PERCEIVED FOREIGN ACCENT IN THREE VARIETIES OF NON-NATIVE ENGLISH*

Elizabeth A. McCullough
Ohio State University

Abstract

What aspects of the speech signal cause listeners to perceive a foreign accent? While many studies have explored this question for a single variety of non-native speech, few have simultaneously considered non-native speech from multiple native language backgrounds. In this perception study, American English-speaking listeners rated stop-vowel sequences extracted from English words produced by L1 American English, L1 Hindi, L1 Korean, and L1 Mandarin talkers on a continuous scale of degree of foreign accent. Stepwise linear regression models revealed that VOT, vowel quality, f0, and vowel duration contributed significantly to the ratings. Additionally, listeners rated productions by all varieties of non-native talkers as sounding foreign-accented to some degree, with those by L1 Hindi talkers as most foreign-accented, and those by L1 Mandarin talkers as more foreign-accented than those by L1 Korean talkers. The results suggest that several acoustic properties contribute substantially to the perception of foreign accent, at least for stop-vowel sequences, and that some varieties of non-native English sound more accented than others.

---

*I thank Mary Beckman, Cynthia Clopper, and Jeff Holliday for valuable input as this project developed, and audiences at the 161st meeting of the ASA and LabPhon13 for helpful comments on previous versions of this work.

Elizabeth A. McCullough. Perceived foreign accent in three varieties of non-native English. *OSUWPL* Volume 60, Spring 2013, 51-66.

1 Introduction

Native listeners of a language can recognize when someone is speaking that language with a foreign accent, even on the basis of very short samples of speech (Flege, 1984). However, they need not to do so accurately to motivate the study of foreign accent perception. Even when listeners misidentify a native talker as being foreign, or an L2 talker as being native, they are basing these judgments on some principled idea of what constitutes "foreign accent." Further, these judgments have repercussions, as sounding accented can have negative social consequences for the talker (Gluszek & Dovidio, 2010), and a listener viewing a talker as "other" may be biased against understanding the talker's speech (Rubin, 1992). The goals of the present investigation are to identify which acoustic properties contribute to the perception of foreign accent by native listeners of American English, and to explore such listeners' implicit views about the relative degrees of foreign accent in different varieties of non-native speech.

In many studies of perceived foreign accent, native listeners quantify the degree of foreign accent they hear in each production, and relationships between these ratings and talker-specific characteristics are examined. For instance, Oyama (1976) found a clear relationship between perceived foreign accentedness and age of arrival of Italian immigrants to the United States: the earlier an individual had immigrated, the weaker a foreign accent he was later judged to have. Flege, Munro, and MacKay (1995) conducted a similar investigation of Italian-accented English, and found that age of learning accounted for 59% of the variance in perceived foreign accent ratings, with earlier learners sounding more native. Such studies, however, do not address the question of which characteristics directly influence native listeners in their assignments of perceived foreign accent ratings. Listeners have no knowledge of an individual talker's language history, and must be attending to properties of the acoustic signal.

What acoustic properties might contribute to the perception of foreign accent? Traditional accounts of L2 acquisition (e.g., Lado, 1957) refer to cross-language phonological differences and the role of L1 "interference" in production of the L2. This suggests that it might be possible to measure and compare acoustic properties that are likely to differ between a talker's L1 and L2, and correlate the degree of perceived foreign accent with the degree of difference on these specific measures.

While this approach is not new, investigations of the signal have generally focused on single acoustic properties. For instance, VOT has been found to contribute to the perception of foreign accent for voiceless stops in L2 Spanish productions by L1 English speakers (Gonzalez-Bueno, 1997), and in L2 English productions by L1 Brazilian Portuguese speakers (Major, 1987) and L1 Japanese speakers (Riney & Takagi, 1999). Pronunciations of liquids and vowels seemed to influence the perception of foreign accent in L2 English productions by L1 Japanese speakers in Riney, Takagi, and Inutsuka (2005), although this was determined by a phonetically trained listener's auditory analysis rather than any acoustic measurements. Munro and Derwing (2001) found that speaking rate influenced the perception of foreign accent in the speech of L2 English talkers from 12 different L1s.

In addition, such investigations, with the notable exception of Munro and Derwing (2001), have nearly exclusively focused on single varieties of non-native speech. Thus, it is not clear that these studies investigated "foreign accent" generally as opposed to more specific scales such as “Brazilian Portuguese accent” or “Japanese accent.” If the task demanded only comparisons between native talkers and L2 talkers from a single L1 background, listeners might have implemented the more specific scale regardless of the instructions given.

To focus on “foreign accent” in general as opposed to some particular accent, the present study, like Munro and Derwing (2001), uses L2 English productions by talkers of multiple L1s. In addition, multiple acoustic properties are measured and evaluated in relation to listeners’ ratings of the degree of foreign accent in each of these productions. The analysis below will identify acoustic properties that seem to influence foreign accent perception, and determine whether some L1 backgrounds give rise to a stronger percept of foreign accent in L2 English than do others.

2 Methods

2.1 Acoustic stimuli

The acoustic stimuli in this experiment were extracted from recordings in the Buckeye GTA Corpus (Hardman, 2010). This corpus includes productions of 64 of the Bamford-Kowal-Bench sentences revised for American English (Bamford & Wilson, 1979) by 24 L1 American English talkers, 19 L1 Hindi talkers, 20 L1 Korean talkers, and 20 L1 Mandarin talkers. All non-native talkers were of a reasonably high English proficiency level, in that they were certified as a Graduate Teaching Associate at The Ohio State University (by a score of 230/300 on the SPEAK test or an “unconditional pass” on the university’s Mock Teaching Test) or had scored at least 26/30 on the speaking section of the TOEFL iBT. For this study, 4 female talkers from each of the 4 L1 groups were used, for a total of 16 talkers. Including 4 talkers from each L1 ensured that there was variation within each L1 group, such that any effects of L1 would not be confounded with individually idiosyncratic pronunciations. Although English is commonly spoken in India, none of the 4 L1 Hindi talkers chosen identified English as a native language.

