Second language comprehensibility revisited: Investigating the effects of learner background

Abstract

The current study investigated first language (L1) effects on listener judgment of comprehensibility and accentedness in second language (L2) speech. The participants were 45 university-level adult speakers of English from three L1 backgrounds (Chinese, Hindi, Farsi), performing a picture narrative task. Ten native English listeners used continuous sliding scales to evaluate the speakers' audio recordings for comprehensibility, accentedness, as well as 10 linguistic variables drawn from the domains of pronunciation, fluency, lexis, grammar, and discourse. While comprehensibility was associated with several linguistic variables (segmentals, prosody, fluency, lexis, grammar), accentedness was primarily linked to pronunciation (segmentals, word stress, intonation). The relative strength of these associations also varied as a function of the speakers' L1, especially for comprehensibility, with Chinese speakers influenced chiefly by pronunciation variables (segmental errors), Hindi speakers by lexicogrammar variables, and Farsi speakers showing no strong association with any linguistic variable. Results overall suggest that speakers' L1 plays an important role in listener judgments of L2 comprehensibility and that instructors aiming to promote L2 speakers' communicative success may need to expand their teaching targets beyond segmentals to include prosody-, fluency-, and lexicogrammar-based targets.

Keywords: Comprehensibility; accentedness; L1 Influence; pronunciation learning

Second language comprehensibility revisited: Investigating the effects of learner background

With languages such as English, Spanish, Arabic, or Chinese becoming ever more prominent in international trade, education, and popular culture, particularly in communication among non-native speakers, understanding various components of second language (L2) speaking ability emerges as an important goal for both language researchers and teachers. One key component of speaking ability is pronunciation, which has typically been discussed with reference to two broad constructs, namely, understanding and nativelikeness (see Derwing & Munro, 2009; Levis, 2005). Understanding embraces various aspects of speakers' ability to make themselves understood. Following common research and assessment practice, understanding is often measured as comprehensibility or listeners' perception of how easy or difficult it is for them to understand L2 speech, rated on 7- or 9-point scales. Nativelikeness, which broadly refers to speakers' ability to approximate speech patterns of the target-language community, is usually operationalized as a listener-based rating of accentedness, also using 7- or 9-point scales.

Although researchers have consistently underscored comprehensibility as a more realistic goal for ensuring communicative success, compared to accent reduction or nativelikeness (e.g., Derwing & Munro, 2009; Levis, 2005), there still remains a need to distinguish how various aspects of L2 speech (at the level of phonology, fluency, lexis, grammar, or discourse) feed into comprehensibility and how they impact accentedness. Previous research has shown that while accent is linked primarily to phonology- and fluency-based characteristics of L2 speech, comprehensibility is additionally linked to grammatical and lexical variables (e.g., Saito, Trofimovich, & Isaacs, forthcoming, in press; Trofimovich & Isaacs, 2012). However, despite the fact that several (if not most) theoretical perspectives in L2 pronunciation learning ascribe an important role to speakers' first language (L1) in determining the rate and ultimate success of learning (e.g., Eckman, 2004; Escudero & Boersma, 2004; Flege, 2003), previous research on comprehensibility and accent has paid little attention to speakers' L1. In fact, most current evidence on L2 speech rating comes from studies that have either treated speakers of various L1s as a single group (e.g., Derwing, Munro, & Wiebe, 1998) or focused only on a single L1 group (e.g., Munro & Derwing, 1999). This limitation makes it unclear to what extent the linguistic variables that feed into comprehensibility and accentedness are specific to the speaker's L1. Therefore, the main objective of this study was to clarify the relationship between comprehensibility and accentedness, investigating the effect of speakers' L1 on listener perception of L2 comprehensibility and accentedness.

Disentangling Comprehensibility from Accent

For many L2 speakers and their teachers, the ideal ultimate learning goal is often to acquire the linguistic ability of a native speaker, characterized by native or near-native accent (Tokumoto & Shibata, 2011). However, adult speakers rarely pass for native speakers (Bongaerts, van Summeren, Planken, & Schils, 1997), so accented L2 speech is generally seen as normal and often unavoidable, even for speakers who begin learning at an early age (Flege, Munro, & MacKay, 1995). Considering the difficulty of acquiring nativelike L2 speech, adopting a more realistic learning goal has been encouraged, with a particular focus on comprehensibility or ease of understanding (Derwing & Munro, 2009; Levis, 2005). Indeed, even a heavy L2 accent does not preclude speakers from being highly comprehensible (Munro & Derwing, 1999). A focus on comprehensibility also seems sensible from a practical perspective, given that the interlocutor's goal in most real-world contexts is to get their message across rather than to pass

for native speakers. Thus, to make informed decisions about future learning goals and to address these goals through instruction, L2 speakers and their teachers need to know which aspects of language contribute to comprehensible speech and which are tied to foreign accent.

A focus on comprehensibility (rather than accentedness) is also motivated from a theoretical standpoint. For instance, the Interaction Hypothesis (e.g., Long, 1996) posits that language learning primarily takes place during communication breakdowns in conversations involving L2 speakers. These breakdowns often lead to negotiation for meaning, in which interlocutors make an effort to repair communication through the use of such discourse moves as clarification requests or confirmation checks. According to the Interaction Hypothesis, negotiation for meaning facilitates L2 development by promoting speakers' attention to various linguistic dimensions which may have caused a communication breakdown (see Mackey & Goo, 2007). And because communication breakdowns occur as a result of some linguistic dimensions more than others (Mackey, Gass, & McDonough, 2000), those dimensions tied to comprehensibility, rather than those that are uniquely linked to accentedness, will be more beneficial in helping learners notice and repair their nontarget production. Thus, to understand which linguistic dimensions of speech are beneficial for development, it is necessary to distinguish the dimensions feeding into comprehensibility from those uniquely tied with accent.

Previous research examining linguistic influences on listener perception of L2 speech has primarily targeted phonology and fluency. When it comes to understanding, for example, stress (Field, 2005), speech rate (Munro & Derwing, 2001), as well as pitch range and pause or syllable length (Kang, Rubin, & Pickering, 2010; Winters & O'Brien, 2013) have all been shown to influence how listeners extract meaning from an utterance. Although there is little research focusing on domains other than phonology and fluency, poor grammar and inappropriate lexical choice also appear to compromise listener understanding (Fayer & Krasinski, 1987). With respect to accentedness, segmental accuracy (Derwing et al., 1998), pausing and articulation rate (Trofimovich & Baker, 2006), and various suprasegmental measures such as pitch range, stress, and pause length (Kang, 2010) have been linked to L2 accent. In sum, listener judgments of L2 speech, which include comprehensibility and accent, are tied to many overlapping linguistic measures from the domains of phonology, fluency, grammar, and lexis.

