Linguistic Research 37(Special Edition): 113-134
DOI: 10.17250/khisli.37..202009.005

The acquisition of acoustic correlates of politeness by native Chinese speakers*

Eunhae Oh · Mao Cui
(Konkuk University)

Oh, Eunhae and Mao Cui. 2020. The acquisition of acoustic correlates of politeness by native Chinese speakers. *Linguistic Research* 37(Special Edition): 113-134. The current paper explores the acoustic correlates of polite speech in Chinese to examine the similarities and differences of phonetic features contributing to the conveyance of politeness between Korean and Chinese. Building on the previous research on different levels of politeness in Korean, Experiment 1 investigated the phonetic characteristics of deferential and non-deferential Chinese utterances produced by eight native Chinese speakers. The results showed that F0, intensity, H1-H2, HNR and duration played important roles in distinguishing deferential from non-deferential speech, which conform to the patterns shown in Korean. Experiment 2 further assessed deferential and non-deferential speech in Korean produced by Chinese learners of Korean under the assumption that Chinese learners will express deference in the L2 without much difficulty due to the shared phonetic knowledge in the L1. The implications of these findings are discussed in terms of cross-cultural multimodal politeness and the implicit acquisition of sociopragmatic knowledge in the L2. **(Konkuk University)**

Keywords politeness, deferential and non-deferential production, phonetic features, second language acquisition

1. Introduction

Both verbal and nonverbal communication requires learning to decode layers of cues conveyed multimodally in a given language community. Understanding how messages are signalled through gesture, facial expressions, and phonetic features is a crucial aspect of social and emotional information processing. Studies have shown how body languages are inherently linked to speech as they unconsciously reveal our internal state (Winters 2005), aid listener comprehension (Knapp and Hall 2009) and augment our ability to process emotional expressions

* This study was supported by Konkuk University in 2019. We are grateful to the anonymous reviewers for their constructive feedback and suggestions.

(Brauer et al. 2001).

The topic on multimodal politeness has recently attracted more attention as it not only allows us to explore the emotional, relational and sociolinguistic elements of language in general but also provides valuable insight into the language-universal or language-specific characteristics of politeness used across languages with various politeness levels. Up to date, politeness markings have been mostly examined through the studies on verbal morphology. In particular, languages that have morphological and lexical honorific markings to express their social hierarchy such as *sup-ni, eyo, sey-yo* in Korean and *desu, masu, sourou* in Japanese have explicit cues for politeness-related meanings (see Lee and Ramsey 2000; Kim 2006; Brown 2011).

Along with the morphological markers, the sound of a speaker's voice can reveal much about the social status of the person being spoken to, regardless of the conversation topic. Recently, the phonetic characteristics of (non-) polite speech have reported how deference and formality can be encoded and decoded in speech production and perception (Shin 2005; Winter and Grawunder 2012; Brown et al. 2014; Ideamru et al. 2019a; Ideamru et al. 2019b). Besides F0 (Loveday 1981; Ofuka et al. 2000; Ito 2004; Campbell 2004; Winter and Grawunder 2012), intensity (Ziarko 2019; Idemaru et al. 2019a), speech rate (Ofuka et al. 2000), and voice quality related features such as shimmer (a measure of frequency instability) (Ito 2004), breathy phonation (Campbell 2004) were also shown to play an important role in expressing politeness.

Among various acoustic cues that are associated with expression of politeness, high pitch, in particular, has been shown to reliably express politeness-related meanings. Especially, Frequency Code (Ohala 1984) and Effort Code (Gussenhouven 2002) proposed that low pitch speech signals dominance and high pitch speech signals subdominance across all languages. Their proposals were supported by languages produced with higher pitch in polite speech such as Japanese (Loveday 1981; Ofuka et al. 2000; Tsuji 2004), Tamil (Brown and Levinson 1987), Dutch, English (Chen et al. 2004), and Mexican Spanish (Orozco 2010).

On the other hand, other studies have shown evidence of some languages using low pitch and intensity to represent a strategy of mitiation (i.e., "prosodic mitigation") (Shin 2005; Winter and Grawunder 2012; Brown et al. 2014; Idemaru

et al. 2019a). In Winter and Grawunder (2012), deferential (to a professor) and non-deferential (to a friend) utterances by 16 Korean speakers and 9 German speakers were compared based on 14 acoustic parameters including F0, intensity, speech rate and voice-quality measurements. They found a similar use of features such as F0, F0 variability, intensity, shimmer and jitter (a measure of intensity instability) between Korean and German speakers, suggesting that some phonetic features in polite speech may be universal. Contrary to previous proposals, however, they found that both Korean and German speakers consistantly lowered F0 and intensity for polite speech.

In Idemaru et al. (2019a) Korean speakers were shown to use intensity as a more consistant and reliable cue than pitch in politeness judgments. They manipulated F0 and intensity of Korean speakers' production of deferential and non-deferential speech to examine the role of pitch and loudness in determining the politeness-related cues in Korean. The results showed greater individual differences across the effects of F0 compared to intensity manipulation, suggesting that intensity may be a more reliable and robust cue in identifying deferential and non-deferential speech. In addition, when Japanese and Korean speakers' production and perception of (non-) deferential speech were compared, both languages shared the phonetic features of a quiet (low intensity) and soft (breathier) voice for deferential speech: pitch in deferential voices was consistently lower in Korean but not in Japanese (Idemaru et al. 2019b).