While many previous studies have used sentences as audio stimuli in perceived foreign accent rating tasks, an exhaustive acoustic investigation of sentences would be a daunting undertaking. Units as large as sentences have a relatively large number of possible segmental and suprasegmental cues, many of which might be expected to influence ratings of perceived foreign accent. In order to limit the number of potentially relevant acoustic cues, stop-vowel sequences were chosen as the acoustic stimuli for the present study. Stop-vowel sequences are even smaller than the words rated in Gonzalez-Bueno (1997) and Major (1987), and as such have a comparatively small number of possible cues to explore, although listeners are still able to perceive foreign accent in units this size (Flege, 1984).

Sequences containing stops are particularly interesting given the language backgrounds included in the Buckeye GTA Corpus. Stops in some of these languages differ from those in American English phonologically: at each place of articulation, Korean distinguishes 3 stops (Lee, 1999) and Hindi, 4 (Ohala, 1999), while American English and Mandarin have only 2 (Ladefoged, 1999; Lee & Zee, 2003). Additionally, the stop contrasts in Hindi, Korean, and Mandarin all differ acoustically from the American English contrast. Such differences might be reflected to some degree in the English productions of these non-native talkers.

The stop-vowel productions used as stimuli were extracted from utterance-medial, word-initial contexts in the Buckeye GTA Corpus recordings. Three stop-vowel sequences were chosen for each of the 6 American English stops, for a total of 18 stop-vowel sequences. The limited number of word types in the corpus did not allow for control of the vowel, nor for the exclusion of stop-vowel sequences that were themselves lexical items (/pi/ and /tu/). All 18 stop-vowel sequences and the words they were extracted from are included in Table 1 in the Appendix.

Each of the 16 talkers produced each of the 18 stop-vowel sequences, for a total of 288 audio stimuli. The mean intensity of all audio stimuli was normalized. No fillers were used.

2.2 Participants

28 monolingual American English-speaking listeners participated in the rating study for linguistics course credit.

2.3 Task

The audio stimuli were presented through headphones at individual computer stations. Listeners were asked to rate the degree of foreign accent in each audio stimulus by sliding a bar along a continuous Visual Analog Scale (VAS). The endpoints of the VAS were labeled “no foreign accent” and “strong foreign accent,” as shown in Figure 1. Listeners were encouraged to use the scale continuously rather than categorically. VAS was chosen because listeners have fairly sensitive responses to foreign accent rating tasks, as confirmed by the continuous rating scale used by Flege et al. (1995), which might not be sufficiently captured by a Likert scale with a fixed number of intervals. In addition, a continuous rating scale allowed for the possibility of better correlation with the continuous acoustic cues discussed below.

`<-- no foreign accent                    strong foreign accent -->`

Figure 1. VAS rating line

Just before each stimulus played, the word it had been extracted from was displayed orthographically on the computer screen so that listeners had some perceptual target for the upcoming production. The audio stimuli were blocked by stop-vowel sequence; that is, all 16 productions of the same stop-vowel sequence were played consecutively. Listeners were permitted to take breaks between each block. Block order was randomized, as was stimulus order within each block. A practice block with an additional stop-vowel sequence not used as a test stimulus (/be□/ from baby) was administered before the test blocks so that listeners could experience the types of variation in the stimuli and practice using the rating scale.

2.4 Acoustic properties

Four acoustic properties were considered in the analysis: VOT, vowel quality, vowel duration, and f0 (fundamental frequency). VOT and vowel quality (midpoint F1 and F2) were chosen as fairly standard measures that have been previously shown to be related to perceived foreign accent ratings (Gonzalez-Bueno, 1997; Major, 1987; Riney & Takagi, 1999; Riney et al., 2005). Vowel duration was included in response to remarks from listeners that some stimuli were too short to rate properly. Similarly, f0 was included in response to remarks from listeners that differences in pitch made the task difficult.

Speaking rate, although found to play a role in foreign accent perception by Munro and Derwing (2001), was not investigated here. Because a temporal measure of each of the two segments in each stop-vowel stimulus was already included, speaking rate would have been largely redundant.

2.5 Analysis

For the production measures, VOT and vowel intervals were marked by the author in Praat (Boersma & Weenink, 2012), and their durations were extracted automatically. F1 and F2 values at the vowel midpoint and mean f0 values over the vowel were extracted automatically using a Praat script that displayed a spectrogram with overlaid formant and pitch tracks, such that the author could check token-by-token for tracking errors. In the rare case of such an error, the author re-measured the token in question with modified settings.

For the perception measure, each listener rating was recorded by the experiment presentation software as a numerical value representing the position on the rating line. In total, 8064 ratings were assigned (16 talkers x 18 stop-vowel sequences x 28 raters).