Moving away from a focus on individual linguistic variables, researchers have recently begun to investigate their combined contribution to listener judgment. For instance, Trofimovich and Issacs (2012) analyzed the speech of 40 L1 French speakers of English targeting 19 coded linguistic measures (divided into phonology, lexis/grammar, fluency, and discourse categories), with the goal of identifying links between these measures and both comprehensibility and accentedness. Comprehensibility was best explained using word stress, type frequency (a measure of lexical richness), and grammar accuracy, while word stress and rhythm best defined accent. A follow-up study, in which the speech of the same speakers was rated for 11 linguistic measures (Saito, Trofimovich, & Isaacs, forthcoming), revealed again that comprehensibility was associated with many variables, including pronunciation (word stress, speech rate, rhythm), lexis, and grammar, but that accentedness was mainly linked to pronunciation (segmental errors, word stress). A study targeting 120 L1 Japanese speakers of English similarly showed that comprehensibility was tied to segmental, prosodic, temporal, lexical, and grammatical aspects of speech, while accentedness was related to pronunciation, especially segmental accuracy and word stress (Saito, Trofimovich, & Isaacs, in press). Thus, comprehensibility seems to encompass a range of linguistic dimensions while accentedness involves mainly pronunciation and fluency factors.

L1 Effects on L2 Comprehensibility and Accentedness

One issue that remains to be resolved concerns the extent to which linguistic correlates of comprehensibility and accentedness are specific to speakers' L1 background. On the one hand, the findings reviewed above indicate that comprehensibility and accentedness are distinct constructs, with comprehensibility associated with a broader range of variables. On the other hand, these results show that several variables (e.g., word stress, rhythm) contribute to both comprehensibility and accentedness for speakers from different L1 backgrounds. This raises a question of whether the linguistic variables linked to comprehensibility and accentedness are unique to the particular L1 groups targeted in prior research, or whether at least some of these variables apply to L2 speakers from many L1 backgrounds. From a theoretical perspective, there is considerable evidence of L1 effects on L2 development, especially in the realm of pronunciation. Such evidence spans decades of research, starting from early attempts to describe L1 influences on L2 pronunciation as a perceptual "sieve" biasing learners (Trubetzkoy, 1939), to the Contrastive Analysis Hypothesis used to predict and explain speech patterns that may cause particular learning difficulties (Lado, 1957), to more recent conceptualization of L1 effects, including Eckman's (2004) structural conformity hypothesis, Escudero and Boersma's (2004) optimality-theoretic model, and Flege's (2003) speech learning model. Indeed, there appears to be little debate that L1 influence plays a significant role in L2 pronunciation learning (Eckman, 2004), which implies that linguistic correlates of comprehensibility and accentedness might in fact be specific to speakers' linguistic background.

Compared to the vast literature documenting L1 influences on the perception and production of specific aspects of L2 phonology, such as segmental contrasts or voice-onset time (e.g., Davidson, 2011), there is little research exploring L1 effects on listener ratings of L2 comprehensibility and accentedness. The majority of studies considering rater perception of L2 speech have either focused solely on L2 speakers from a single L1 group (e.g., Winters & O'Brien, 2013) or conflated multiple L1s into a single group (e.g., Kang et al., 2010), and the few studies that have compared different L1 groups have yielded mixed findings. For instance, Anderson-Hsieh, Johnson, and Koehler (1992) found that ratings of speech prosody had a strong positive correlation with speakers' L2 pronunciation scores regardless of their L1 background, whereas segmental and syllable structure errors were dependent on speakers' L1. In terms of other prosodic factors, Baker et al. (2011) reported that word duration and word reduction patterns were negatively associated with accentedness ratings for both Chinese and Korean speakers of English. In contrast, Kang (2010) identified Chinese and Japanese speakers over speakers of other L1s (i.e., Arabic, Russian, Hindi) as having strong L2 accents due to frequent, inappropriate word emphasis. And in a longitudinal study comparing speech ratings of Mandarin and Slavic L2 learners of English, Derwing, Munro, and Thomson (2008) found that, while both groups had equal proficiency levels at the outset of the study, only the Slavic group improved over time, implying that there could be a possible L1 transfer effect benefiting the Slavic speakers. This limited evidence thus identifies a pressing need to consider how speakers' L1 background affects listener judgments of L2 speech, a point which will be crucial in enabling learners and their teachers to set appropriate learning goals.

Research Objectives

In sum, there is growing research interest in identifying linguistic influences on comprehensibility (understanding) and distinguishing such influences from those tied to accent (nativelikeness). This study contributes to this research agenda by investigating L1 background effects on the relationship between various linguistic dimensions of L2 speech and both

comprehensibility and accentedness. The study had two specific objectives: (a) to further clarify which linguistic variables in L2 speech contribute to listener perception of comprehensibility and accentedness and (b) to determine whether and to what degree the relative contributions of these linguistic variables remain generally problematic across a range of speakers or differ as a function of their L1 background.

Method

Participants

Speakers. The L2 participants were 45 speakers from an unpublished corpus of L2 speech which included audio recordings by 143 speakers from 19 different L1s completing five speaking tasks (Isaacs and Trofimovich, 2011). The speakers, who at the time of the recording were international students in undergraduate (8) and graduate (37) programs at an English-medium university in Montreal, Canada, were organized in three groups (n = 15) based on their L1 background (Chinese, Hindi/Urdu, Farsi). The speakers of Hindi and Urdu were combined into one group because the principal difference between these languages is script-based (King, 1994). All speakers in the Chinese group spoke Mandarin as their L1. The Farsi, Hindi/Urdu, and Chinese groups represented the three largest cohorts in the corpus, with a total of 32, 17, and 15 speakers, respectively.

In creating the final groups, the speakers in the two larger cohorts (Farsi, Hindi/Urdu) were matched as much as possible to the Chinese speakers for several background variables (see Table 1). The only exception was the 14:1 male-female ratio in the Hindi/Urdu group, which reflected the gender composition of Hindi/Urdu speakers in the larger university community. All speakers, who were within the first term of their studies and had recent TOEFL and IELTS test scores, represented a comparable level of L2 oral ability. They had all demonstrated at minimum a speaking score of 17 for TOEFL iBT or 5 for IELTS, which was considered sufficient for them to pursue academic degrees. According to one-way ANOVAs, there were no significant differences between the three groups in their TOEFL and IELTS total scores, Fs < 1.49, p > .25, or in listening and speaking subscores, Fs < 3.19, p > .06.

