Production of coda voicing contrast of L2 English by native Mandarin Chinese speakers in comparison with native Korean speakers

Cite as: JASA Express Lett. **2**, 085202 (2022); https://doi.org/10.1121/10.0013006
Submitted: 15 March 2022 • Accepted: 23 June 2022 • Published Online: 01 August 2022

Ziqian Du, Sahyang Kim and Taehong Cho

JASA EXPRESS LETTERS | ARTICLE | asa.scitation.org/journal/jel

Production of coda voicing contrast of L2 English by native Mandarin Chinese speakers in comparison with native Korean speakers

Ziqian Du,[1] Sahyang Kim,[2] and Taehong Cho[1,a)]
[1]*Department of English Language and Literature, Hanyang Institute for Phonetics and Cognitive Sciences of Language, Hanyang University, Seoul, Korea*
[2]*Department of English Education, Hongik University, Seoul, Korea*
ziqiandu@163.com, sahyang@hongik.ac.kr, tcho@hanyang.ac.kr

Abstract: This study investigates native language effects on phonetic encoding of coda voicing contrast in L2 English by Chinese versus Korean speakers. Results show much smaller phonetic differences in both vowel duration and F0 in marking coda voicing contrast for Chinese speakers than Korean speakers, despite native Chinese speakers' experience with lexical tones. They suggest that producing an F0-related cue in L2 is conditioned by position-specific phonological richness and use of F0 in the speaker's L1. The results are discussed in terms of contrast maximization and effort minimization with reference to the information structure occurring in both L1 and L2.

[Editor: Donald Derrick] https://doi.org/10.1121/10.0013006

Received: 15 March 2022 **Accepted:** 23 June 2022 **Published Online:** 1 August 2022

1. Introduction

This study examines how phonetic implementation of phonological coda voicing contrast in English, as reflected in both the temporal (vowel duration) and spectral (F0) dimensions, can vary as a function of L2 speakers' native language experience (Mandarin Chinese versus Korean). In addition, the study discusses the specific cross-linguistic differences and linguistic principles that might underlie the observed native language effects on production of coda voicing contrast in L2 English.

The phonological contrast of coda voicing in English (e.g., /t/-/d/ as in *bet* versus *bed*) is manifested by phonetic duration, accompanied by secondary phonetic correlates along the spectral dimension, such as vowel F0 and formants. For example, the duration of the preceding vowel is longer before a voiced coda consonant than before a voiceless one [e.g., Chen (1970), Crowther and Mann (1992), and de Jong (2004)], all else being equal [e.g., Hawkins and Nguyen (2004)]. The phonetic implementation of coda voicing contrast observed in L1 English has been the locus of L2 phonetic research to understand how the sound structure of L2 speakers' native language can condition the fine phonetic detail of coda voicing contrast in L2 English.

Crowther and Mann (1992), for example, showed that, although both Chinese and Japanese speakers of L2 English use vowel duration consistently to mark the coda voicing contrast, Japanese speakers produce a durational difference with a larger magnitude (which is closer to that of native speakers) than do Chinese speakers. As the Japanese language has a moraic structure that uses the temporal dimension to mark phonological length contrast, the authors suggested that Japanese learners would be more sensitive to temporal difference of the coda voicing contrast than would Chinese learners, which could account for the observed native language effect on L2 English. Flege *et al.* (1992) showed native language effects on phonetic implementation of coda voicing contrast by Spanish versus Chinese speakers of L2 English, with Spanish speakers using vowel duration in a larger magnitude than Chinese speakers. Although the authors did not elaborate on the cause of the difference, it could be attributed to phonological differences of the L2 speakers' native languages. That is, Mandarin Chinese (henceforth Chinese) does not allow phonological contrasts in the coda position (in which only nasals can occur), but Spanish permits quite a few consonants (such as /s/, /d/, /n/, /r/, and /l/) in a word-final coda position and more consonants (with both voiced and voiceless stops in the consonant clusters) in a word-medial coda position. It is therefore reasonable to assume that the phonological richness of the coda position in Spanish could lead to increased awareness of the phonological contrast in that position in L2 English, which could be positively reflected in Spanish speakers' voicing contrast in L2 English compared with Chinese speakers of L2 English.