3 Predictions

While variation in both acoustic measurements and perceived foreign accent ratings can of course be expected in productions by different talkers from the same L1 background and even in different productions by the same talker, this study focuses mainly on variation across L1 backgrounds. Actual language background is expected to align to some degree with perceived foreign accent in a talker's speech, in that productions from native talkers are likely to be judged as exhibiting little to no foreign accent, and productions by non-native talkers are likely to be judged as exhibiting some degree of foreign accent. In this section, the effect of L1 background on each acoustic property is examined for the stimuli used in the present study, and subsequent predictions are made. If there are differences between native and non-native productions for a particular acoustic property, then listeners might use that property to cue foreign accent. Additionally, if a particular variety of non-native speech differs from native speech on multiple acoustic properties, then it might be judged as strongly foreign-accented. As explained further in Section 4.1, stimuli with voiced versus voiceless stop targets will be analyzed separately to facilitate comparison with previous studies; thus, these acoustic measurements are also presented separately for the two groups of stimuli.

3.1 Differences by L1 background

Figure 2 shows VOT values for the productions by talkers of each variety of English. Separate one-way ANOVAs with talkers as subjects and L1 as a between-subjects factor revealed significant effects of L1 on VOT for voiced ($F(3,12) = 6.50$; $p < 0.01$) and voiceless ($F(3,12) = 12.80$; $p < 0.001$) stop targets. Separate Tukey post-hoc tests for voiced and voiceless stop targets showed that these effects were driven by the differences between L1 Hindi talkers' values and all others ($p < 0.05$); no other pairwise comparisons were significant. Indeed, the most striking details of Figure 2 are the VOT ranges for L1 Hindi talkers' productions. Unlike the voiced stop productions by talkers from other L1 backgrounds, which exhibited short lag voicing, a substantial portion (56%) of the L1 Hindi talkers' voiced stop productions were actually prevoiced. Additionally, L1 Hindi talkers' voiceless stop productions had much shorter VOT values than those for talkers of all other L1s.

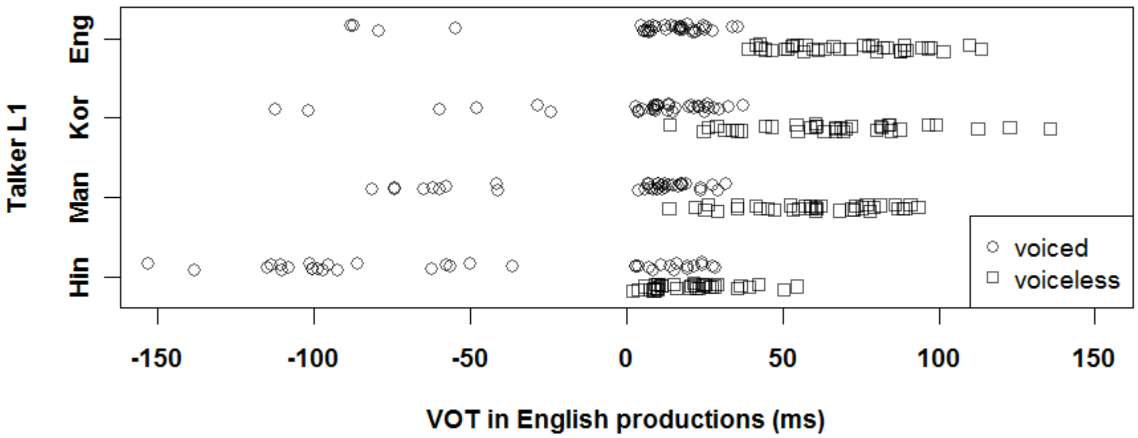

Figure 2. VOT values by target voicing and talker L1

The use of multiple varieties of non-native English in this study presents a problem for linear regression using raw values: productions by native talkers can be expected to have values near one extreme of the range for perceived foreign accent ratings ("no foreign accent"), but the same assumption cannot necessarily be made about the distribution of their acoustic measurements. That is, it would be possible for productions by talkers from one non-native background to have values lower than those observed in productions by native talkers, while productions by talkers from a different non-native background might have values higher than those observed in productions by native talkers. This problem is avoided here by using values that represent the difference between a production and the comparable native productions on some acoustic property. Specifically, all acoustic cues considered below are parameterized as the absolute difference from the American English mean, calculated separately for each of the 18 stop-vowel sequences. For instance, rather than examining the raw VOT value for some L1 Korean talker's production of /pi/ from the word people, the value considered is the absolute difference between the VOT value of this production of /pi/ and the mean VOT of the 4 L1 American English talkers' productions of /pi/. Difference values for productions by L1 American English talkers are also calculated, and should generally be small. Figure 3 shows VOT differences for the productions by each talker. Talkers within each L1 group are presented in a consistent order across all plots.

Separate one-way ANOVAs with talkers as subjects and L1 as a between-subjects factor revealed significant effects of L1 on VOT difference for voiced ($F(3,12) = 7.45$; $p < 0.01$) and voiceless ($F(3,12) = 38.38$; $p < 0.001$) stop targets. Again, this result was due to the values for productions by L1 Hindi talkers being different from those for productions by talkers from all other L1 backgrounds, as shown by separate Tukey post-hoc tests for voiced and voiceless stop targets ($p < 0.05$). In the case of VOT, the effect of L1 background is the same regardless of whether raw values or values that denote the difference from native production means are used.