TABLE 1

The resulting three L1 groups were considered to provide an appropriate comparison of possible L1 effects on accent versus comprehensibility because these groups represented typologically different languages, belonging to the Sino-Tibetan language family (Chinese) or to Indo-Aryan (Hindi/Urdu), and Iranian (Farsi) sub-branches of the Indo-European language family. The three L1s also crucially differ in their segmental inventories (e.g., Duanmu, 2007; Shackle, 2001; Wilson & Wilson, 2001) as well as prosody, particularly in terms of rhythm, thus allowing for direct comparisons between the speakers of syllable-timed French tested by Trofimovich & Isaacs (2012) and the speakers of non-Romance syllable-timed Hindi (Shackle, 2001), stress-timed Farsi (Jun, 2005), and tonal Chinese (Jun, 2005).

Raters. The raters, who were educated entirely in English, included 10 native English speakers (M_{age} = 32.7 years, SD = 10.2) born and raised in English-speaking homes with at least one native English-speaking parent (with seven reporting both parents as native speakers). The raters, who resided in Montreal (a bilingual French-English city), reported speaking English on average 89% of the time (SD = 8.8), interacting with native English speakers 73% of the time (SD = 14.9), and listening to English media 85% of the time (SD = 13.5) daily. The raters had on average 6.6 years of L2 teaching experience (1-23) and were either enrolled in (9) or recently completed (1) their graduate studies (7 MA, 3 PhD) in applied linguistics at a local English-medium university. Because listeners' familiarity with L2 speech can impact their judgments

(Bent & Holt, 2013; Winke & Gass, 2013), only the raters who reported high familiarity with accented English were selected. Using a 9-point scale (1 = "not at all familiar", 9 = "very familiar"), the raters, who were members of the same university community, reported high familiarity with accented English (M = 8.6, SD = .7), including L2 speech by speakers of the target languages (i.e., Farsi, Hindi, Chinese). Raters with linguistic and teaching backgrounds were chosen because Saito et al. (forthcoming) showed that experienced raters, compared to inexperienced ones, were more consistent in evaluating complex and less intuitive linguistic variables in a similar rating task.

Materials

As part of the original corpus, each speaker completed five speaking tasks but only the picture narrative task was chosen for analysis in this study because it was the same task used in Trofimovich & Isaacs (2012), which allowed for direct comparisons of findings. In the picture narrative task, speakers were presented with an eight-frame colored picture story featuring two strangers bumping into each other while rounding a corner, then accidently exchanging their identical suitcases, and finally realizing their mistake upon returning home (Derwing et al., 2008). All narratives were recorded directly onto a computer using a Plantronics (DSP-300) microphone, stored as digital audio files, and then normalized by matching peak amplitude across files. For each recording, all fillers and false starts at the beginning of the file were removed before it was edited down to the initial 30 s, in line with previous research using 20-60 s recordings to evaluate L2 speech (e.g., Derwing et al., 1998, 2008). All samples were also transcribed by a trained research assistant and subsequently verified. The audio recordings and transcripts served as the stimuli for judgments of accent and comprehensibility as well as for linguistic coding using 10 rated categories spanning the dimensions of phonology, fluency, lexis, grammar, and discourse.

Speech Rating

All ratings were collected as part of a larger project evaluating speaker performance in three speaking tasks, including the picture narrative task. The project involved four individual 2 hour sessions, all occurring within three weeks of each other, during which each rater evaluated audio recordings or transcripts blocked by task in a counterbalanced order (e.g., Task 1-2-3, 2-3-1, etc.), with audio recordings or transcripts presented in a unique randomized order. Session 1 was devoted to providing global judgments of accent and comprehensibility based on audio recordings. Session 2 and part of Session 3 were dedicated to rating audio recordings for five phonology- and fluency-based categories. The remainder of Session 3 and Session 4 were spent evaluating orthographic transcripts for five lexical, grammatical, and discourse categories.

All ratings were carried out using a computer-based scale developed by Saito et al. (forthcoming), with each measure evaluated on a 1000-point continuous sliding scale and endpoints clearly marked on a horizontal plane. The scales were run through the MATLAB software, and the raters used a free moving slider on a computer screen to assess each category. The rating was recorded as "0" if the slider was placed at the leftmost (negative) end of the continuum, marked with a frowning face. The rating was recorded as "1000" if the slider was set at the rightmost (positive) end, marked with a smiley face. The slider initially appeared in the middle (rating of 500), and the raters were informed that even a small movement of the slider may represent a fairly large difference in the rating. Apart from brief verbal descriptions for the endpoints of each category and the frowning and smiley faces to indicate the directionality of the scale, no numerical labels or marked intervals were included in the scale.

At the beginning of each session, the raters were seated in front of a personal laptop displaying the rating interface and received training on the relevant categories and use of the scale. Each was supplied with a written description of each measure, including examples illustrating the scalar endpoints, and was given the opportunity to discuss each measure with the researcher (for all training materials and onscreen labels, see Online Supporting Documentation). The raters then performed four practice judgments by listening to audio recordings or viewing transcripts, using the appropriate scales, and the rationale for each judgment was discussed with the researcher after each individual practice rating, to ensure that each measure had been accurately understood. The raters were informed that recordings were 30 s in duration, with the possibility that some speakers may have been cut off in the middle of a phrase, but should not be penalized for this when being rated.

Rated Categories

Accent and comprehensibility. Following Trofimovich and Isaacs (2012), accent was defined as raters' perception of how different the speaker sounded from a native speaker of North American English (1 = "heavily accented", 1000 = "no accent at all"), while comprehensibility was defined as the degree of ease or difficulty in raters' understanding of L2 speech (1 = "hard to understand", 1000 = "easy to understand"). Comprehensibility (rather than intelligibility) was chosen as the measure of understanding as it reflects a more typical and practical approach to measuring understanding in a variety of assessment contexts (such as oral proficiency scales) and in research settings (Isaacs & Trofimovich, 2012). Consistent with previous research on listener-based ratings of accent and comprehensibility, the raters were allowed to listen to each recording only once before making their judgment, which ensured that the ratings were comparable across studies.