[a)]Author to whom correspondence should be addressed.

This study explores native language effects on production of coda voicing contrast in L2 English by newly examining L2 production by native speakers of Chinese compared to native speakers of Korean. The phonology of the two languages differs in various aspects, but one of the cross-linguistic differences relevant here is that Korean phonotactics permits numerous consonants in the coda position, including nasals, liquid /l/, and obstruents, whereas Chinese phonotactics permits only the nasals /n/ and /ŋ/. Although the underlying phonological contrast of Korean stops relevant for voicing is neutralized in word-final coda position (Kim and Jongman, 1994), stops contrast in place of articulation in coda position. Given the relative richness of phonological contrast in the coda in Korean, Korean speakers of L2 English, presumably similar to Spanish learners of L2 English, might have heightened phonological awareness of the coda voicing contrast and show more robust use of vowel duration compared to Chinese speakers of L2 English.

Another cross-linguistic difference relevant here is that Chinese is a tone language, while Korean is not. This leads to a question: whether Chinese speakers, who have native language experience using F0 systematically in marking lexical tones, will have an advantage over Korean speakers in learning the F0 difference that correlates with coda voicing contrast in English. Some previous studies suggest that native speakers of tonal languages might better perceive some F0 differences than will native speakers of non-tonal languages [e.g., So and Best (2010), Wayland and Guion (2004); see Holliday (2015) for related discussion]. It is therefore conceivable that Chinese speakers are more attuned to the F0 difference used for the coda voicing contrast in English than are speakers of non-tonal languages such as Korean. On the other hand, even though Korean is not a tone language, younger generations use F0 as the primary phonetic feature of phonological contrast among stops (categorically lower F0 for lenis stops and higher F0 for aspirated stops) (Choi *et al.*, 2020). This opens the possibility that Korean learners of English might also use F0 efficiently in marking the coda voicing contrast in L2 English, though it is unknown whether their performance will match that of Chinese learners of L2 English. As discussed in Holliday (2015), native language experience in a tone language might not translate directly into better non-native tone perception (Francis *et al.*, 2008). In fact, in an L2 phonetic study, Holliday (2015) showed that Chinese speakers are not necessarily better than English speakers in using F0 to mark the lenis-aspirated stop contrast in L2 Korean. Thus, whether experience with a native tone language will confer benefits to Chinese speakers of L2 English remains an open question to be explored in this study.

In testing those possible native language effects, we also consider focus-induced prominence in phonetic realization of phonological contrast with reference to information structure (with the focused condition being the locus of information) that can be applicable to both L1 and L2. Focus-induced prominence in L1 speech is generally considered to engender *localized hyperarticulation* of the segments under focus to result in an enhancement of phonological contrast and lexical contrast (de Jong, 2004; Cho, 2016). Contrastive focus in Chinese induces an increase in syllable duration as well as an F0 range expansion in a direction to magnify F0 contours [e.g., Chen and Gussenhoven (2008)]. Contrastive focus in Korean also induces an increase in syllable duration, but it does not show an F0 range expansion similar to Chinese (Choi *et al.*, 2016). Choi *et al.* (2016) adopted the notion of localized hyperarticulation in L2 phonetic research by examining focus-conditioned phonetic realization of coda voicing contrast produced by Korean speakers of L2 English in comparison with native English speakers. Their results showed that the non-native phonological contrast of coda voicing in L2 English was subject to localized hyperarticulation, phonetically expressed in a much clearer fashion—for example, with a larger difference in vowel duration in the focused condition than in the unfocused one. The authors suggested that segmental and suprasegmental features associated with a phonological contrast in L2 speech are realized in a communicatively optimized way, by making reference to higher-order information structure (p. 14). This study extends Choi *et al.* (2016) to L2 speech by Chinese speakers and to the F0 dimension and compares the results with those obtained from the Korean speakers in Choi *et al.* (2016). This will allow us to understand how L2 production by Chinese versus Korean speakers interacts with focus-related prominence driven by information structure, and it will also provide a linguistic context in which any phonetic difference caused by the L2 speakers' native language experience can be observed in terms of vowel duration and F0.