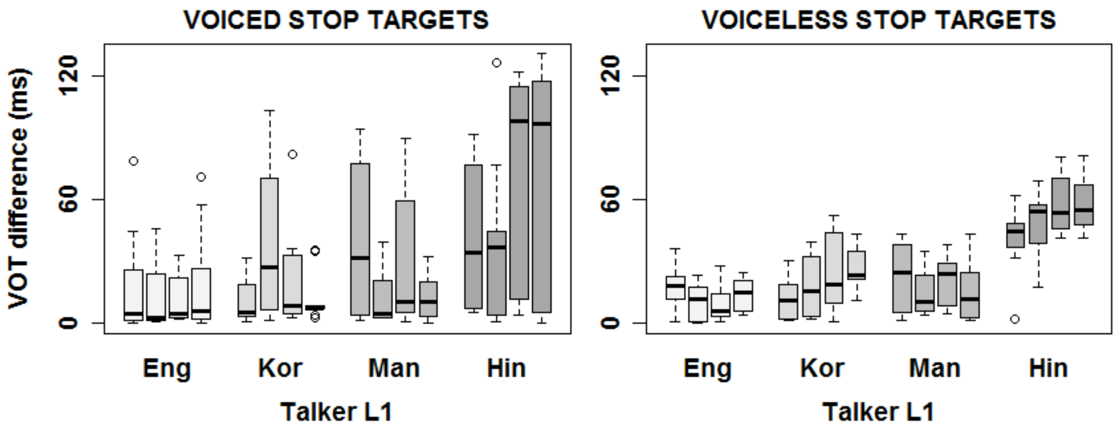

Figure 3. VOT differences by target voicing, talker L1, and talker identity

Vowel quality is defined in terms of two dimensions, F1 and F2. To reduce this to a single dimension, vowel quality difference is quantified as the Euclidean distance in F1/F2 space from the position of a single production to the mean position for the productions of that stop-vowel sequence by the L1 American English talkers. Figure 4 shows vowel quality differences for the productions by each talker. A one-way ANOVA with talkers as subjects and L1 as a between-subjects factor showed a significant effect of L1 on vowel quality difference for stimuli with voiceless stop targets ($F(3,12) = 8.04$; $p < 0.01$); the effect was not significant for stimuli with voiced stop targets. A Tukey post-hoc test for stimuli with voiceless stop targets revealed that vowel quality difference in productions by L1 Korean and L1 Mandarin talkers was different from that in productions by L1 American English talkers ($p < 0.05$); difference values were greater for productions by the non-native talkers. The comparison between difference values for L1 Hindi talkers' productions and L1 American English talkers' productions, in the same direction, approached significance ($p = 0.05$).

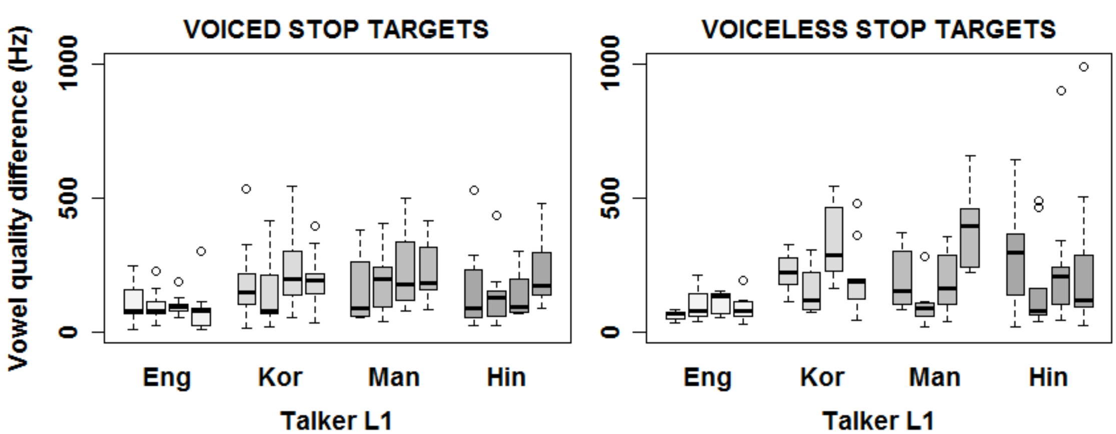

Figure 4. Vowel quality differences by target voicing, talker L1, and talker identity

Figure 5 shows vowel duration differences for the productions by each talker. A one-way ANOVA with talkers as subjects and L1 as a between-subjects factor revealed a significant effect of L1 on vowel duration difference for stimuli with voiced stop targets ($F(3,12) = 10.12$; $p < 0.01$); the effect was not significant for stimuli with voiceless stop targets. A Tukey post-hoc test for stimuli with voiced stop targets showed that vowel duration difference in productions by L1 Korean and L1 Mandarin talkers was different from that in productions by L1 American English talkers ($p < 0.05$); again, difference values were greater for productions by the non-native talkers.

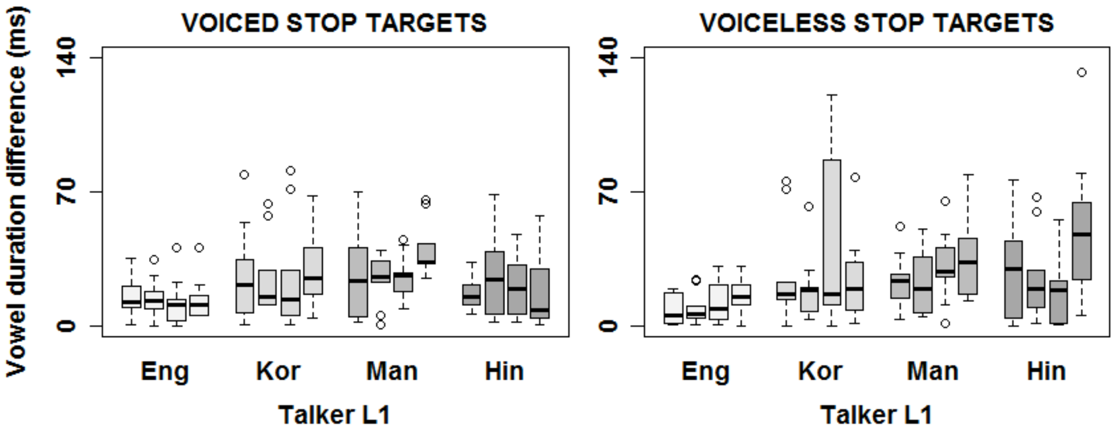

Figure 5. Vowel duration differences by target voicing, talker L1, and talker identity

Figure 6 shows f0 differences for the productions by each talker. Separate one-way ANOVAs with talkers as subjects and L1 as a between-subjects factor did not reveal a significant effect of L1 on f0 difference for either set of stimuli.