Phonology and fluency. The raters evaluated each audio recording for the following five segmental, prosodic, and temporal categories (described and illustrated in full in Online Supporting Documentation):

1. Segmental errors (1 = "frequent", 1000 = "infrequent or absent"), defined as errors in the pronunciation of individual consonants and vowels within a word (e.g., *dat* instead of *that*; *pin* instead of *pen*), as well as any segments erroneously deleted from or inserted into words (e.g., *'ouse* instead of *house*; *supray* instead of *spray*).[1]
2. Word stress errors (1 = "frequent", 1000 = "infrequent or absent"), defined as errors in the placement of primary stress (e.g., *com-pu-TER* instead of *com-PU-ter*, where capitals designate primary stress) or the absence of discernible stress, such that all syllables receive equal prominence (e.g., *com-pu-ter*).
3. Intonation (1 = "unnatural", 1000 = "natural"), defined as appropriate pitch moves that occur in native speech, such as rising tones in yes/no questions (e.g., Will you be home tomorrow↑) or falling tones at the end of statements (e.g., Yeah, I'll stay at home↓).
4. Rhythm (1 = "unnatural", 1000 = "natural"), defined as the difference in stress (emphasis) between content and function (grammatical) words. For instance, in the sentence "They RAN to the STORE", the words "ran" and "store" are content words and therefore are stressed more than the words "they", "to", and "the", which are grammatical words featuring reduced vowels.
5. Speech rate (1 = "too slow or too fast", 1000 = "optimal"), defined as a speaker's overall pacing and the speed of utterance delivery.

The judgments of phonology and fluency likely require an in-depth analysis of the speech signal. Therefore, to ensure rating quality, raters had the option to listen to the same speech sample multiple times until they felt satisfied with their judgment.

Lexis, grammar, and discourse structure. To remove pronunciation and fluency as possible confounds in judgments of lexis, grammar, and discourse, the raters evaluated written transcripts of the audio files (Crossley, Salsbury, & McNamara, 2014). The transcripts had been modified to remove hesitation markers (e.g., um, uh), spelling clues signaling pronunciation-specific errors (e.g., *when,* although pronounced as *ven,* was spelled as *when*), and punctuation to avoid transcriber influence (Ochs, 1979). The raters evaluated written transcripts for the following five lexical, grammatical, and discourse categories (described and illustrated in full in Online Supporting Documentation):

6. Lexical appropriateness (1 = "many inappropriate words used", 1000 = "consistently uses appropriate vocabulary"), defined as the speaker's choice of words to accomplish the task. Poor lexical choices include incorrect, inappropriate, and non-English words (e.g., "A man and a woman bumped into each other on a walkside").
7. Lexical richness (1 = "few, simple words used", 1000 = "varied vocabulary"), defined as the sophistication of the vocabulary used by the speaker. Simple words with little variety correspond to poor lexical richness (e.g., "The girl arrived home her dog was happy she arrived home", compared to "The girl arrived home to find her dog overjoyed at her return").
8. Grammatical accuracy (1 = "poor grammar accuracy", 1000 = "excellent grammar accuracy"), defined as the number of grammar errors made by the speaker. Examples included errors of word order (e.g., "What you are doing?"), morphology (e.g., "She go to school every day"), and agreement (e.g., "I will stay there for five day").
9. Grammatical complexity (1 = "simple grammar", 1000 = "elaborate grammar"), defined as the sophistication of the speaker's grammar. Grammatical complexity is low if the speaker uses simple, coordinated structures without embedded clauses or subordination (e.g., "The man wore a black hat and he enjoyed his coffee", compared to "The man that was wearing a black hat was enjoying his coffee").
10. Discourse richness (1 = "simple structure, few details", 1000 = "detailed and sophisticated"), defined as the richness and sophistication of the utterance content. Discourse richness is low if the entire narrative is simple, unnuanced, bare, and lacks sophisticated ideas or details, but high if the speaker produces several distinct ideas or details so that the statement sounds developed and sophisticated.

As was the case with phonology and fluency judgments, the raters were allowed to spend as much time as needed with each transcript to allow for accurate judgments.

Understanding and use of rated categories. Upon completion of each set of ratings, the raters used 9-point scales to assess the extent to which they understood the categories (1 = "I did not understand at all", 9 = "I understand this concept well") and to which they could comfortably and easily use them (1 = "very difficult", 9 = "very easy and comfortable"). The raters indicated that they could understand all categories well (M = 8.3; SD = .5) and could use them easily (M = 7.8; SD = .9).

Results

Rater Consistency

The 10 raters were overall consistent in their global judgments, revealing high reliability indexes (Cronbach's alpha) for accent (a = .93) and comprehensibility (a = .86). Therefore, mean

accent and comprehensibility scores were calculated for each speaker by averaging across all listener ratings. Although the perception of individual linguistic categories is presumably less intuitive and more complex, compared to accent and comprehensibility, the raters were nevertheless fairly consistent, demonstrating reliability indexes that exceeded the benchmark value of .70-.80 (Larson-Hall, 2009) for pronunciation ($a_{\text{segmentals}}$ = .92; $a_{\text{word stress}}$ = .78; $a_{\text{intonation}}$ = .78; a_{rhythm} = .85), fluency ($a_{\text{speech rate}}$ = .90), vocabulary ($a_{\text{appropriateness}}$ = .80; $a_{richness}$ = .88), grammar ($a_{\text{acccuracy}}$ = .80; $a_{complexity}$ = .89), and discourse ($a_{richness}$ = .90). One adjustment was made to lexical appropriateness due to a low corrected item-total correlation (.21) specific to one rater by removing this rater's data ($a_{\text{appropriateness}}$ = .81). The raters' scores were therefore considered sufficiently consistent and were averaged across the 10 raters (nine for lexical appropriateness) to derive a single mean score per speaker for each rated category.