2. Method

2.1 Participants, speech materials, and recording procedure

Twenty native speakers of Chinese who were born and raised in mainland China participated in this study. They were either college students or recent graduates (less than one year after graduation) at the time of data collection. Participants were divided into advanced and intermediate learner groups based on English proficiency, estimated by TOEFL (Test of English as a Foreign Language, a standardized test by ETS) score. The average TOEFL score of the 10 advanced learners of English (6 females, 4 males, mean age 23.7 years, range 21–26) was 106.6; that of the 10 intermediate learners of English (8 females, 2 males, mean age 21.9 years, range 19–24) was 79.7. The TOEFL scores are based on reading, listening, speaking and writing tests combined. The advanced learners with their TOEFL Score above 100 (mean 106.6) are ranked around 85 percentile among those who have taken the test, and the intermediate learners with the TOEFL score below 90 (mean 79.7) were around 25 percentile. For comparison between Chinese and Korean learners' English, Korean data originally collected for Choi *et al.* (2016) were adopted. (The Korean learners all had lived in Korea at the time of recording.) Although that study did not report F0 data, both vowel duration and F0 data are available in an OSF repository (OSF, 2022). Our Chinese participants and the Korean participants from Choi *et al.* (2016) were comparable in terms of age (mean age 23 and 24 years, respectively, for the advanced and intermediate Korean learners of English) and average TOEFL score (110 for the advanced and 75 for the intermediate Korean learners of English). It is not known whether the participants spoke other languages besides their mother tongue and L2 English, but our general understanding is that it is not common for Chinese or Korean learners of English, living in China or Korea, to speak a third language.

To directly compare speech production between the Chinese and Korean learners of English, the speech materials in Choi *et al.* (2016) were used in this study. Four minimal pairs of English CVC words, which contrast in terms of the voicing of coda stops, were used (*bed-bet, ped-pet, bad-bat, pad-pat*). As shown in Table 1, each target word was placed in a carrier sentence that was an answer to a question in a mini discourse situation. The mini discourse was used to induce the planned focus conditions—i.e., the target words were either phonologically contrasted or not (focused versus unfocused) within the discourse. The target words also occurred in two boundary-related locations, being placed either at the beginning or in the middle of an intonational phrase (IP-initial versus IP-medial). Notice that analysis of the boundary-related effects is beyond the scope of this study, and Choi *et al.* (2016) reported no effect of the preceding boundary on coda production in any meaningful way. Nevertheless, we included this condition to examine focus-related effects in multiple contexts, which were reflected in the statistical analyses.

Before the recording sessions, all participants spent about 25 min practicing the 30 trials to get familiar with the focus types in the different mini discourse situations. During recording, participants first silently read the mini discourse on a computer screen, and then listened to the pre-recorded prompt question as spoken by a female native speaker of American English. They were asked to respond vocally to the question using given sentences. The speed of the mini-discourse presentation and prompt questions was controlled by the experimenter (the first author Z.D.).

The speech data of 5 participants were recorded in a soundproof booth at Hanyang Institute for Phonetics and Cognitive Sciences of Language in Seoul, South Korea, with a Tascam HP-Ps digital recorder and a SHURE KSN44 microphone at a sampling rate of 44.1 kHz. Due to the outbreak of the Covid-19 pandemic, the speech data of the other 15 participants were collected online. The participants lived in China and were asked to record their responses in a quiet room. For the online recordings, participants were asked to install three software programs on their laptops before the experiment: AUDACITY or PRAAT for sound recording, Zoom or Voov meeting for mini-discourse presentation and on-line communication with the experimenter, and TEAM VIEWER or SUNFLOWER to allow the experimenter to control the recording procedure on the participant's laptop. The recording was saved directly onto each participant's laptop using AUDACITY or PRAAT with 32-bit quantization and a 44.1 kHz sampling rate.