As for Chinese, studies on the acoustic features of utterances carrying social, emotional and expressive meanings have reported different results. On the one hand, some studies found that Chinese may prefer high pitch when friendliness or informality is expressed (Chen et al. 2004; Cheang and Pell 2009; Zhang and Gu 2011). For example, Cheang and Pell (2009) elicited expressions of sarcasm from 6 native Cantonese speakers and analyzed several acoustic features of sarcastic and sincere utterances, namely, mean F0, F0 range, mean intensity, intensity range, speech rate and HNR (harmonics-to-noise ratio). The results showed that there was a specific set of acoustic cues to express sarcasm in Cantonese such as higher mean F0, narrower F0 range and narrower intensity range compared to the sincere utterances. Zhang and Gu (2011) examined the prosodic characteristics of entire utterances from 16 Chinese undergraduate students and found that there were significant differences in the mean F0 and F0

range between polite and impolite utterances. Stressed words in polite utterances were produced with overall lower F0 values than than their counterparts in impolite utterances.

On the other hand, some suggested that pitch may not be systematically correlated with polite speech in Chinese. Fan and Gu (2016) manipulated mean F0, F0 range and duration of the stimuli in the perceptual experiments of the polite and impolite speech of Mandarin. The results showed that there were very limited effects of manipulated mean F0 and no meaningful effect of manipulated F0 range. However, different sentential duration had a significant effect on the perception of politeness: speeding up the original speech resulted in a continuous decrease in the degree of politeness. On a similar note, Gu, Zhang and Fujisaki (2011) found that polite utterances had a significantly lower speech rate than impolite utterances. However, there was no salient difference in the mean F0 or F0 range on the utterance-level prosodic settings, suggesting that Chinese may not use pitch as an intrinsic feature of politeness speech.

The results of the previous studies seem to suggest that the link between phonetic features and politeness may not be as strong or straightforward as one may have suggested. Rather, politeness is communicated through multimodal channels and one channel could make a greater contribution than the others in certain languages. To the best of our knowledge, not many studies have made thorough observations of the acoustic features of Chinese deferential and non-deferential speech. Thus, we aim to contribute to the existing body of knowledge with the analyses on the correlates of politeness used in Chinese.

Additionally, we examined the Korean deferential and non-deferential utterances produced by Chinese learners of Korean with two-fold aims: to investigate how the cross-linguistic similarities or differences affect the L2 learners' ability to utilize the subtle acoustic differences between the two languages and to examine whether L2 experience can help learners attend to the language-specific cues in the L2. Recently, studies have shown that the ability to recognize certain emotions from the tone of voice in a given language needs to be learned through social interactions. For instance, Chronaki et al. (2018) provided evidence of the role of language experience in emotion recognition during childhood and adolescence. Speech discrimination testing was conducted on three age groups (children, adolescents, adults), using five target emotions

produced by native speakers of four different languages: English, Spanish, Arabic and Chinese. They found that English groups, regardless of age, performed much better at recognizing emotions in their native language, which was interpreted as an "in-group advantage" in vocal emotion. More importantly, the accuracy of vocal emotion recognition was significantly higher for adults than children or adolescents, suggesting that children develop emotional and pragmatic skills in much the same way that they acquire language. The findings provide insight into the implication of language learning experience on the acquisition of implicit sociolinguistic knowledge in the L2. Through the cross-linguistic comparisons, the present study will offer a comprehensive understanding of how the concept of politeness is manifested in speech sound across languages through multimodal speech channels.

2. Experiment 1: Chinese production by native Chinese speakers

2.1 Participants

Eight native Chinese speakers (4 male, 4 female; age range = 17-20), average age = 18.5) were tested in Shandong province in China. None of the participants reported having prior experience of studying or living abroad. The speakers indicated Mandarin as their only native language and English as their second language. None of them reported having a third language, or having any speech and hearing problems. Of the eight speakers, seven were from Shandong province, and one from Gansu, northern China. Participants were paid upon completion of a 30-minute task.

2.2 Speech stimuli and procedure

The Chinese stimuli were the translation of the scenarios used in Brown et al. (2014). Along with ten request scenarios (see Appendix A), Chinese participants were given a picture of a male professor in a suit for a deferential condition and a picture of a young college student, who were described to be

the participant's imaginary best friend, for a non-deferential situation. They were given the instructions and reading materials one day in advance for familiarization. Both the instructions and reading materials were written in Chinese.

For the experiment, the two conditions ("professor" and "friend") were produced back-to-back in order to change the order of the condition with the scenarios randomized within each condition. Participants were seated in front of the microphone and the instructions and reading materials were presented in a laptop via E-Prime software tools. The script was accompanied by a picture of the imagined interlocutor and repetitions were allowed, if needed. The productions were recorded in a recording studio in Shandong, China, using a Samson C03 Multi Pattern Condenser Microphone and an M-audio Profire 610 at a 44.1 kHz sampling rate with 16-bit quantization. Ten sets of two verbally identical clauses were elicited from 8 participants (10 scenarios * 2 conditions * 8 speakers * 2 times), giving a total of 320 utterances.

2.3 Analysis

Using Praat version 6.0.35 (Boersma and Weenink 2017), each utterance was extracted from the dialogues and analyzed with the same statistical methods reported in Idemaru et al. (2019a) for comparisons with Korean. In the current study, a total of 13 acoustic features were measured and analyzed. First, the features related to the vocal folds vibration include median F0, maximum F0, minimum F0 and standard deviation of F0 (measured in Hz). Second, the features concerning voice loudness include median intensity, maximum intensity, minimum intensity and standard deviation of intensity (measured in dB). Standard deviation was measured to assess the variability in the acoustic features across the utterances. Additionally, acoustic parameters of voice quality such as shimmer, jitter, H1-H2 (a measure of breathiness), and the harmonic-to-noise ratio (HNR: a measure of noise in the voice signal) were examined. Lastly, utterance durations were compared between the two politeness conditions.