Comprehensibility and Accent

The first analysis examined possible group-based differences in global ratings (shown in Table 2). For this, comprehensibility and accentedness ratings were submitted to a two-way ANOVA with group (Chinese, Hindi/Urdu, Farsi) as a between-subjects factor and perceptual judgment (comprehensibility, accentedness) as a within-subjects factor. The ANOVA yielded a significant main effect of group, $F(2, 42) = 13.75$, $p < .0001$, $\eta_p^2 = .40$, a significant main effect of perceptual judgment, $F(1, 42) = 53.30$, $p < .0001$, $\eta_p^2 = .56$, and a significant two-way interaction, $F(2, 42) = 3.22$, $p = .05$, $\eta_p^2 = .13$. Tests of interaction effects (Bonferroni adjusted $a = .006$) further showed that the Farsi ($p < .0001$) and Hindi/Urdu ($p < .0001$) groups, but not the Chinese group, were rated higher in comprehensibility than in accent, with large effect sizes (Cohen's d = .92-1.22). Tests of interaction effects also revealed that the Chinese group was rated as being more accented than the Farsi ($p = .001$) group, with a large effect size ($d = 1.35$), and that the Chinese group was rated as being less comprehensible than the other two groups ($p < .0001$), with large effect sizes (d = 1.68-2.12). In sum, the three groups differed in comprehensibility and accentedness, with the Chinese group rated as more accented than the Farsi group and as less comprehensible than the remaining two groups. No differences existed between the Farsi and Hindi speakers.

TABLE 2

Linguistic Categories

The next analyses investigated how the global speech ratings of comprehensibility and accentedness related to the 10 rated linguistic categories. First, the linguistic scores for all speakers were submitted to an exploratory Principal Component Analysis (PCA) with Oblimin rotation to determine if the 10 rated linguistic categories showed any underlying patterns based on their clustering. Despite a relatively low sample size ($N = 45$), the Kaiser-Meyer-Oklin value was .82, exceeding the required .60 for sampling adequacy and indicating excellent factorability of the correlation matrix (Hutcheson & Sofroniou, 1999). In addition, a significant Bartlett's test of sphericity, $\chi^2(45) = 501.13$, $p < .0001$, showed that the correlations between the categories were sufficiently large for PCA. As shown in Table 3, the PCA revealed two factors accounting for 81.2% of total variance. Factor 1, labeled "Pronunciation", consisted of the four pronunciation categories, plus speech rate (with the relevant variables intercorrelated at $r(58)$ = .54-.87); Factor 2, labeled "Lexicogrammar", consisted of all vocabulary, grammar, and discourse-level categories and speech rate (with all relevant variables intercorrelated at $r(58)$ = .46-.95). Thus, the 10 linguistic categories patterned along two dimensions (pronunciation and lexicogrammar). Speech rate was common to both, suggesting that both pronunciation and lexicogrammar are linked to fluency (Segalowitz, 2010).

TABLE 3

The resulting pronunciation and lexicogrammar PCA scores, derived through the Anderson-Rubin method of obtaining non-correlated factor scores, were then used as predictor variables in two separate stepwise multiple regression analyses to examine the contribution of pronunciation and lexicogrammar to accent and comprehensibility. While the two regression models accounted for a similar amount of total variance (64% for accent, 70% for comprehensibility), the ratio explained by the two factors differed (Table 4). The variance in accent was entirely explained by the pronunciation factor (64%), while both the lexicogrammar (49%) and pronunciation (21%) factors contributed to comprehensibility.

TABLE 4

L1 Background

The next analysis examined whether the relationship between linguistic categories and speech ratings depended on the speakers' L1. Pearson correlation coefficients were computed separately for each group between the two PCA factor scores (pronunciation and lexicogrammar) and comprehensibility and accentedness ratings (a = .0025). As shown in Table 5, accentedness was linked to pronunciation for all L1 groups, while the relationship between comprehensibility and the two factor scores differed as a function of group. Comprehensibility was associated with pronunciation for the Chinese group, with lexicogrammar for the Hindi/Urdu group, and with neither factor for the Farsi group.

TABLE 5

The final analysis further explored which linguistic dimensions were associated with comprehensibility and which were linked to accentedness, separately for each L1 group. Pearson correlation coefficients were computed for each L1 group between the 10 linguistic categories and comprehensibility and accentedness (a = .0025). As summarized in Table 6, accentedness was associated with various rated pronunciation (but not lexicogrammar) categories for all L1 groups, from segmental issues for the Chinese group, segmental issues in combination with intonation and word stress for the Hindi/Urdu speakers, and a combination of segmental issues and word stress for the Farsi speakers. In contrast, comprehensibility was linked uniquely to segmental issues for the Chinese group but to a variety of lexicogrammar variables (lexical appropriateness and richness, grammatical complexity and discourse richness) for the Hindi/Urdu group. Confirming an earlier finding, no rated category showed a strong association with comprehensibility for the Farsi speakers.

TABLE 6

Discussion

Conceptualized as an investigation of L1 influences on the relationship between linguistic dimensions of speech and listener ratings of comprehensibility and accentedness, this study showed that comprehensibility and accentedness are overlapping yet distinct constructs and that linguistic dimensions feeding into comprehensibility vary as a function of speakers' L1. While accentedness was uniquely linked to pronunciation (especially segmental errors) for all L1 groups, linguistic correlates of comprehensibility depended on the speakers' L1.

Linguistic Variables and Accentedness

When it comes to accent, the relationship between pronunciation variables and accentedness ratings is not surprising in light of previous research identifying pronunciation- and fluency-based influences on accent (e.g., Trofimovich & Baker, 2006; Derwing et al., 1998; Kang, 2010). For all L1 groups, segmental errors represented the strongest influence, consistent with previous studies targeting L1 speakers of French (Saito et al., forthcoming) and Japanese

(Saito et al., in press). Segmental errors may thus be particularly salient to the listener, regardless of L1 background, making the greatest contribution to accent. However, beyond similarities in segmental influence, each L1 group differed in which of the four remaining pronunciation variables were associated with their accent. For the Farsi group, word stress had the second strongest link with accent, likely due to the more predictable stress in Farsi, which generally stresses word-final syllables (Wilson & Wilson, 2001), leading to a possible overgeneralization in stress placement for these speakers. For the Hindi-Urdu speakers, after segmentals, intonation followed by word stress had the strongest relationships with accent, perhaps due to Hindi/Urdu speakers' use of substantially raised pitch to indicate stress in their native language (Shackle, 2001), compared to English speakers' heavier articulation in stressed syllables using greater energy through a combination of pitch, volume, and duration (Ladefoged & Johnson, 2014). And for the Chinese speakers, segmental errors (i.e., substitutions, omissions), which reflect these speakers' well-documented challenge with English segments and syllable structure (e.g., Chang, 2001; Rau, Chang, & Tarone, 2009), were likely so salient that they made it difficult for listeners to isolate other potential prosodic and fluency-based influences on accent.