The set of sentences was repeated three times in randomized order for each participant. In that way, 1920 tokens (20 speakers × 8 target words × 2 focus types × 2 positions × 3 repetitions) were obtained from the Chinese learners of English. Fourteen tokens were discarded due to inadequate prosodic juncture around the test word or misplacement of focus. For Korean learners of English, Choi *et al.* (2016) collected 2304 tokens (24 speakers × 8 target words × 2 focus types × 2 positions × 3 repetitions) and discarded 5 due to misplacement of boundary or focus, leaving 2299 tokens for analysis.

2.2 Measurements and statistical analysis

The duration and F0 of the vowels preceding a voiceless or voiced coda were measured using PRAAT (Boersma and Weenink, 2021). In line with Choi *et al.* (2016), vowel duration (henceforth V-duration) was the interval from the onset of the voicing of the vowel to the vowel offset, which was defined as the offset of F2. The F0 value was measured at the mid-point of the vowel and converted to semitones relative to the base frequency of 100 Hz using the Eq. $(12) \times \log2$ (f0 value)/100. It is worth noting that vowel duration excluding a release (aspiration) component may undermine the validity of a cross-linguistic comparison, especially if the two language groups produce different degrees of aspiration for the voiceless onset consonants. We carried out a separate statistical analysis to examine the interaction between Native Language (Chinese, Korean) and Onset Voicing (voiceless, voiced), and confirmed that there was neither a main effect of Native Language nor an interaction between the two factors (see supplementary materials[1]).

Table 1. Example sentences with the target "bad." The target word is underlined. Focused words are in uppercase.

Focus	Boundary	Test sentences
Focused	IP-initial	Q: *Did you write "BAT fast again"?* A: *Not exactly. "<u>BAD</u> fast again" was what I wrote.*
	IP-medial	Q: *Did you write "say BAT fast again"?* A: *No, I wrote "say <u>BAD</u> fast again."*
Unfocused	IP-initial	Q: *Did you write "bad SLOWLY again"?* A: *Not exactly. "<u>bad</u> FAST again" was what I wrote.*
	IP-medial	Q: *Did you write "say bad SLOWLY again"?* A: *No, I wrote "say <u>bad</u> FAST again."*

A series of linear mixed-effects models was fitted to each measure using the lmerTest (Bates *et al.*, 2015) package in R (R Core Team, 2020) to examine the effects of coda voicing, prosodic focus, and language proficiency on V-duration and F0. The fixed effects were Focus (focused, unfocused), Coda Voicing (voiceless, voiced), L2 Proficiency (intermediate, advanced), Native Language (Chinese, Korean), and their interactions, while the two dependent variables were V-duration and F0 of the vowel preceding the coda. As explained in Sec. 2.1, the Boundary factor (IP-initial versus IP-medial) was added only as a control factor and was therefore not analyzed in this study. All fixed effects were contrast coded. The random intercept was set for Subject and Item, and the slopes included all the fixed effects as long as a model converged. The detailed statistical results are provided in supplementary materials.[1]

3. Results

3.1 Coda voicing effects on V-duration (duration of the preceding vowel)

There was a significant effect of Coda Voicing on V-duration ($\beta = 14.6$, $t = 7.998$, $p < 0.001$), indicating that it is significantly longer before the voiced than the voiceless coda. Crucially, while the Native Language factor itself did not return significance ($\beta = 2.96$, $t = 0.48$, $p = 0.63$), there was a significant interaction between Coda Voicing and Native Language ($\beta = 14.6$, $t = 7.998$, $p < 0.001$). As shown in Fig. 1(a), the interaction was due to the fact that the Coda Voicing effect was weaker for the Chinese than the Korean learners (mean difference, 6.5 vs 22.7 ms). The Coda Voicing effect also interacted with Focus and Native Language, with a two-way interaction of Coda Voicing × Focus ($\beta = 14.19$, $t = 5.02$, $p < 0.001$) and a three-way interaction of Coda Voicing × Focus × Native Language ($\beta = 12.57$, $t = 2.22$, $p < 0.05$). Those interactions stem from the finding that, although the Coda Voicing effect is larger in the focused condition than the unfocused condition in both the Chinese and Korean learner groups, the Korean learners showed a more robust effect even in the unfocused condition, as can be inferred from Figs. 1(b)–1(c). It is also worth noting that the L2 Proficiency level (intermediate versus advanced) was not involved in any additional interaction effect, indicating that the Native Language effect was not further conditioned by learners' L2 level.