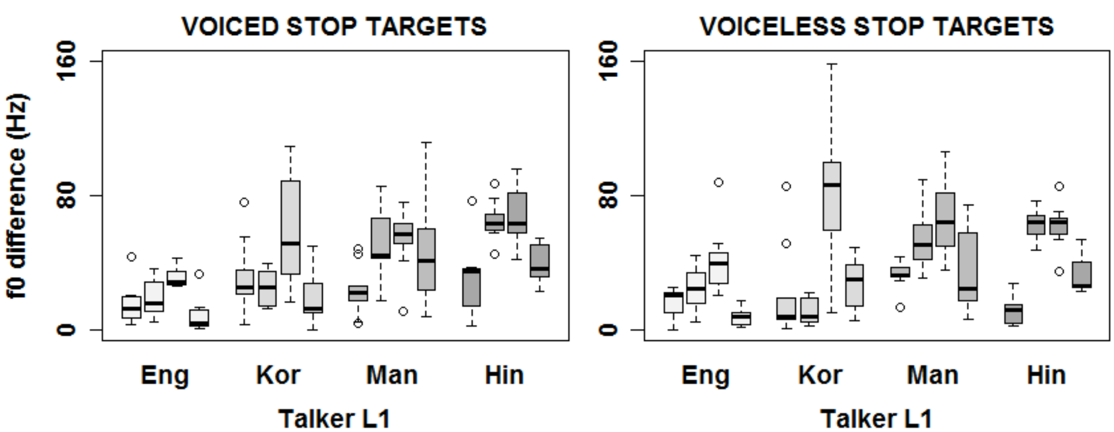

Figure 6. f0 differences by target voicing, talker L1, and talker identity

3.2 Acoustic cues

As differences were observed between native and non-native talkers in VOT, vowel quality, and vowel duration, these acoustic properties might play a role in foreign accent perception. In support of this prediction, the relationship of VOT to foreign accent perception has been demonstrated previously by acoustic measurements (Gonzalez-Bueno, 1997; Major, 1987; Riney & Takagi, 1999), and vowel quality's role has been suggested previously by auditory analysis (Riney et al., 2005). In contrast, f0 did not reliably differ by L1 background and may be unlikely to influence the perception of foreign accent.

3.3 Overall rating patterns

Productions by L1 Hindi talkers differed from native talkers' productions in VOT and marginally in vowel quality, while productions by L1 Korean and L1 Mandarin talkers differed from native talkers' productions in vowel quality and vowel duration. Due to these acoustic differences, it is expected that productions from all non-native backgrounds should sound accented to some degree. As VOT has repeatedly been shown to correlate with perceived foreign accent ratings, and as productions from L1 Hindi talkers are the only ones that came close to showing effects for properties of both consonants and vowels, it is possible that productions from L1 Hindi talkers might sound more foreign-accented than productions by other non-native talkers. This would align well with feedback from listeners, who often commented that the "Indian accent" was the most obvious.

4 Results

In this section, the possible acoustic cues to foreign accent perception are evaluated, and then the relative strength of foreign accent in different varieties of non-native English is explored.

4.1 Acoustic cues

The regression models detailed in this section test the relationship between perceived foreign accent ratings and acoustic cues directly, and the stepwise models evaluate the contributions of all measured acoustic cues simultaneously. A significant positive relationship between any acoustic cue and perceived foreign accent ratings means that as values of that acoustic cue deviate more from the L1 American English talkers' mean, listeners perceive a higher degree of foreign accent.

The dependent variable is the mean perceived foreign accent rating for each of the 288 stimuli. Thus, while the predictions above were based on differences in acoustic properties across L1 groups, this portion of the analysis should also account for differences in ratings among talkers in the same L1 group, as well as differences in ratings among the 18 productions by a single talker. To the extent that productions varied in some significant acoustic cue, in terms of deviation from the native mean, the ratings for those productions should be different, regardless of the identity of the talker.

Previous studies have found clear relationships between VOT and ratings of perceived foreign accent (Gonzalez-Bueno, 1997; Major, 1987; Riney & Takagi, 1999). Notably, however, these studies only explored this relationship with voiceless stop targets. Thus, the results below are presented separately for voiced and voiceless stop targets, to make comparison with previous findings possible.

A stepwise linear regression model of the data for stimuli with voiced stop targets only identified vowel quality, VOT, f0, and vowel duration, in that order, as significant cues. Four linear regression models were built, such that the first included only the most significant cue, the second included the two most significant cues, etc. Full statistical details are given in Table 2 in the Appendix. For the present discussion, it is important to highlight that the model with vowel quality accounted for 18% of the variance in the perceived foreign accent ratings. With the addition of VOT, this rose to 36% (18% improvement); with the addition of f0, this rose to 42% (6% improvement); and with the addition of vowel duration, this rose to 44% (2% improvement). Thus, the best model accounted for less than half of the variance in ratings, and included an acoustic property, f0, that was not predicted to play a role.

