the role of language experience in the development of L2 speaking ability. She has extensive experience in teaching English as a second and foreign language.
Contact: sara.kennedy@mcgill.ca

Pavel Trofimovich is Assistant Professor of Applied Linguistics at the TESL Centre/Department of Education at Concordia University in Montreal. His research and teaching focus on cognitive aspects of L2 processing, L2 phonology, sociolinguistic aspects of SLA, and computer-assisted language learning.

Acknowledgements

Notes

1   Other studies, which are not directly relevant to this discussion, have examined listener L1 background as a factor in understanding L2 speech (e.g., Major, Fitzmaurice, Bunta, & Balasubramanian, 2002; Smith & Bisazza, 1982). While listener L1 background is an important aspect of a listener's experience, listener experience is defined in the present study as the extent of listeners' previous exposure to L2 speech.

2   An anonymous *CMLR* reviewer suggested that the difference in accentedness ratings among lists for the L1 Chinese speakers may have been due to unequal distribution of difficult L2 segments in different word positions. For example, English/z/may have been more frequent in the word-initial position in the semantically anomalous and semantically meaningful lists than in the T-F list, where it appeared mostly word-finally. Assuming that English/z/is more problematic for native Chinese speakers word-initially than word-finally, the speakers may therefore have had more difficulty in producing such words, leading to less native-like accentedness ratings for the semantically meaningful and anomalous sentences. This interesting explanation should be addressed in future research.

3   Some factors to consider in this future research might include task characteristics and individual listener differences (e.g., age, education level, gender). For example, as suggested by an anonymous *CMLR* reviewer, some listeners may be more skilled than others at completing transcription tasks, which could misrepresent listeners' levels of understanding.

References

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, *42*, 529–555.

Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning*, *38*, 561–613.

Bradlow, A., & Bent, T. (2003). Listener adaptation to foreign-accented English. In M.J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1581–1583). Barcelona: Universitat Autònoma de Barcelona.

Bradlow, A., Nygaard, L., & Pisoni, D. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception and Psychophysics*, *61*, 206–219.

Burda, A.N., Casey, A.M., Foster, T.R., Pilkington, A.K., & Reppe, E.A. (2006). Effects of accent and age on the transcription of medically related utterances: A pilot study. *Communication Disorders Quarterly*, *27*, 110–116.

Burda, A., Overhake, D., & Thompson, K. (2005). Familiarity and older adults' transcriptions of native and nonnative speech. *Perceptual and Motor Skills*, *100*, 939–942.

Burda, A.N., Scherz, J.A., Hageman, C.F., & Edwards, H.T. (2003). Age and understanding speakers with Spanish or Taiwanese accents. *Perceptual and Motor Skills*, *97*, 11–20.

Celce-Murcia, M., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. New York: Cambridge University Press.

Clarke, C. (2000). Perceptual adjustments to foreign accented English. *Research on Spoken Language Processing*, *24*, 321–335.

Derwing, T., Rossiter, M., & Munro, M. (2002). Teaching native speakers to listen to foreign-accented speech. *Journal of Multilingual and Multicultural Development*, *23*, 245–259.

Dongilli, P. (1994). Semantic context and speech intelligibility. In J.A. Till, K.M. Yorkston, & D.R. Beukelman (Eds.), *Motor speech disorders: Advances in assessment and treatment* (pp. 175–191). Baltimore, MD: Paul H. Brooks.

Dörnyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly*, *29*, 55–85.

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*, 83–87.

Elson, N. (1992). Unintelligibility and the ESL learner. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 229–236). Oxford: Oxford University Press.

Gass, S., & Varonis, E. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, *34*, 65–89.

Goldinger, S.D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1166–1183.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.

Kenworthy, J. (1987). *Teaching English pronunciation*. New York: Longman.

Klyhn, J. (1986). International English: Communication is the name of the game. *TESOL Newsletter*, 20(2), 1–6.

Krech Thomas, H. (2004). *Training strategies for improving listeners' comprehension of foreign-accented speech*. Unpublished doctoral dissertation, University of Colorado, Boulder.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*, 1121–1134.

Littlewood, W. (1984). *Foreign and second language learning: Language acquisition and its implications for the classroom*. Cambridge: Cambridge University Press.

Mack, M. (1988). Sentence processing by non-native speakers of English: Evidence from the perception of natural and computer-generated anomalous L2 sentences. *Journal of Neurolinguistics*, 3, 293–316.

Mack, M. (1992). How well is computer-processed speech understood? A cross-linguistic and cross-dialectal analysis. *World Englishes*, 11, 285–301.

Mack, M., Tierney, J., & Boyle, M.E.T. (1990). The intelligibility of natural and LPC-vocoded words and sentences presented to native and non-native speakers of English. MIT Lincoln Laboratory Technical Report No. 869.

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*, 173–190.

McGarr, N. (1981). The effect of context on the intelligibility of hearing and deaf children's speech. *Language and Speech*, 24, 255–264.

McGarr, N. (1983). The intelligibility of deaf speech to experienced and inexperienced listeners. *Journal of Speech and Hearing Research*, 26, 451–458.

Meara, P. (1992). *The Swansea vocabulary tests*. Swansea, UK: University of Swansea.

Munro, M., & Derwing, T. (1995a). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning*, 45, 73–97.

Munro, M., & Derwing, T. (1995b). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, 38, 289–306.

Munro, M., & Derwing, T. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning*, *48,* 159–182.

Munro, M., & Derwing, T. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition*, *23,* 451–468.