3.2 Coda voicing effects on F0 (semitones) in the preceding vowel

There was a significant effect of Coda Voicing on F0 (semitones) in the preceding vowel ($\beta = 0.34$, $t = 5.28$, $p < 0.001$), in such a way that F0 was lower before the voiced coda than the voiceless one. Unlike the case with V-duration, F0 revealed a significant effect of Native Language ($\beta = 4.12$, $t = 2.77$, $p < 0.01$), indicating that the Chinese learners, on average, produced the target words with a significantly higher F0 than did the Korean learners [Fig. 2(a)]. But as was the case with V-duration, the Coda Voicing effect interacted with Native Language ($\beta = -0.37$, $t = -2.87$, $p < 0.01$). As shown in Fig. 2(a), the interaction was due to a larger Coda Voicing effect on F0 for the Korean than the Chinese learners. *Post hoc* comparisons further indicated that the Coda Voicing effect on F0 was significant in the Korean learner group ($\beta = -0.52$, $t = -4.93$, $p < 0.001$) but not in the Chinese learner group ($\beta = -0.15$, $t = -1.32$, $p = 0.19$), indicating that the Chinese learners did not use F0 to distinguish coda voicing, at least not as robustly as did the Korean learners. As was the case for V-duration, the Coda Voicing effect on F0 interacted with Focus and Native Language in a two-way interaction of Coda Voicing × Focus ($\beta = 0.52$, $t = 4.41$, $p < 0.001$) and a three-way interaction of Coda Voicing × Focus × Native Language ($\beta = -0.77$, $t = -3.32$, $p < 0.01$). Those interactions can be accounted for by the following patterns. As in Figs. 2(b) and 2(c), the Coda Voicing effect on F0 was significant in the focused condition but not in the unfocused condition in both the Chinese and Korean learner groups, accounting for the Coda Voicing × Focus interaction. This indicates that both Chinese and Korean L2 learners use F0 only when focus-induced prominence is expressed. But as can also be seen in Figs. 2(b) and 2(c), the focus-induced prominence effect is much more robust in the Korean than in the Chinese learner

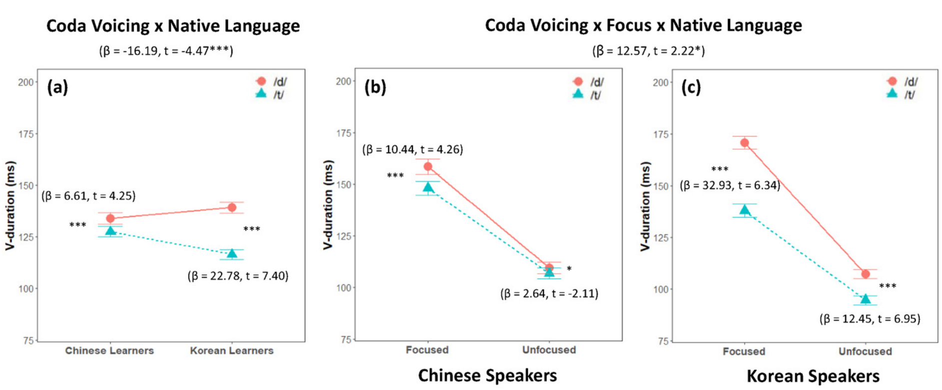

Fig. 1. Effects of Coda Voicing on V-duration in interaction with Native Language (a) and Focus and Native Language (b), (c). The data for Korean learners were obtained from OSF (2022) in connection with Choi *et al.* (2016).