Similar models were created for stimuli with voiceless stop targets only; full statistical details are given in Table 3 in the Appendix. In the first model, VOT alone explained 40% of the variance in the ratings of perceived foreign accent—nearly as much as all four acoustic cues combined in the model for voiced targets. The addition of vowel quality brought this value to 45% (5% improvement), and the model that also included f0 accounted for 48% of the variance (3% improvement). Vowel duration, although significant in the stepwise model, did not contribute significantly to the explicitly specified regression model.

4.2 Overall rating patterns

Perceived foreign accent ratings by talker are shown for stimuli with voiced and voiceless stop targets separately in Figure 7. Separate one-way repeated measures ANOVAs with listeners as subjects and L1 as a within-subjects factor revealed significant effects of L1 on perceived foreign accent ratings for stimuli with voiced ($F(3,104) = 8.93$; $p < 0.001$) and voiceless ($F(3,104) = 10.29$; $p < 0.001$) stop targets. Post-hoc Bonferroni-corrected paired t-tests showed differences between all pairwise comparisons for voiced targets ($p < 0.00017$), as well as for voiceless targets ($p < 0.00017$). While there is quite a range of ratings within each L1 group, general patterns do emerge. As predicted, productions by L1 American English talkers were rated as having very little foreign accent, and productions by non-native talkers were rated as having greater degrees of foreign accent. In addition, productions by L1 Hindi talkers were rated as having the highest degree of foreign accent. While there was no basis for predicting a difference between perceived foreign accent ratings for productions by L1 Korean talkers as compared to productions by L1 Mandarin talkers, the latter received higher ratings, as shown in Figure 7.

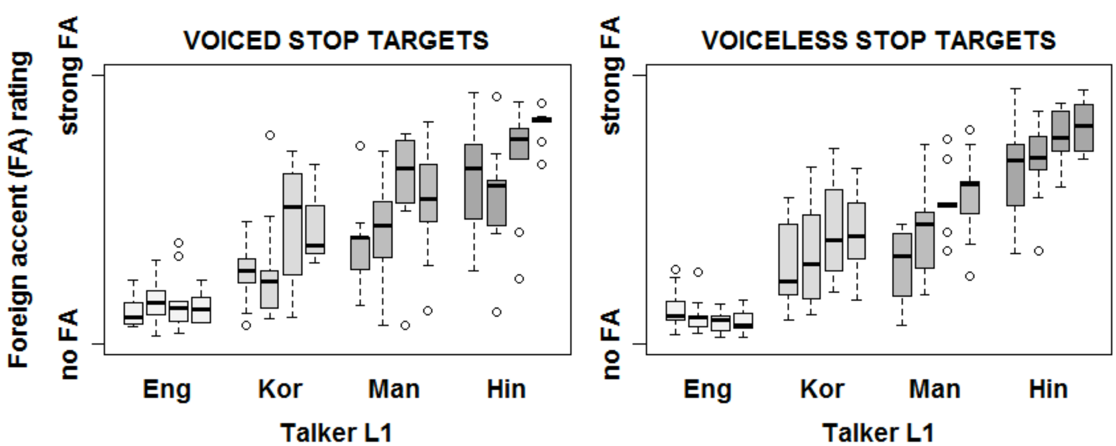

Figure 7. Ratings of perceived foreign accent by target voicing, talker L1, and talker identity

5 Discussion

The first goal of this research was to identify potential acoustic cues to perceived foreign accent. Acoustic properties were measured in very short stop-vowel sequences extracted from sentences in English. The same sequences were used as stimuli in a rating task, and the perceived foreign accent ratings were shown to be correlated with VOT, vowel quality, f0, and vowel duration for at least some of the stimuli. For stimuli with voiced stop targets, the best model included all four acoustic properties and accounted for 44% of the variance in perceived foreign accent ratings. For stimuli with voiceless stop targets, VOT, vowel quality, and f0 accounted for 48% of the variance in perceived foreign accent ratings, with VOT playing a substantial role.

The present study confirmed previous findings that VOT and vowel quality seem to be involved in the perception of foreign accent, suggested that f0 and vowel duration may also play minor roles, and showed that the relative importance of these cues differs somewhat for stimuli containing voiced stops targets as opposed to voiceless stop targets. This last fact is particularly notable given the frequent use of sentence-length stimuli, which contain many segments that may or may not be controlled for voicing, in studies of foreign accent. It remains to be seen whether such differences in relative importance persist in the perception of longer stimuli, or whether segmental cues are washed out by more global prosodic properties.

Although VOT was a significant predictor regardless of the subset of data analyzed, it clearly played a larger role in ratings of stimuli with voiceless stop targets than in ratings of stimuli with voiced stop targets. Figure 2 showed that a number of the voiced stop productions, especially by L1 Hindi talkers, were prevoiced, while voiceless stop productions were characterized by lag voicing. The difference in the contribution of VOT might have resulted from the listener population in this study; American English-speaking listeners are likely to be quite sensitive to lag voicing differences, as most word-initial productions by native speakers of American English are in this range. However, this effect might equally have been a general consequence of perception, as even listeners of languages with prevoiced stops seem to be more sensitive to the presence or absence of prevoicing than to the amount of prevoicing (van Alphen & McQueen, 2006). Thus, a convincing linear relationship between VOT and perceived foreign accent ratings might be unlikely for any stimuli with voiced stop targets. If so, a linear regression analysis would not be expected to perform well on this portion of the data. As previous investigations have generally focused only on voiceless stop targets, this complication has not been widely recognized.