Thus, listener perception of L2 accent appears to be the result of a complex interaction between several pronunciation and fluency variables. Listeners seem to prioritize segmentals regardless of speaker background but also consider relative weights of other pronunciation variables, consistent with the idea that rating accent primarily requires listeners to attend to how speech sounds rather than to what meaning it conveys. In this dataset, the Chinese speakers represented one possible endpoint in a perceptual weighting continuum of this kind, with listener judgments of accentedness tied solely to segmental errors. The Hindi/Urdu and Farsi groups fell further along this continuum, with one (Farsi) or two (Hindi/Urdu) variables beyond segmentals (word stress, intonation) having a strong link to accent judgments. Presumably, for other L1 groups, accentedness may be associated with a combination of segmental errors and several other pronunciation and fluency variables, as was the case for the French speakers in Trofimovich and Isaacs' original study (2012). Future research needs to explore how exactly listeners weigh multiple cues in making seemingly effortless but reliable accent ratings.

Linguistic Variables and Comprehensibility

With respect to comprehensibility, which arguably represents a more realistic L2 learning goal compared to accent reduction (Derwing & Munro, 2009), there were clear L1 background effects. For the Hindi/Urdu group, the relationship between linguistic variables and comprehensibility was restricted to lexicogrammar, suggesting that listener-rated ease of understanding was based on these speakers' lexical, grammatical, and discourse-based choices, rather than on the quality of their pronunciation, fluency, or prosody. This finding underscores an observation that speakers of East Indian languages, while being accented, are often proficient (but not necessarily nativelike) English users (Smith, 1992), and shows the importance of lexicogrammar to comprehensibility for speakers who have little difficulty with L2 segments and prosody in a way that is consequential for comprehensibility. In essence, according to listeners, the comprehensibility of the Hindi/Urdu speakers' speech was associated with these speakers' use of appropriate and rich vocabulary, complex grammar, and rich discourse structure, which suggests that the Hindi/Urdu speakers were not yet fully skilled in using these aspects of L2 speech, at least as far as comprehensibility was concerned. In contrast to the Hindi/Urdu speakers, the comprehensibility of the Chinese group was uniquely associated with pronunciation, particularly with segmental accuracy, likely due to substantial crosslinguistic distance between Chinese and English (Duanmu, 2007). Listeners may have prioritized

segmental accuracy over lexicogrammar because segmental substitutions and errors of syllable structure, such as deletion of coda consonants, are highly noticeable to English listeners and are more frequent for Chinese speakers (Anderson-Hsieh et al., 1992; Rau et al., 2009), compared to other speaker groups.

Finally, for the Farsi group, none of the 10 linguistic variables bore a strong relationship with comprehensibility. It may be that Farsi speakers, who represented the largest cohort in the targeted speech corpus (34/143 or 22%), were familiar to the raters as members of the university community, revealing familiarity effects on ratings (Bent & Holt, 2013; Winke & Gass, 2013). It could also be that for some speakers, such as Farsi speakers here, comprehensibility ratings may be based on a range of variables, with no single factor bearing a particularly strong relationship with comprehensibility. Alternatively, while the Farsi speakers did not differ from the remaining groups in their overall speaking proficiency (as measured through the IELTS and TOEFL tests), they may have been nevertheless more proficient than the other speakers in those aspects of L2 speech that are consequential for comprehensibility. As shown in Table 2, the Farsi speakers were rated as being the most comprehensible and least variable in terms of comprehensibility scoring outcomes. Finally, it is possible that the discrete listener-based perceptual measures that we examined were not sensitive enough to detect strong linguistic ties with comprehensibility for Farsi speakers and that more fine-grained measures might be needed to isolate specific linguistic sources of variability in this group's comprehensibility performance.

Overall, these findings reinforce the view that speakers' L1 plays a role in comprehensibility, extending previous literature on L1 effects in learning pronunciation (e.g., Eckman, 2004). As discussed previously, these findings illustrate several distinct, L1-specific patterns of linguistic influences on comprehensibility: from individual pronunciation effects for the Chinese speakers, to lexicogrammar effects for the Hindi/Urdu speakers, to weak or potentially "distributed" effects of many variables for the Farsi speakers, to combined pronunciation and lexicogrammar effects for the French speakers in Trofimovich and Issacs' initial comprehensibility study (2012). However, a thorough understanding of linguistic influences on comprehensibility – particularly as suggested by the findings for the Farsi group – would certainly require a more nuanced approach, one that besides speakers' L1 background also considers their L2 proficiency level. In fact, in a study targeting 120 Japanese speakers of English with a varying degree of immersion experience in Canada, who represented a wide range of speaking ability, Saito et al. (in press) showed that comprehensibility varied as a function of learners' assessed oral production level (beginner, intermediate, advanced). Briefly, for Japanese speakers, word stress and intonation were equally important for comprehensibility at all ability levels, from beginner to advanced. Attaining a minimum level of segmental accuracy, fluency, lexical appropriateness, and grammatical accuracy was relatively important for low comprehensibility speakers, while segmental precision and grammatical accuracy characterized the highest skill level. It appears, then, that the particular relationships between comprehensibility and linguistic aspects of L2 speech not only might be determined by speakers' L1 background, as shown here, but also might be specific to their proficiency level. It would thus be important to address this issue in future research, with the goal of identifying linguistic correlates of comprehensibility for speakers at different ability levels across the comprehensibility continuum.

Implications and Conclusions

The findings of this study point to two broad conclusions. The first is that, when judging comprehensibility, listeners consider not only pronunciation- and fluency-related aspects of L2

speech, which contribute to the perception of accent, but also grammatical, lexical, and discourse-based variables. In fact, the bulk of shared variance in the speakers' comprehensibility ratings in this study was predicted by the lexicogrammar rather than the pronunciation factor (49% vs. 21%). Although listeners might react to lexicogrammar factors differently when evaluating written transcripts (which was the case here), compared to assessing lexicogrammar in speech, both lexicogrammar and pronunciation variables seem to determine the time and effort needed for listeners to extract meaning from L2 speech. This is in contrast to listener judgment of accent, where ratings appear to be invariably fast, effortless, and intuitive, mostly driven by form-based aspects of speech (Saito et al., in press; Munro, Derwing, & Burgess, 2010). The second conclusion is that linguistic influences on comprehensibility depend on speakers' L1. Comprehensibility, at least for some L2 university students, is broader than simply pronunciation to the ears of the listener.