Linear Mixed-Effects Models were used with packages lme4 version 1.1.13 (Bates et al. 2015) and afex 0.16.1 (Singmann et al. 2016) for the acoustic

analyses. The model included Politeness condition (deference vs. non-deference), Gender and Scenario (10 scenarios) as fixed effects predictors. The model also included the interactions between Politeness condition and Gender as well as Politeness condition and Scenario to assess whether the effects of politeness condition varied by other variables. By-subject and by-scenario varying intercepts and their slopes for Politeness condition was included as Random effects.[1]

2.4. Results

As shown in Table 1, main effects of Politeness condition were found for median F0, median intensity, maximum intensity, minimum intensity, H1-H2, HNR and duration. Because no main effects of Gender, Scenario or interactions were found, the results are interpreted to indicate the effects of Politeness condition were consistent across these phonetic features.

Table 1. Main effects of politeness main effects on the 13 phonetic features and the characteristics of the Chinese deferential production in comparison to the non-deferential production by native Chinese speakers are shown.

Acoustic features	Politeness condition	Deferential speech
Median F0	O	- 8.4 Hz (lower)
Maximum F0	X	-
Minimum F0	X	-
F0 variability	X	-
Median intensity	O	-1.4 dB (quieter)
Maximum intensity	O	-1 dB (quieter)
Minimum intensity	O	-1.8 dB (quieter)
Intensity variability	X	-
Shimmer	X	-
Jitter	X	-
H1-H2	O	+0.6 dB (breathier)
HNR	O	+0.3 (clearer)
Duration	O	+ 142 ms (longer)

Median F0 returned a main effect of Politeness condition ($\chi^2(1) = 9.55$, $p = 0.002$), with substantially lower F0 values (8.4 Hertz on average) in deference

1 In lme4 syntax the full model was as follows: Acoustic features ~ Politeness * (Gender + Scenario) + (1 + Politeness|Subject) + (1 + Politeness|Scenario). P-values are based on likelihood ratio tests.

utterances. The lack of Gender ($\chi^2(1) = 0.00$, p = 0.99) or Scenario ($\chi^2(1) = 0.26$, $p = 0.61$) effects on Politeness condition suggested that lower F0 was used as a consistent cue to indicate deference for Chinese speakers. The results agree with the findings on the lowering of F0 in deferential speech reported in previous research (Winter and Grawunder 2012; Brown et al. 2014; Idemaru et al. 2019b). Although maximum F0 ($\chi^2(1) = 0.61$, $p = 0.43$), minimum F0 ($\chi^2(1) = 1.41$, $p = 0.23$) and variability of F0 ($\chi^2(1) = 0.00$, $p = 0.99$) did not show any main effect of Politeness condition, pitch tended to be less variable (1.2 SD on average) in deference than in non-deferential speech.

For intensity, there was a significant main effect of Politeness condition for median intensity ($\chi^2(1) = 10.828$, $p = 0.0005$) as well as maximum intensity ($\chi^2(1) = 9.14$, $p = 0.0025$) and minimum intensity ($\chi^2(1) = 19.52$, $p = 0.0001$). However, there was no meaningful difference for intensity variability between the two conditions ($\chi^2(1) = 0.27$, $p = 0.60$). Overall, deference speech was produced with lower intensity (1.4 dB quieter on average) than non-deferential speech. Note that the intensity of non-deferential speech ranged from 32.4 dB to as high as 72.5 dB whereas that of deferential speech ranged from as low as 30.7 dB to 71.3 dB, suggesting that Chinese speakers tend to speak more quietly in deferential speech.

As for the features related to voice quality, jitter and shimmer did not return a significant main effect of Politeness condition ($\chi2(1) = 0.00$, p = 0.99 for both). There was, however, a significant main effect of Politeness condition for H1-H2 ($\chi2(1) = 6.92$, p = 0.008) with an average value of H1-H2 around 8.38 dB for deferential speech and 7.7 for non-deferential speech. There was also a marginal but statistically significant effect for HNR ($\chi2(1) = 4.02$, p = 0.04). Considering that breathy voice generally has higher H1-H2 and HNR values than creaky voice (Katz and Assmann 2019), breathy phonation is shown to take an important role of signaling the conveyance of politeness for deferential speech in Chinese. Additionally, utterance duration showed a significant main effect of Politeness condition ($\chi2(1) = 5.6$, p = 0.017). Deference speech was an average 142 ms (0.8 log-transformed unit) longer with less durational variability than non-deference speech.

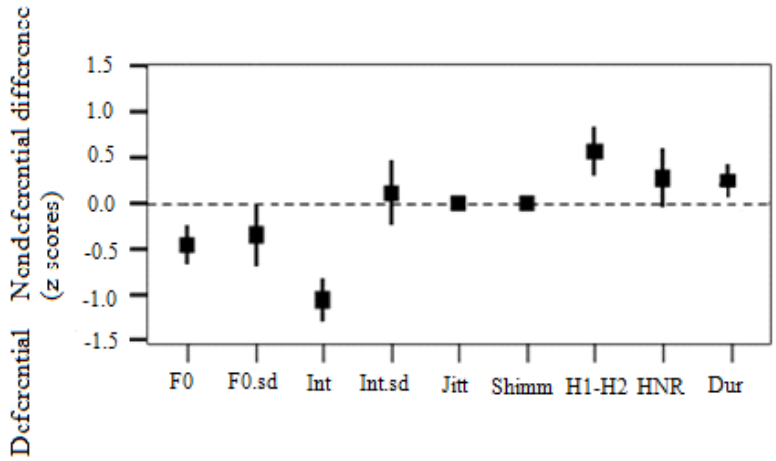

Figure 1. Standardized coefficients from the linear mixed effects model analysis for F0 median, F0 standard deviation, intensity median, intensity standard deviation, jitter, shimmer, H1-H2, HNR, and duration of Chinese production by native Chinese speakers are shown. Positive difference scores indicate that the acoustic variable assumed higher values in deferential speech.