Munro, M.J., Derwing, T.M., & Morton, S.L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, *28,* 111–131.

Nygaard, L., Sommers, M., & Pisoni, D. (1994). Speech perception as a talker-contingent process. *Psychological Science*, *5,* 42–46.

Riney, T., & Flege, J.E. (1998). Changes over time in global foreign accent and liquid identifiability and accuracy. *Studies in Second Language Acquisition*, *20,* 213–243.

Riney, T., Takada, M., & Ota, M. (2000). Segmentals and foreign language accent: The Japanese flap in EFL. *TESOL Quarterly*, *34,* 711–737.

Rosenberg, S., & Jarvella, R.J. (1970). Semantic integration and sentence perception. *Journal of Verbal Learning and Verbal Behavior, 9,* 548–553.

Rossiter, M. (2003). 'It's like chicken but bigger': Effects of communication strategy in the ESL classroom. *The Canadian Modern Language Review, 60,* 105–121.

Schmid, P.M., & Yeni-Komshian, G.H. (1999). The effects of speaker accent and target predictability on perception of mispronunciations. *Journal of Speech, Language, and Hearing Research*, *42,* 56–64.

Smith, L.E., & Bisazza, J.A. (1982). The comprehensibility of three varieties of English for college students in seven countries. *Language Learning*, *32,* 129–269.

Smith, S.C. (1997). *UAB Software*. Birmingham: Department of Rehabilitation Sciences, University of Alabama.

Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics, 25,* 1–24.

Tench, P. (1981). *Pronunciation skills*. London: Macmillan.

Trofimovich, P., & Baker, W. (2006). Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, *28,* 1–30.

Varonis, E., & Gass, S. (1982). The comprehensibility of nonnative speech. *Studies in Second Language Acquisition*, *4,* 114–136.

Varonis, E., & Gass, S. (1985). Miscommunication in native-nonnative conversation. *Language in Society*, *14,* 327–343.

West, M. (1953). *A general service list of English words*. London: Longman, Green & Co.

Wingstedt, M., & Schulman, R. (1984). Comprehension of foreign accents. In W. Dressler (Ed.), *Phonologica 1984: Proceedings of the Fifth International Phonology Meeting*. Warsaw: Adam Mickiewicz.

Appendix A

True–False Sentences (based on Munro & Derwing, 1995b)

Sometimes animals live in caves.
June is the first month of the year.
Crayons come in many colours.
Some people enjoy playing basketball.
Mosquitoes have soft pink fur.
Many lawyers ride cows to work.
Most horses have sixteen eyes.
Some cows like to read books.
Some people have brown eyes.
Trucks drive on the highway.
People play baseball with a piano.
Most tigers live in apartments.
In the winter, the snow is green.
Rice is popular in China.
Some people have sandwiches for lunch.
Apples and oranges are fruit.
Many people can walk on the ceiling.
Cola is a type of soft drink.
Some flowers bloom in the summer.
Planes fly through the air.
Some people like to watch television.
Most fish live on dry land.
Adults are usually younger than children.
Most people wear watches on their lips.
Some people walk on their ears.
Sugar is bad for your teeth.
Mice usually like to chase dogs.
Rocks make a delicious soup.
Most babies like to drink milk.
Police often wear uniforms.
Many dogs like to smoke cigars.
Riding a bicycle is good exercise.
Cabbages are usually highly intelligent.

Killer bees usually make good pets.
Most children like to eat cookies.
Ice cream is very hot.
Dogs usually wear gloves on their feet.
It is very warm in the Arctic.
Most caterpillars turn into butterflies.
Salad is a healthy food to eat.
Some people eat cereal for breakfast.
You can use credit cards at many stores.

Semantically Meaningful Sentences (based on Mack et al., 1990)

Crazy Mary digs a deep hole.
Careful Sharon reported good reasons.
A caring doctor helps the sick patients.
Fast Jenny runs a short race.
A happy zebra checked a dirty river.
A lovely child bought a thick book.
The nervous mother feels the sharp knife.
Pretty Jane throws a pink ball.
A rescued snake visited thirty zoos.
The strong chief kicked a mad dog.
Thoughtful Steve buys soya milk.
Zero chickens join soccer teams.
A big farmer lifts a large load.
A cheap theatre feels very fake.
A confident guy viewed a natural scene.
A fair judge gives a second chance.
The happy girl feeds the thirty turtles.
A hundred ships took a dangerous trip.
A new plan makes John nervous.
Nervous George saw a violent killer.
The requested video played a zillion times.
Shy Jacob plays a boring game.
A thin lady taught a musical language.
Three fans visited a famous stadium.

Semantically Anomalous Sentences (based on Mack et al., 1990)

Bloody Zack costs a north zone.
A broken vacation turns kingly thighs.
A dark nail zaps a ready reason.

Fake Charles finished the happy garage.
Gentle Bill charms a shapely gas.
Hairy Rob serves monthly chairs.
Long Fred throws a lazy town.
Meaty Tim holds a super zero.
A paper nature sees the cool master.
A shapely shop gives a healthy fire.
The total runner judged a short fact.
Zesty jails paint childish vegetables.
A bookish day painted a long thinker.
A Chinese park sings a metal diet.
Dear victor lets rude thighs.
A funny gift jumped a tiny thinker.
A great shoe packs a kind death.
Healthy Ned types a solid rat.
A loud kiss shapes a dead ball.
Normal Mike sings a Japanese value.
A purple voice feels a common box.
A thin jailer chose a major salt.
A valued church counts a zinc purpose.
A zigzag shampoo hides a roomy vitamin.