Although implications of these findings are premature, there are nevertheless several suggestions for researchers and teachers. First, there appears to be no single linguistic variable universally predictive of comprehensibility for speakers from different L1 backgrounds. A given speaker's comprehensibility is likely shaped by L1 background effects interacting with speaker's overall level of proficiency given particular demands of a speaking task. This interaction needs to be explored in future research with speakers from different proficiency levels engaged in various speaking tasks, particularly focusing on contributions of lexicogrammar to comprehensibility in written versus spoken language. Second, researchers and teachers may wish to leave aside the debate contrasting segmentals versus suprasegmentals as preferred targets in pronunciation teaching (e.g., Derwing et al., 1998), since comprehensibility cuts across many linguistic variables, which include both segmentals and suprasegmentals but are not limited to them.

Finally, instructors teaching homogeneous groups of L2 learners may benefit from an understanding of the specific linguistic variables that impact their learners' comprehensibility and accentedness. And while instructors teaching learners from multiple L1 groups cannot take full advantage of L1-specific knowledge, they can still promote learners' communicative success by expanding their teaching targets beyond segmentals to include syllable structure, word stress, fluency phenomena, as well as aspects of grammar and lexis. This change in focus, along with balanced instruction on how fluency affects comprehensibility, how grammatical errors in L2 speech are perceived by listeners, and how lexical knowledge is linked to understanding, should promote communicative improvement. In essence, targeting L2 comprehensibility as a learning goal requires an eclectic, comprehensive approach sensitive to the variety of L1s in a language classroom.

Notes

1. An anonymous reviewer pointed out that the measure of segmental errors likely encompassed both non-native segmental substitutions as well as possible nonstandard or dialectal contextual realizations of English segments. Because such distinction was not drawn for the raters, it would be impossible to speculate as to the relative contribution of each error type to accentedness and comprehensibility. Disentangling non-native and dialectal influences on L2 speakers' speech is certainly a worthwhile topic of future research.

Acknowledgments

This study was funded by grants from the Social Sciences and Humanities Research Council of Canada (SSHRC) and Fonds de recherche sur la société et la culture (FRQSC) awarded to the second author, the Grant-in-Aid for Scientific Research in Japan (No. 26770202) awarded to the third author, and by a Marie Curie Career Integration Grant (European Commission) awarded to the fourth author. We are grateful to participants, Ze Shan Yao for his technical help, and to anonymous TESOL Quarterly reviewers for their helpful input and feedback on the content of this manuscript.

The Authors

Dustin Crowther is an MA graduate of Concordia University in Montreal, Canada, and is currently pursuing a PhD in Second Language Studies at Michigan State University. He has extensive experience teaching English in Canada and Japan. His research interests include the teaching and learning of second language pronunciation, cross-linguistic influence, and intercultural pragmatics.

Pavel Trofimovich is an associate professor of applied linguistics in the Department of Education at Concordia University, Montreal, Canada. His research focuses on cognitive aspects of second language processing, second language phonology, sociolinguistic aspects of second language acquisition, and the teaching of second language pronunciation.

Kazuya Saito is Assistant Professor of English at Waseda University in Japan. His program of research investigates how instruction and corrective feedback can help adult learners develop their L2 oral proficiency, especially in the domains of pronunciation, listening, vocabulary, and grammar.

Talia Isaacs is a senior lecturer in education at the University of Bristol, UK, where she serves as co-coordinator of the Centre for Assessment and Evaluation Research and director of the Second Language Speech Lab. Her research focuses on construct validity in pronunciation research, and, in particular, on how comprehensibility is operationalized in rating scales.

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. doi:10.1111/j.1467-1770.1992.tb01043.x

Baker, R. E., Baese-Berk, M., Bonnasse-Gahot, L., Kim, M., Van Engen, K. J., & Bradlow, A. (2011). Word durations in non-native English. *Journal of Phonetics, 39*, 1-17. doi:10.1016/j.wocn.2010.10.006

Bent, T., & Holt, R. F. (2013). The influence of talker and foreign-accent variability on spoken word identification. *The Journal of the Acoustical Society of America, 133*, 1677-1686. doi:10.112/1.4776212

Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, *19*, 447-465. doi:10.1017/S0272263197004026

Chang, J. (2001). Chinese speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (2nd ed) (pp. 310-324). Cambridge: Cambridge University Press.

Crossley, S. A., Salsbury, T., & McNamara, D. S. (2014). Assessing lexical proficiency using analytic ratings: A case for collocation accuracy. *Applied Linguistics*. doi:10.1093/applin/amt056

Davidson, L. (2011). Phonetic and phonological factors in the second language production of phonemes and phonotactics. *Language and Linguistics Compass, 5*, 126-139. doi:10.1111/j.1749-818X.2010.00266.x

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching, 42*, 476-490. doi:10.1017 /S026144480800551X

Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics, 29*, 359-380. doi:10.1093/applin/amm041

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48*, 393-410. doi:10.1111/0023-8333.00047

Duanmu, S. (2007). *The phonology of standard Chinese: The Phonology of the world's languages.* Oxford: Oxford University Press.

Eckman, F. (2004). From phonemic differences to constraint rankings: Research on second language phonology. *Studies in Second Language Acquisition, 26*, 513-549. doi:10.1017/S027226310404001X

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*, 551-585. doi:10.1017/S0272263104040021

Fayer, J. M., & Krasinski, E. (1987). Native and nonnative judgments of intelligibility and irritation. *Language Learning, 37*, 313-326. doi:10.1111/j.1467-1770.1987.tb00573.x

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39*, 399-423. doi:10.2307/3588487

Flege, J. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), Phonetics and phonology in language comprehension and production: Differences and similarities (pp. 319-355). Berlin: Mouton de Gruyter.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication, 16*, 1-26. doi:10.1016/0167-6393(94)00044-B

Hutcheson, G. D., & Sofroniou, N. (1999). *The multivariate social scientist*. London: Sage.

Isaacs, T., & Trofimovich, P. (2011). *International students at Canadian universities: Validating a pedagogically-oriented pronunciation scale*. Unpublished corpus of second language speech.

Issacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition, 34*, 475-505.

Jun, S.-A. (2005). Prosodic typology. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 430-458). Oxford: Oxford University Press.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System, 38*, 301-315. doi:10.1016/j.system.2010.01.005

Kang, O., Rubin, D., & Pickering, L. (2010). Supresegmental measures of accentedness and judgments of language learner proficiency in oral English. *Modern Language Journal, 94*, 554-566. doi:10.1111/j.1540-4781.2010.01091.x

King, C. (1994). *One language, two scripts: The Hindi movement in nineteenth century North India*. Delhi: Oxford University Press.