While all four acoustic cues measured were found to contribute significantly to the model for at least one of the two target voicing categories, collectively they accounted for around half of the variance in perceived foreign accent ratings, leaving half of the variance unexplained. Stop-vowel stimuli were chosen so as to limit the number of acoustic cues listeners were able to attend to in performing the task, but it seems that listeners used more than the four acoustic cues measured here. As an individual's L1 is expected to influence L2 production, additional acoustic properties to consider in future analyses might be those that distinguish stops in the non-native talkers' L1s. For instance, voice quality and noise offset time are often invoked in phonetic descriptions of Hindi stops (Davis, 1994), and in Korean, voice quality and f0 at the onset of the following vowel are relevant to the 3-way stop distinction (Kong, Beckman, & Edwards, 2011). Although such cues are not needed for phonemic distinctions in American English, this does not mean that they were not perceived by American English-speaking listeners. If detectable in the stimuli, such cues might clearly identify certain productions as foreign and augment their ratings of perceived foreign accent.

The second goal of this research was to investigate whether some varieties of non-native English sound generally more accented than others. The large ranges of ratings for productions by talkers of each L1 indicate that listeners were not simply categorizing the productions into four discrete piles on the scale. Acoustic measurements differ among productions by talkers of different L1s, but also among productions by talkers from the same L1 and even among productions by the same talker. These multiple levels of variation might explain why f0 turned out to contribute to the regression models for both voiced and voiceless stop targets despite there being no effect of talker L1 on its difference values. Nonetheless, it seems clear that not all L1 backgrounds are equal when it comes to the perception of foreign accent. While productions by all non-native talkers were judged as sounding foreign-accented to some degree, those by L1 Hindi talkers were most accented, while those by L1 Mandarin talkers were more accented than those by L1 Korean talkers. These results suggest that listeners have some quantifiable concept of "foreign accent" more abstract than the "Brazilian Portuguese accent" or "Japanese accent" concepts that might have been explored in earlier studies, though the relationship between the abstract "foreign accent" and the various more specific incarnations is yet unclear.

These two sets of results also can be related to each other. For instance, the acoustic analysis showed that productions by L1 Hindi talkers differed significantly from productions by native talkers in their VOT values. Thus, there is a clear relationship between the importance of VOT as an acoustic cue and the high perceived foreign accent ratings for productions by L1 Hindi talkers. The directionality, however, is uncertain. Are productions by L1 Hindi talkers rated as most foreign-accented because they differ in VOT, or does VOT matter because English produced by L1 Hindi talkers is thought to sound highly foreign-accented? When asked if they could identify the accents present in the stimuli, 19 of 28 listeners identified "Indian" (15), "Hindi" (3), or "South Asian" (1), indicating that this variety of speech may have been easily identifiable. However, 18 of 28 listeners identified "Chinese" (9), "Mandarin" (1), "Asian" (7), or "East Asian" (1), while this variety of speech was rated as sounding significantly less foreign-accented. Only 2 of 28 listeners specifically identified "Korean," although some of the "Asian" and "East Asian" responses may have been at least partly in response to the productions by L1 Korean talkers. Further research is needed to investigate the role that stereotypes and familiarity with particular non-native varieties might play in the perception of foreign accent.

The results of this study are particularly interesting in light of the sociolinguistic situation of English in India, where "it is the 'associate official' language of the country and it also serves as a link language between the educated" (Gargesh, 2004, p. 992). Although they did not claim English as a native language, the L1 Hindi talkers in this study reported that they first began studying English between ages 3 and 4. In contrast, the L1 Korean and L1 Mandarin talkers in this study reported that they did not begin to learn English until at least age 11. Compared to the other non-native groups, the L1 Hindi talkers' experience with English started earlier and lasted longer. Their target variety of English was Indian rather than American, however, and the ratings from this experiment suggest that this variety can sound quite foreign-accented to American English-speaking listeners. Thus, judgments about foreign accent are critically shaped by the listener's ideas about the world.

A listener's impressions regarding foreign accent develop over the course of day-to-day language use. However, most communication requires units much more complex than the stop-vowel sequences used as stimuli in this study, and the cues discovered here would not necessarily play the same roles in ratings of larger units like words or sentences. As longer stimuli would exhibit a much wider variety of potential acoustic cues, a considerably more complicated relationship between perceived foreign accent ratings and acoustic cues would seem inevitable. Nonetheless, the present work has taken the initial steps in addressing the question of what aspect of the speech signal causes listeners to perceive a foreign accent.

References

Bamford, John, and Ian Wilson. 1979. Methodological considerations and practical aspects of the BKB sentence lists. In *Speech-hearing tests and the spoken language of hearing-impaired children*, ed. by John Bench and John Bamford, 148-187. London: Academic Press.

Boersma, Paul, and David Weenink. 2012. *Praat: Doing phonetics by computer* (Version 5.3) [Computer program]. <http://www.praat.org/>.

Davis, Katharine. 1994. Stop voicing in Hindi. *Journal of Phonetics* 22.177-193.

Flege, James Emil. 1984. The detection of French accent by American listeners. *Journal of the Acoustical Society of America* 76.692-707.

Flege, James Emil, Murray Munro, and Ian MacKay. 1995. Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America* 97.3125-3134.

Gargesh, Ravinder. 2004. Indian English: Phonology. In *A Handbook of varieties of English*, ed. by Edgar W. Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie, and Clive Upton, 992-1002. Berlin: Mouton de Gruyter.