Differences between z-scores ($z = (x-\mu)/\sigma$) of the acoustic features measured in deferential and non-deferential speech are illustrated in Figure 1 (also see Idemaru et al. (2019b) for direct comparisons with Japanese and Korean). The farther away each coefficient is from zero, the greater the difference between the two conditions and positive values indicate that the scores were higher in deferential speech. Overall, median F0 and intensity values were lower for deferential speech and the values for voice quality and duration were higher for deferential speech. Namely, deferential speech was overall lower, quieter, breathier (sounding softer) and slower compared to non-deferential speech.

3. Experiment 2: Korean production by experienced Chinese speakers

In Experiment 1, median F0, median intensity, maximum and minimum intensity, H1-H2, HNR and duration were shown to be the acoustic cues used to distinguish Chinese deferential and non-deferential speech in Chinese. In order to examine whether Chinese learners of Korean transfer these language-specific cues to convey politeness in Korean, we examined Korean deference and

non-deferential utterances produced by experienced Chinese learners of Korean in Experiment 2.

3.1 Participants

Eight experienced Mandarin Chinese learners of Korean (4 male, 4 female; age range = 21-24, average age = 23) participated. A self-report on language background showed that the experienced Chinese speakers, who were studying at a university in Seoul, had resided in Korea for approximately three years (Range = 31-38 months, Mean = 35 months) at the time of testing. They all had received a level between 3 and 4 (out of 6) on the Test of Proficiency in Korean (TOPIK). The participants were taking Korean language classes on a daily basis and showed an intermediate level in Korean proficiency in speaking and reading. All participants were free of speech and hearing problems as determined by self-report.

3.2 Speech stimuli

The stimuli were part of those used in the studies reported in Brown et al. (2014). Five scenarios shown in Appendix B were chosen from ten scenarios used in Brown et al. (2014) in order to make the task easier for the Chinese learners of Korean. Each scenario involved two social conditions: one to an unfamiliar professor and the other to a close friend. Items were blocked by politeness condition and the five scenarios were randomly ordered within each condition. The script was presented with a picture of a male professor in a suit for the "professor" situations and a picture of a young college student for the "friend" situations. The Chinese speakers read each item at least three times and they were allowed to read the same sentence again, if needed. Each script followed a specific design, illustrated by the following example from the "professor" situations:

(1) kyoswu-**n**im
professor-HON

'professor(HON)'

(2) cinan pen-ey malssumha-si-n khemphyuthe phulokulaym-ul
last time-at words:HON-do-HON-MOD computer program-ACC
kwuha-yss-supnita
buy-PAST-HON
'I've bought(HON) that computer program you mentioned(HON) last time.'

(3) kulentey sayongpep-i elyew-ese kule-nuntey
but instructions-NOM difficult-therefore like that-CONJ
'But the instructions are difficult'

(4) pappu-si-kyess-ciman camkkan-man kaluchy-e cwu-si-l swu iss-na-yo?
busy-HON-must-but briefly-only teach-BEN-HON-can-INT-HON?
'I know you must be busy(HON), but can(HON) you teach(HON) me how to use it?'

(from Brown et al. 2014)

As shown above, lines 1, 2 and 4 contain a number of honorific morphemes, such as the *–nim* suffix on the address form kyoswunim 'professor' (line 1), referent honorific *-si-* (lines 2 and 4) and the addressee honorific verb endings *-supnita* (line 2) and *-yo* line 4). These markers were included in the "professor" versions of the scripts, but absent in the "friend" versions. In contrast, line 3 was purposefully designed so that it contained no morphological honorific marking and was scripted identically in the "professor" and "friend" versions. Five scenarios were elicited from eight speakers three times in two politeness conditions (5 scenarios * 2 conditions * 8 speakers * 3 times), giving a total of 240 utterances.

The productions were recorded using a Shure SM 10A head-mounted microphone on a Marantz PMD670 solid-state recorder at a 44.1 kHz sampling rate with 16-bit quantization.

3.3 Procedure

The Experiment was conducted in the sound-attenuated booth in Seoul,

Korea. Participants were given the instructions and reading materials one day in advance to familiarize themselves with the procedures and sentences. Participants were seated in front of a computer monitor and the instructions and reading materials were presented using E-Prime.

3.4 Results

The same linear mixed effects model analysis with Politeness condition, Gender and Scenario as fixed effects and Subject and Scenario as Random effects was used: acoustic features ~ Politeness * (Gender + Scenario) + (1 + Politeness|Subject) + (1 + Politeness|Scenario). Only a significant main effect of Politeness condition was interpreted as a consistent predictor of Politeness across the given acoustic cues. Any interactions between Politeness and Gender or Pliteness and Scenario indicate that the effects of Politeness varied by other variables.

Table 2. Main effects of politeness main effects on the 13 phonetic features and the characteristics of the Korean utterances of deferential production in comparison to the non-deferential production by Chinese speakers are shown.