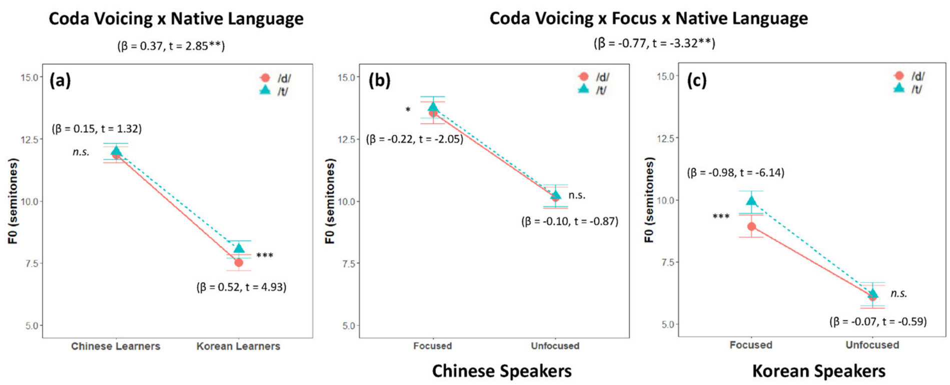

Fig. 2. Effects of Coda Voicing on F0 (semitone) in interaction with Native Language (a) and Focus and Native Language (b), (c). The data for Korean learners were obtained from OSF (2022) in connection with Choi *et al.* (2016).

group, accounting for the three-way interaction of Coda Voicing × Focus × Native Language. Again, as was the case with V-duration, the L2 proficiency level (intermediate versus advanced) was not involved in any additional interaction effect, indicating that the observed patterns for the Native Language effect were not further conditioned by learners' L2 level.

4. Summary and general discussion

One of the most important findings of this study is that L2 speakers with different native language backgrounds, Chinese versus Korean, differ in the details of phonetic implementation of coda voicing contrast in both temporal and spectral dimensions—i.e., vowel duration and F0. Along the temporal dimension, although both Chinese and Korean speakers encode coda voicing contrast in the preceding vowel duration in a way similar to how native English speakers do (de Jong, 2004; Choi *et al.*, 2016), the magnitude of the effect is much smaller for Chinese speakers (5.5 ms difference) than for Korean speakers (25.7 ms difference). A valid question then arises as to whether this native language effect is an artifact of a possible speech rate difference in the production of the two native language groups. However, the overall mean V-duration was comparable across the native language groups (no main effect of Native Language on V-duration), indicating that the effect does not reflect speech rate differences between the native language groups. Moreover, the results show that the difference as a function of L2 speakers' native language is crystalized along the spectral (F0) dimension in a rather unexpected direction. As the Coda Voicing × Native Language interaction indicates, whereas the Korean speakers use F0 quite robustly in encoding the coda voicing contrast, the Chinese speakers barely use it, despite the fact that Chinese speakers have native language experience with tonal contrasts along the F0 dimension (see below for further discussion). These results suggest that Korean speakers use both primary and secondary phonetic features (vowel duration and F0, respectively) much more robustly in marking coda voicing contrast in L2 English than do Chinese speakers. Here, again, the question of whether such a native language background effect can be at least partially accounted for by a possible difference in L2 proficiency between the Korean and the Chinese learner groups is valid. Although we cannot entirely rule out that possibility, it is unlikely because the groups had comparable TOEFL scores, and the native language effect on coda voicing contrast (as evidenced by the Coda Voicing × Native Language interaction) did not interact with speakers' L2 proficiency level—i.e., there was neither a three-way interaction of Coda Voicing × L2 Proficiency × Native Language nor any other interaction that involved L2 Proficiency.

As mentioned in the Introduction, one possible attribute of observed native language effects could have to do with position-specific phonological richness—i.e., the way that each language uses coda position for phonological voicing contrasts. Given that Korean allows multiple phonological contrasts in the coda position (though they are neutralized in voicing-related features) and Chinese does not, it is reasonable to assume that Korean speakers of L2 English have increased phonological awareness of the coda voicing contrast in English, allowing them to be attuned to the non-native coda voicing contrast better than Chinese speakers. This possibility is in line with the more extensive use of vowel duration found in Spanish speakers of L2 English compared with Chinese speakers (Flege *et al.*, 1992), which could be driven by the phonological differences between the two languages—i.e., Spanish allows richer phonological contrast in coda position than does Chinese. More crucially, in contrast with a general assumption that speakers of a tonal language are better attuned to F0-related features in non-native perceptions [e.g., Wayland and Guion (2004); see Holliday (2015) for related discussion], this study shows that the Chinese speakers used F0 in marking the coda voicing contrast in much less magnitude than the Korean speakers, even in the focused condition where hyperarticulation was expected. This could be in part because the F0 difference used for coda voicing contrast in L2 English is rather subtle and only one of many secondary features; thus, it might not be categorically substantial enough to be comparable to the phonological tonal events used in Chinese. Holliday (2015) in fact demonstrated that Chinese speakers' use of rather substantial F0 differences in association with a three-way stop contrast in L2 Korean was no better than its use by English speakers of L2 Korean. This study builds on that finding and indicates that native Chinese speakers might not necessarily benefit from their experience with native tones in producing L2 contrast cued by non-native F0 differences, whether primary or secondary. Rather, the results dovetail better with the view of So and Best (2010) that "[T]he phonemic status and the phonetic features (similarities or dissimilarities) between the tonal systems of the target language and the listeners' native languages play critical roles in the perception of non-native tones" (p. 1).