Gluszek, Agata, and John F. Dovidio. 2010. The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review* 14(2).214-237.

Gonzalez-Bueno, Manuela. 1997. Voice onset time in the perception of foreign accent by native listeners of Spanish. *International Review of Applied Linguistics in Language Teaching* 35(4).251-262.

Hardman, Jocelyn B. 2010. The intelligibility of Chinese-accented English to international and American students at a US university. Columbus, OH: OSU dissertation.

Kong, Eun Jong, Mary E. Beckman, and Jan Edwards. 2011. Why are Korean tense stops acquired so early: The role of acoustic properties. *Journal of Phonetics* 39(2).196-211.

Ladefoged, Peter. 1999. American English. In *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*, 41-44. Cambridge: Cambridge University Press.

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

Lee, Hyun Bok. 1999. Korean. In *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*, 120-123. Cambridge: Cambridge University Press.

Lee, Wai-Sum, and Zee, Eric. 2003. Illustrations of the IPA: Standard Chinese (Beijing). *Journal of the International Phonetic Association* 33(1).109-112.

Major, Roy C. 1987. English voiceless stop production by speakers of Brazilian Portuguese. *Journal of Phonetics* 15.197-202.

Munro, Murray J., & Tracey M. Derwing. 2001. Modeling perceptions of the accentedness and comprehensibility of L2 speech. *Studies in Second Language Acquisition* 23.451-468.

Ohala, Manjari. 1999. Hindi. In *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*, 100-103. Cambridge: Cambridge University Press.

Oyama, Susan. 1976. A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research* 5.261-285.

Riney, Timothy J., and Naoyuki Takagi. 1999. Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning* 49(2).275-302.

Riney, Timothy J., Naoyuki Takagi, and Kumiko Inutsuka. 2005. Phonetic parameters and perceptual judgments of accent in English by American and Japanese listeners. *TESOL Quarterly* 39(3).441-466.

Rubin, Donald L. 1992. Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education* 33(4).511-531.

van Alphen, Petra M., and James M. McQueen. 2006. The effect of Voice Onset Time differences on lexical access in Dutch. *Journal of Experimental Psychology: Human Perception and Performance* 32(1).178-196.

Appendix

Table 1. Stop-vowel sequences used as audio stimuli

Sequence	*Word*
/bæ/	back
/bʌ/	buckets
/bɑ/	boxes
/dɪ/	dish
/dɪ/	dinner
/dɔ/	dog
/gɚ/	girl
/gʊ/	good
/gɑ/	got
/pi/	people
/pɪ/	picture
/pæ/	packed
/teɪ/	table
/tu/	two
/tɑʊ/	towel
/kɪ/	kitchen
/keɪ/	came
/koʊ/	coat

Table 2. Regression models (stimuli with voiced stop targets)

Variable	*Value*	*Statistical test*
Model 1	$R^2 = 0.18$	$F(1,142) = 32.67$, $p < 0.001$
Vowel quality	$m_1 = 0.09$	$t(142) = 5.72$, $p < 0.001$
Model 2	$R^2 = 0.36$	$F(2,141) = 41.06$, $p < 0.001$
Vowel quality	$m_1 = 0.10$	$t(141) = 6.72$, $p < 0.001$
VOT	$m_2 = 0.32$	$t(141) = 6.36$, $p < 0.001$
Model 3	$R^2 = 0.42$	$F(3,140) = 35.93$, $p < 0.001$
Vowel quality	$m_1 = 0.09$	$t(140) = 6.04$, $p < 0.001$
VOT	$m_2 = 0.28$	$t(140) = 5.78$, $p < 0.001$
f0	$m_3 = 0.28$	$t(140) = 4.07$, $p < 0.001$
Model 4	$R^2 = 0.44$	$F(4,139) = 28.91$, $p < 0.001$
Vowel quality	$m_1 = 0.08$	$t(139) = 5.64$, $p < 0.001$
VOT	$m_2 = 0.28$	$t(139) = 5.90$, $p < 0.001$
f0	$m_3 = 0.27$	$t(139) = 3.99$, $p < 0.001$
Vowel duration	$m_4 = 0.20$	$t(139) = 2.21$, $p < 0.05$

Table 3. Regression models (stimuli with voiceless stop targets)

Variable	*Value*	*Statistical test*
Model 1	$R^2 = 0.40$	$F(1,142) = 96.21$, $p < 0.001$
VOT	$m_1 = 0.85$	$t(142) = 9.81$, $p < 0.001$
Model 2	$R^2 = 0.45$	$F(2,141) = 60.02$, $p < 0.001$
VOT	$m_1 = 0.76$	$t(141) = 8.82$, $p < 0.001$
Vowel quality	$m_2 = 0.04$	$t(141) = 3.82$, $p < 0.001$
Model 3	$R^2 = 0.48$	$F(3,140) = 44.91$, $p < 0.001$
VOT	$m_1 = 0.73$	$t(140) = 8.69$, $p < 0.001$
Vowel quality	$m_2 = 0.03$	$t(140) = 3.35$, $p < 0.01$
f0	$m_3 = 0.17$	$t(140) = 2.90$, $p < 0.01$
Model 4	$R^2 = 0.48$	$F(4,139) = 34.57$, $p < 0.001$
VOT	$m_1 = 0.71$	$t(139) = 8.25$, $p < 0.001$
Vowel quality	$m_2 = 0.03$	$t(139) = 2.66$, $p < 0.01$
f0	$m_3 = 0.17$	$t(139) = 2.98$, $p < 0.01$
Vowel duration	$m_4 = 0.11$	$t(139) = 1.51$, n.s.