Acoustic features	Politeness condition	Deferential speech
Median F0	O	- 7.4 Hz (lower)
Maximum F0	X	-
Minimum F0	X	-
F0 variability	O	-2.5 Hz (less variable)
Median intensity	O	-0.9 dB (quieter)
Maximum intensity	O	-1.1 dB (quieter)
Minimum intensity	O	-1.4 dB (quieter)
Intensity variability	X	-
Shimmer	X	-
Jitter	O	- 0.003 (higher stability in freq.)
H1-H2	O	+1.4 dB (breathier)
HNR	O	+0.3 (clearer)
Duration	X	-

As shown in Table 2, results showed a significant main effect of Politeness condition ($\chi^2(1) = 8.59$, $p = 0.003$) on median F0. There was also a significant main effect of Scenario ($\chi^2(1) = 4.52$, $p = 0.03$), suggesting that median F0 was

higher for some scenarios than the others. However, no main effect of Gender ($\chi^2(1) = 2.73$, $p = 0.10$) or an interaction between Politeness condition and Scenario ($\chi^2(1) = 0.05$, $p = 0.82$) was shown. Similarly, standard deviation of F0 returned a significant main effect of Politeness condition ($\chi^2(1) = 5.18$, $p = 0.028$) and Scenario ($\chi^2(1) = 7.38$, $p = 0.007$). The interaction of the two variables was not significant ($\chi^2(1) = 0.07$, $p = 0.79$). To sum, Chinese speakers lowered their pitch (7.4 Hz on average) in deferential speech with considerably less variability (2.5 Hz on average). For maximum and minimum F0, however, there was no main effects of Politeness condition, Gender or Scenario, which suggest that the F0 range was kept relatively constant regardless of the variables.

For intensity, median intensity showed a significant main effect of Politeness condition ($\chi^2(1) = 6.64$, $p = 0.001$) and Scenario ($\chi^2(1) = 4.97$, $p = 0.03$) but no main effect of Gender ($\chi^2(1) = 1.45$, $p = 0.23$). Intensity maximum ($\chi^2(1) = 9.52$, p = 0.002) and minimum ($\chi^2(1) = 14.18$, $p = 0.0002$) also returned a main effect of Politeness condition. The variability for intensity was not significant between the two Politeness conditions ($\chi^2(1) = 0.000$, $p = 0.99$). On average, median intensity was 0.9 dB lower (quieter) in deferential speech than non-deferential speech and both maximum and minimum intensity values were notably lower in deferential speech.

For voice quality, Jitter ($\chi^2(1) = 5.32$ $p = 0.021$), H1-H2 ($\chi^2(1) = 10.38$, $p = 0.001$), and HNR ($\chi^2(1) = 4.04$, $p = 0.03$) showed a significant main effect of Politeness condition. Shimmer, on the other hand, did not show a meaningful main effect of Politeness condition ($\chi^2(1) = 2.75$, $p = 0.10$). None of the these features returned main effects of Gender, Scenario or interactions ($p < .99$). In sum, deferential speech was produced with less variability in frequency (Jitter), a greater degree of breathiness (H1-H2) and a clearer voice (HNR), which are indicative of higher voice quality for speech perception.

Duration was not significantly different between the two Politeness conditions ($\chi^2(1) = 1.84$ $p = .17$). In absolute terms, deferential utterances were longer than non-deferential utterances on average (3512 ms vs. 3153 ms). When log-transformed, however, the difference was insignificant ($\chi^2(1) = 0.19$ $p = .66$).

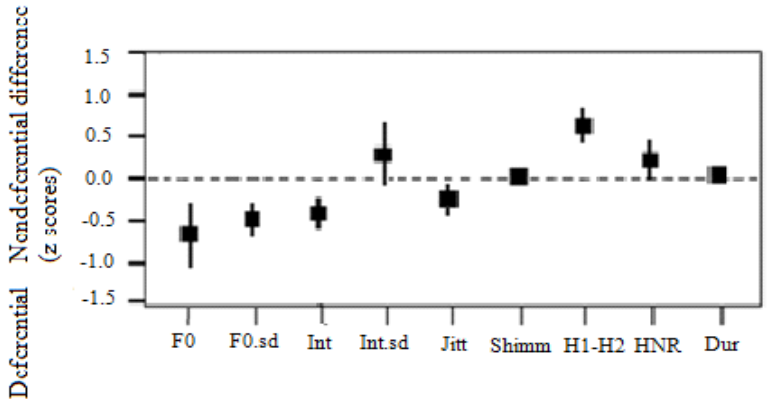

Figure 2. Standardized coefficients from the linear mixed effects model analysis for F0 median, F0 standard deviation, intensity median, intensity standard deviation, jitter, shimmer, H1-H2, HNR, and duration of Korean production by native Chinese speakers are shown. Positive difference scores indicate that the acoustic variable assumed higher values in deferential speech.

Again, differences between z scores of the acoustic features measured in deferential and non-deferential speech are illustrated in Figure 2. Negative values for median F0, F0 variability, median intensity and Jitter indicate that the values for these acoustic variables were lower in deferential speech than non-deferential speech. On the other hand, H1-H2 and HNR revealed higher values in deferential speech. Overall, Chinese learners of Korean produced deferential speech with a lower, quieter, breathier and clearer voice compared to non-deferential speech in Korean.

4. Discussion and conclusion

The current study in Experiment 1 examined the general phonetic characteristics associated with the concept of politeness in Chinese by comparing deferential and non-deferential speech in Chinese. Native Chinese speakers were shown to produce Chinese deferential speech with lower median F0, lower intensity, higher H1-H2, HNR and longer duration. Considering that the just-noticeable-differences (JND) for pitch and intensity are around 7 Hz

(Jongman et al. 2017) and 1 dB (Long 2014) respectively, the differences of 8.4z and 1.4 dB between deferential and non-deferential speech appear to be above the perceptibility threshold. The results are in line with previous studies on Korean politeness which showed low pitch being assoicated with deference and formality (Shin 2005; Winter and Grawunder 2012; Brown et al. 2014; Idemaru et al. 2019b). Idemaru et al. (2019a) showed that low intensity was a more consistent and reliable cue than low pitch in Korean. The main effect of Politeness condition having a significant effect on a greater number of intensity-related cues (i.e., median, maximum and minimum intensity) than pitch-related cues (median F0) suggests that intensity may be a more important and consistent cue compared to pitch in distinguishing deferential from non-deferential speech in Chinese.