Nonetheless, the question remains as to why the Korean speakers showed more extensive use of the spectral F0 cue. Although we have no compelling explanations to offer, it might reflect the different levels at which F0 is used in the two languages—i.e., Korean uses F0 at the segmental level for phonological stop contrast, whereas Chinese uses it at the lexical level. Because F0 is used at the segmental level in L2 English, it might be that the Korean learners benefit from their native use of F0 at the segmental level, which might have translated into their use of F0 at the segmental level in L2 English. In addition, stops are never released in the coda in Korean, and Korean listeners tend to pay particular attention to the spectral cues available in the preceding vowel to recover some phonological information about coda stops in non-native speech (Cho and McQueen, 2006). This language-specific aspect might have been reflected to some extent in the Korean learners' production of L2 English. To draw more firm conclusions, speakers of other native language backgrounds (tonal versus non-tonal) using F0 in L2 English should be compared with the Chinese and Korean speakers tested in this study.

Finally, the results presented here show a prominence-related effect in relation to information structure, as reflected in the focused versus unfocused conditions. That is, speakers' Native Language effect on Coda Voicing contrast in L2 English interacts with Focus in a three-way interaction (Coda Voicing × Native Language × Focus). The Coda Voicing effects on vowel duration and F0 are much larger in magnitude in the focused condition than the unfocused condition in both native language groups. As discussed in Choi *et al.* (2016), the focus-related results could be driven by two interactive principles of the linguistic communicative system: *contrast maximization* and *effort minimization* [e.g., Lindblom (1990) and Flemming (1995)]. In other words, the speaker expends effort to maximize a certain phonological contrast (*hyperarticulation*) when needed (i.e., in the focus-induced context driven by information structure), whereas speaker effort is minimized (*hypoarticulation*) in the unfocused condition to which less or no attention is required. That interaction between communicative principles appears to operate differentially on primary and secondary phonetic features—i.e., vowel duration and F0, respectively. On the one hand, although the magnitude of the difference in vowel duration is substantially attenuated in the unfocused condition, the difference remains significant, showing minimal maintenance of phonological contrast along the primary phonetic dimension even in the unfocused condition. On the other hand, along the secondary phonetic dimension, the F0 difference is not maintained in the unfocused condition in either the Chinese or Korean speaker group. This suggests that the focus-induced communicative efficacy driven by information structure is obtained by weighting the phonetic realization of primary versus secondary features differentially in L2 speech, as is likely the case in L1 speech [see Choi *et al.* (2016) for related discussion].

In summary, this study builds on previous L2 phonetic research and provides some specific phonetic evidence for how differentially coda voicing contrast in L2 English is phonetically encoded along the primary (vowel duration) versus the secondary (F0) phonetic dimensions as a function of L2 speakers' native language, Chinese versus Korean. It is proposed that specific phonetic features of the coda voicing contrast phonetically implemented in L2 speech can be understood in relation to the position-specific phonological richness of a speaker's native language. The results also confirm that the speakers of a tonal language might not necessarily learn better an F0-related cue for non-native phonological contrast than do speakers of a non-tonal language, and they might even use it less efficiently when the use in L2 is not comparable to that in their native language. A broader implication of this study is that L2 speech is modulated not only by the phonology of L2 speakers' native language, but also by the detailed phonetic realization on both the segmental and suprasegmental dimensions in a communicatively optimized way as in L1 speech.