Ladefoged, P., & Johnson, K. (2014). *A Course in phonetics (7th ed.)*. Stamford, CT: Cengage.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.

Larson-Hall, J. (2009). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*, 369-377. doi:10.2307/3588485

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of language acquisition: Second language acquisition* (pp. 413-468). New York: Academic Press.

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22*, 471-497. doi:10.1017/S0272263100004022

Mackey, A. & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407-452). Oxford: Oxford University Press.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 49*, 285-310. doi:10.1111/0023-8333.49.sl.8

Munro, M. J., & Derwing, T. M. (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., & Burgess, C. (2010). Detection of nonnative speaker status from content-masked speech. *Speech Communication, 52*, 626-637. doi:10.1016/j.specom.2010.02.013

Ochs, E. (1979). Transcription as theory. In E. Orchs & B. B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 43-72). New York: Academic Press.

Rau, D. V., Chang, H.-H. A., & Tarone, E. E. (2009). Think or sink: Chinese learners' acquisition of the English voiceless interdental fricative. *Language Learning, 59*, 581-621. doi:10.1111/j.1467-9922.2009.00518.x

Saito, K., Trofimovich, P., & Isaacs, T. (forthcoming). Using listener judgements to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study.

Saito, K., Trofimovich, P., & Isaacs, T. (in press, 2015). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.

Shackle, C. (2001). Speakers of South Asian languages. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (pp. 310-324). Cambridge: Cambridge University Press.

Smith, L. E. (1992). Spread of English and issues of intelligibility. In B. B. Kachru (Ed.), *The other tongue: English across cultures* (pp. 75-90). Chicago: University of Illinois Press.

Tokumoto, M., & Shibata, M. (2011). Asian varieties of English: Attitudes towards pronunciation. *World Englishes, 30*, 392-408. doi:10.1111/j.1467-971X.2011.01710.x

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

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, *15*, 905-916.

Trubetzkoy, N. S. (1939). *Grundzüge der phonologie.* Los Angeles: University of California Press.

Wilson, L., & Wilson, M. (2001). Farsi speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (pp. 310-324). Cambridge: Cambridge University Press.

Winke, P., & Gass, S. (2013). The influence of L2 experience and accent familiarity on oral proficiency rating: A qualitative investigation. *TESOL Quarterly, 47*, 762-789. doi:10.1002/tesq.73

Winters, S., & O'Brien, M. G. (2013). Perceived accentedness and intelligibility: The relative contributions of F0 and duration. *Speech Communication, 55*, 486-507. doi:10.1016/j.specom.2012.12.006

Table 1

Means and Standard Deviations for Participant Background Characteristics

Background variable	Chinese	Hindi/Urdu	Farsi
Gender (m/f)	6/9	14/1	9/6
Age	22.5 (2.9)	23.5 (2.0)	25.2 (2.4)
Years of English study	10.3 (2.9)	14.3 (6.0)	8.5 (4.8)
Years in Canada	0.7 (.3)	0.4 (.2)	0.4 (.2)
TOEFL iBT total score	84.8 (5.9)	92.6 (4.8)	87.8 (7.1)
IELTS total score	6.3 (.5)	6.7 (.6)	6.8 (.4)
English use at home[a]	17.0 (16.9)	40.0 (26.5)	21.0 (34.1)
English use at school[a]	72.7 (21.5)	83.3 (20.6)	50.0 (30.5)

Note. [a]Self-rating on a 0-100% scale.

Table 2
Means and Standard Deviations for Accentedness and Comprehensibility Ratings (1000-Point Sliding Scale)

L1 group	Accentedness	Comprehensibility
Chinese	343 (115)	417 (115)
Hindi-Urdu	434 (170)	611 (116)
Farsi	524 (150)	638 (092)

Table 3
Summary of a Two-Factor Solution Based on a Principal Component Analysis of the 10 Rated Linguistic Variables

Factor 1 (Pronunciation)	Word stress errors (.98), Intonation (.91), Rhythm (.86), Segmental errors (.85), Speech rate (.46)
Factor 2 (Lexicogrammar)	Discourse richness (.95), Grammatical complexity (.94), Lexical richness (.93), Grammatical accuracy (.86), Lexical appropriateness (.84), Speech rate (.60)

Note. All eigenvalues > 1

Table 4
Results of Multiple Regression Analyses Using the Factors of Pronunciation and Lexicogrammar as Predictors of Accent and Comprehensibility

Predicted variable	Predictor variables	*Adjusted* R^2	R^2 *change*	$F(1, 44)$	*p*
Accent	Pronunciation	.64	.64	80.76	.0001
Comprehensibility	Lexicogrammar	.49	.49	42.59	.0001
	Pronunciation	.70	.21	52.05	.0001

Note. The variables entered into the regression equation were the two factors obtained in the PCA reported in Table 3.

Table 5
Pearson Correlations Between the Pronunciation and Lexicogrammar Factors and Accentedness and Comprehensibility by L1 Group

Factor	Accentedness			Comprehensibility		
	Chinese	Hindi/Urdu	Farsi	Chinese	Hindi/Urdu	Farsi
Pronunciation	.68*	.79*	.83*	.71*	.51	.35
Lexicogrammar	-.09	.19	-.14	.44	.74*	.47

Note. *significant correlation (a = .0025).

Table 6
Pearson Correlations Between the 10 Rated Linguistic Categories and Accentedness and Comprehensibility by L1 Group

Category	Accentedness			Comprehensibility		
	Chinese	Hindi/Urdu	Farsi	Chinese	Hindi/Urdu	Farsi
Segmentals	.91*	.92*	.93*	.71*	.57	.65
Word stress	.64	.70*	.85*	.59	.29	.31
Rhythm	.51	.52	.63	.62	.57	.24
Intonation	.66	.77*	.56	.64	.52	.13
Speech rate	.18	.41	.23	.67	.60	.33
Lexical appropriateness	.19	.31	-.06	.18	.77*	.47
Lexical richness	-.15	.27	-.18	.34	.73*	.47
Grammatical accuracy	.08	.18	.13	.34	.65	.40
Grammatical complexity	-.22	.18	-.11	.30	.71*	.37
Discourse richness	-.17	.05	-.24	.36	.72*	.36

Note. *significant correlation (a = .0025).