The voice characteristics of deferential speech such as higher H1-H2 (an index of breathy voicing) and HNR (an index of vocal clarity) produced in longer utterances are also consistent with the findings in Winter and Grawunder (2012), which reported a greater degree of breathiness and clarity across longer deferential speech in Korean. In Author et al. (under review), when naive Chinese listeners were asked to distinguish deferential speech from non-deferential speech in Korean, H1-H2 and HNR along with pitch were the three most important acoustic features for both Chinese and Korean listeners (see Brown et al. 2014). Under the assumption that the use of similar phonetic information may be due to the similarity in the reliance of vocal politeness cues between Korean and Chinese, Chinese learners of Korean were expected to use of their L1 knowledge to exploit the relevant phonetic features in acquiring politeness cues in the L2 (also see Park 2020).

In Idemaru et al. (2019b), 8 phonetic features were shown to represent the characteristics of deferential speech in Korean by native Korean speakers: median F0, standard deviation of F0, median intensity, standard deviation of intensity, H1-H2, Shimmer, HNR, and duration. In Chinese, median F0, median intensity, H1-H2, HNR and duration made a significant contribution to the perception of deferential speech. Provided that Chinese and Korean speakers pay particular attention to language-specific phonetic features to polite speech, Chinese speakers were expected to transfer most of the common features (i.e., median F0, median intensity, H1-H2, HNR and duration) to their deferential speech in Korean.

The most noticeable difference between the results of Experiment 1 and Experiment 2 is that the Chinese speakers exploited a greater number of phonetic features in Korean utterances (L2) than in Chinese utterances (L1). Along with median F0, median intensity, H1-H2 and HNR, two significant features (i.e., standard devation of F0, and Jitter) newly emerged. Chinese speakers' utterances showed lower pitch variability and a greater control of vocal fold vibration (Jitter) for deferefential speech, both of which concern the degree of pitch perturbation. Namely, deferential speech in Korean was produced with more stable and smoother tone by Chinese speakers. As pitch variation has been found to be greater in casual, non-deferential speech in Korean as well (Winter and Grawunder 2012; Idemaru et al. 2019b), the new addition of the feature may be due to the acquisition of the language-specific acoustic correlates of politeness after some experience in the L2. Dromey et al. (2005) reported that multilingual English L1 users who had the experience of acquiring a second language may be capable of fostering an additional sensitivity to the prosodic characteristics in addition to the linguistic differences compared to the monolingual English L1 users. In return, the additional sensitivity to certain aspects of speech could also be carried over to native language tasks.

Against our expectations for duration to be longer in deferential speech, there was no difference in the length of utterances between deferential and non-deferential speech. However, considering that Chinese speakers read sentences in Korean at considerably slower speech rates than those in Chinese (average time of extracted utterances: 3500 ms vs. 2000 ms), it is likely that the speakers were more focused on reading the passages than "performing" as intended due to their proficiency in Korean. With more experience and higher reading proficiency, duration and speech rate are expected to contribute to the realization of politeness L2 production.

As for the main effects of Scenario for F0 and intensity, post-hoc tests on five scenarios confirmed that Chinese speakers produced deferential speech with significantly lower F0 and intensity for Scenario 2 and Scenario 3. Compared to Scenario 1, 4, and 5, which are about borrowing a book, sending an email and asking for directions, Scenario 2 (canceling and postponing the meeting with Professor) and 3 (asking for Professor Kim's phone number) carry a more apologetic and personal tone, which require greater attention to the polite cues

in the speech. Producing these specific messages with lower F0 and intensity indicate that Chinese speakers were sensitive to these phonetic information as a signal of emotion and were vocally expressed with correlates of politeness in the L2.

The present study explored the phonetic features of the politeness-related speech in Chinese and examined deferential speech in Korean by Chinese learners of Korean. The results provide a glimpse of the similarities and differences between the two languages and how the L1 may influence the acquisition of language-specific phonetic features of politeness in the L2. Compared to pitch, intensity-related features tend to be a stronger and more reliable cue associated with politeness in Chinese - more so than Korean. Chinese learners exhibited a greater number of phonetic features, including F0 variability and jitter, in the production of Korean deferential utterances. More importantly, however, low intensity was found to be a constant indicator of deference in both languages. Deferential speech produced by Chinese learners was overall lower in pitch (with less variability), quieter, and breathier than non-deferential speech. Notwithstanding the notable foreign accents in their utterances, they were able to acquire and produce native-like features in signaling politeness-related meaning in the L2. With higher fluency attained through extended L2 experience, Chinese learners are expected to make fine-tuned adjustments to their perception and production of language-specific cues to politeness.

References

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1-48.

Boersma, Paul and David Weenink 2017. Praat: Doing phonetics by computer [Computer program]. Version 6.0.35.

Brauer, Markus, Jamin Halberstadt, and Åse Innes-Ker. 2001. When did her smile drop? Facial mimicry and the influences of emotional state on the detection of change in emotional expression. *Cognition and Emotion* 15: 853-864.

Brown, Penelope and Stephen C. Levinson. 1987. *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Brown, Lucien. 2011. *Korean honorifics and politeness in second language learning*. Amsterdam: John Benjamins.

Brown, Lucien, Bodo Winter, Kaori Idemaru, and Sven Grawunder. 2014. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues. *Journal of Pragmatics* 66: 45-60.

Caffi, Claudia. 2007. *Mitigation*. Oxford: Elsevier.