Acknowledgments

We would like to thank the Mandarin Chinese speakers who participated in our experiment. We thank an anonymous reviewer for insightful and constructive comments, and the Editor Donald Derrick for giving us sufficient time to revise this manuscript. This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (Grant No. NRF-2020S1A5A2A03042969). This work was also supported by research funds from Hanyang University (Grant No. HY-2021).

References and links

[1]See supplementary material at https://www.scitation.org/doi/suppl/10.1121/10.0013006 for detailed statistical results.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (**2015**). "Fitting linear mixed-effects models using lme4" J. Stat. Software **67**(1), 1–48.
Boersma, P., and Weenink, D. (**2021**). "Praat: Doing phonetics by computer (version 6.2.04) [computer program]," https://www.fon.hum.uva.nl/praat/ (Last viewed July 26, 2022).
Chen, M. (**1970**). "Vowel length variation as a function of the voicing of the consonant environment," Phonetica **22**, 129–159.
Chen, Y., and Gussenhoven, C. (**2008**). "Emphasis and tonal implementation in Standard Chinese," J. Phon. **36**(4), 724–746.
Cho, T. (**2016**). "Prosodic boundary strengthening in the phonetics-prosody interface," Lang. Linguist. Compass **10**(3), 120–141.

Cho, T., and McQueen, J. (**2006**). "Phonological versus phonetic cues in native and nonnative listening: Korean and Dutch listeners' perception of Dutch and English consonants," J. Acoust. Soc. Am. **119**(5), 3085–3096.
Choi, J., Kim, S., and Cho, T. (**2016**). "Phonetic encoding of coda voicing contrast under different focus conditions in L1 vs. L2 English," Front. Psychol. **7**, 1–17.
Choi, J., Kim, S., and Cho, T. (**2020**). "An apparent-time study of an ongoing sound change in Seoul Korean: A prosodic account," PLoS One **15**(10), e0240682–29.
Crowther, C., and Mann, V. (**1992**). "Native language factors affecting use of vocalic cues to final consonant voicing in English," J. Acoust. Soc. Am. **92**, 711–722.
de Jong, K. J. (**2004**). "Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration," J. Phon. **32**, 493–516.
Flege, J. E., Munro, M. J., and Skelton, L. (**1992**). "Production of the word-final English /t/-/d/contrast by native speakers of English, Mandarin, and Spanish," J. Acoust. Soc. Am. **92**(1), 128–143.
Flemming, E. (**1995**). "Auditory representation in phonology," Ph.D. dissertation, University of California, Los Angeles, CA.
Francis, A. L., Ciocca, V., Ma, L., and Fenn, K. (**2008**). "Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers," J. Phon. **36**, 268–294.
Hawkins, S., and Nguyen, N. (**2004**). "Influence of syllable-coda voicing on the acoustic properties of syllable-onset /l/ in English," J. Phon. **32**(2), 199–231.
Holliday, J. J. (**2015**). "A longitudinal study of the second language acquisition of a three way stop contrast," J. Phon. **50**, 1–14.
Kim, H., and Jongman, A. (**1994**). "Complete neutralization of manner of articulation in Korean," J. Acoust. Soc. Am. **95**(5), 2924.
Lindblom, B. (**1990**). "Explaining phonetic variation: A sketch of the H and H theory," in *Speech Production and Speech Modeling*, edited by W. J. Hardcastle and A. Marchal (Kluwer Academic, Dordrecht).
OSF (**2022**). https://osf.io/fbms4 (Last viewed July 26, 2022).
R Core Team (**2020**). "R: A Language and Environment for Statistical Computing," https://www.R-project.org/ (Last viewed July 26, 2022).
So, C., and Best, C. T. (**2010**). "Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences," Lang. Speech **53**(2), 273–293.
Wayland, R. P., and Guion, S. G. (**2004**). "Training English and Chinese listeners to perceive Thai tones: A preliminary report," Lang. Learn. **54**(4), 681–712.