Campbell, Nick. 2004. Accounting for voice-quality variation. *Proceedings of 2nd International Conference on Speech Prosody*, 217–220. Nara, Japan.

Cheang, Henry and Marc Pell. 2009. Acoustic markers of sarcasm in Cantonese and English. *The Journal of the Acoustical Society of America* 126: 1394-1405.

Chen, Aoju, Carlos Gussenhoven, and Toni Rietveld. 2004. Language-specificity in the perception of paralinguistic intonational meaning. *Language and Speech* 47(4): 311-349.

Chronaki, Georgia, Michael Wigelsworth, Mark D. Pell, and Sonja A. Kotz. 2018. The development of cross-cultural recognition of vocal emotions during childhood and adolescence. *Scientific Reports* 8: 8659.

Dromey, Christopher, Jose Silveira, and Paul Sandor. 2005. Recognition of affective prosody by speakers of English as a first or foreign language. *Speech Communication* 47: 351–359.

Fan, Ping and Wentao Gu. 2016. Prosodic cues in polite and rude Mandarin speech. In *Proceedings of International Symposium on Chinese Spoken Language Processing*. Tianjin, China.

Gu, Wentao, Ting Zhang, and Hiroya Fujisaki. 2011. Prosodic Analysis and Perception of Mandarin Utterances Conveying Attitudes. *Proceedings of Interspeech 2011*, 1069–1072. Firenze, Italy.

Gussenhoven, Carlos. 2002. Intonation and interpretation: Phonetics and phonology. Aix-en Provence. *Proceedings of the Speech Prosody*, 47-57.

Idemaru, Kaori, Bodo Winter, Lucien Brown, and Grace E. Oh. 2019a. Loudness trumps pitch in politeness judgments: Evidence from Korean deferential speech. *Language and Speech*. Advance online publication.

Idemaru, Kaori, Bodo Winter, and Lucien Brown. 2019b. Cross-cultural multimodal politeness: The phonetics of Japanese deferential speech in comparison to Korean. *Intercultural Pragmatics*. 16(5): 517-555.

Ito, Mika. 2004. Politeness and voice quality - The alternative method to measure aspiration noise. In *Proceedings of the Second International Conference on Speech Prosody*, 213-216.

Jongman, Allard, Zhen Qin, Jie Zhang, and Joan A. Sereno. 2017. Just noticeable differences for pitch direction, height, and slope for Mandarin and English listeners. *The Journal of the Acoustical Society of America* 142: EL163.

Kim, Hyun-Oak. 2006. *Grammaticalization in sentence-final politeness marking in Korean and Japanese*. In Susumo Kuno et al. (eds.), *Harvard studies in Korean linguistics*, 72-85. Seoul, Korea: Hanshin.

Knapp, Mark L. and Judith A. Hall. 2009. *Nonverbal communication in human interaction*. Boston, MA: Wodsworth Cengage Learning.

Lee, Iksop and Robert S. Ramsey. 2000. *The Korean language*. Albany: State University of New York Press.

Long, Marshall. 2014. Human perception and reaction to sound. In *Architectural acoustics*, 81-127. Boston: Academic Press.

Loveday, Leo. 1981. Pitch, politeness and sexual role: An exploratory investigation into the pitch correlates of English and Japanese politeness formulae. *Language and Speech* 24(1): 71-89.

Ofuka, Etsuko, J. Denis McKeown, Mitch G. Waterman, and Peter J. Roach. 2000. Prosodic cues for rated politeness in Japanese speech. *Speech Communication* 32(3): 199-217.

Ohala, John J. 1984. An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* 41(1): 1-16.

Orozco, Leonoro. 2010. *Estudio sociolingüístico de la cortesía en tratamientos y peticiones*. PhD Dissertation. El Colegio de México: Datos de Guadalajara.

Park, Hae In. 2020. How do speakers of different languages differ in the encoding of complex motion events? *Linguistic Research* 37(1): 95-120.

Shin, Suin. 2005. *Grammaticalization of politeness: A contrastive study of German, English and Korean*. Berkeley: University of California.

Singmann, Henrik, Ben Bolker, Jake Westfall, Frederik Aust, S. Højsgaard, J. Fox, M. A. Lawrence, U. Mertens, and J. Love. 2016. Afex: Analysis of factorial experiments. R package version 0.16-1. https://CRAN.R-project.org/package=afex.

Tsuji, Akiko. 2004. The case study of high pitch register in English and in Japanese: Does high pitch register relate to politeness? *Seijo English Monographs* 37: 227–260.

Winter, Bodo and Sven Grawunder. 2012. The phonetic profile of Korean formal and informal speech registers. *Journal of Phonetics* 40(6): 808-815.

Winters, Allison. 2005. Perceptions of body posture and emotion: A question of methodology. *The New School Psychology Bulletin* 3: 35–45.

Zhang, Ting and Gu Wentao. 2011. Corpus design and prosody study for Mandarin affective speech. *Journal of School of Chinese Language and Culture* 2011(3): 33-44.

Ziarko, Ethan. 2019. Prosodic properties of formality in conversational Japanese. *Journal of the International Phonetic Association* 49: 331-352.

Appendix A

Stimuli list

	English translation	Deferential	Non-deferential
1	Do you know the book you mentioned last time? It seems like they don't have the book in the library; can I borrow yours?	教授，还记得您上次提到的书吗？图书馆好像没有那本书。您能把您的借给我看看吗？	李明，还记得你上次提到的书吗？图书馆好像没有那本书。能把你的借给我看看吗？
2	We were going to meet this afternoon (in your office).But something urgent has come up ; can we meet tomorrow instead?	教授，我们原本约好今下午在您办公室见面的，但是我突然有点急事儿，能改到明天见面吗？	李明，我们原本约好今下午在我们教室见面的，但是我突然有点急事儿，能改到明天见面吗？
3	I am taking prof. Kim's class this term. But I have to talk on the phone with Professor Kim today; can you tell me his telephone number?	教授，我这学期选了金教授的课，但是我需要今天给金教授打一个电话。您能告诉我他的电话号码吗？	李明，我这学期选了金教授的课，但是我需要今天给金教授打一个电话。你能告诉我他的电话号码吗？
4	You said that you sent an e-mail to all students/friends last week. But I don't think I received that e-mail; can you send it again?	教授，您说你上周给全班同学发过一封邮件，但是我好像没有收到那封邮件。您能重新发一遍给我吗？	李明，你说你上周给全班同学发过一封邮件，但是我好像没有收到那封邮件。你能重新发一遍给我吗？
5	We are having MT this weekend. But I don't know the MT location exactly; can you tell me where it is?	教授，我们这个周末举行MT活动，但是我不知道MT活动的具体地点。您能告诉我是在哪儿吗？	李明，我们这个周末举行MT活动，但是我不知道MT活动的具体地点。你能告诉我是在哪儿吗？
6	Last class/study group meeting you used the word 'pragmatics' a lot. I don't really know what this word means; can you explain it again?	教授，上次小组讨论时您多次提到"语用学"这个词，但我不清楚这个词到底是什么意思。您能再解释一下吗？	李明，上次小组讨论时你多次提到"语用学"这个词，但我不清楚这个词到底是什么意思。你能再解释一下吗？
7	You know that website you mentioned last time. But I can't find that site; can you let me know the URL one more time?	教授，您上次提到了一个网站，但是我找不到那个网址。您能再告诉我一下那个网络链接吗？	李明，你上次提到了一个网站，但是我找不到那个网址。你能再告诉我一下那个网络链接吗？
8	You said that that book is in the library. But I can't find that book; can you tell me	教授，您说图书馆有那本书。但是我没找到那本书。您能告诉我在哪个区吗？	李明，你说图书馆有那本书。但是我没找到那本书。你能告诉我在哪个区吗？

	where it is?		
9	I've bought that computer program you mentioned last time. But the instructions are difficult; can you teach me how to use it?	教授，我买了您上次提到的那个软件。但是操作说明对我来说有点难，您能教我如何使用那个软件吗？	李明，我买了你上次提到的那个软件。但是操作说明对我来说有点难，你能教我如何使用那个软件吗？
10	I will go to the library and fetch that book.But my bag is a bit heavy; can I leave my bag here?	教授，我要去图书馆拿回那本书，但是我的书包有点沉，我能把书包放在您这里吗？	李明，我要去图书馆拿回那本书，但是我的书包有点沉，我能把书包放在你这里吗？

Appendix B

Stimuli list

	English translation	Deferential	Non-deferential
1	Do you know the book you mentioned last time? It seems like they don't have the book in the library: can I borrow yours?	교수님, 지난 번에 말씀하신 책 있잖아요. 그 책이 도서관에 없는 거 같은데 빌려 주실수 있을까요?	친구야, 지난 번에 말한 책 있잖아. 그 책이 도서관에 없는 거 같은데 빌려줄 수 있어?
2	We were going to meet this afternonn (in your office). But something urgent has come up: can we meet tomorrow instead?	교수님, 오늘 오후에 교수님을 뵙기로 했잖아요. 갑자기 급한 일이 생겨서 그러는데 오늘 대신 내일 뵐 수 있을까요?	친구야, 오늘 오후에 너랑 만나기로 했잖아. 갑자기 급한 일이 생겨서 그러는데 오늘 말고 내일 만날 수 있어?
3	I am taking prof. Kim's class this term. But I have to talk on the phone with Professor Kim today: can you tell me this telephone number?	교수님, 제가 김교수님의 수업을 듣습니다. 그런데 오늘 김 교수님하고 통화해야 할 것 같은데 혹시 김 교수님의 전화번호를 아세요?	친구야, 내가 김 교수님 수업 듣잖아. 그런데 오늘 김 교수님하고 통화해야 할 것 같은데 혹시 김 교수님 전화번호 알아?
4	You said that you sent an e-mail to all students last week. But I don't think I received that e-mail: can you send it again?	교수님, 지난 주에 모든 학생들에게 이메일을 보내셨다고 하셨잖아요. 그런데 그 이메일을 못받은 거 같은데 죄송하지만 다시 한번 보내주실 수 있나요?	친구야, 지난 주에 친구들에게 다 이메일을 보냈다고 했잖아. 그런데 그 이메일을 못받은 거 같은데 미안하지만 다시 한번 보내 줄 수 있어?
5	You said that you sent an e-mail last week. But I don't know the MT location exactly: can you tell me where it is?	교수님, 이번 주말에 엠티 가잖아요. 그런데 엠티 장소를 정확히 몰라서 그러는데 가는 길 좀 알려주세요.	친구야, 이번 주말에 엠티 가잖아. 그런데 엠티 장소를 정확히 몰라서 그러는데 가는 길 좀 알려줘.

* Underlined parts of the utterances are used for analyses.

Eunhae Oh
Associate Professor
Dept. of English Language and Literature
Konkuk University
120 Neungdong-ro, Gwangjin-gu
Seoul, 05029, South Korea
E-mail: grace1111@konkuk.ac.kr

Mao Cui
PhD Student
Dept. of English Language and Literature
Konkuk University
120 Neungdong-ro, Gwangjin-gu
Seoul, 05029, South Korea
E-mail: 623219460@qq.com

Received: 2020. 04. 14.
Revised: 2020. 06. 30.
Accepted: 2020. 08. 29.