Available online at www.sciencedirect.com

ScienceDirect

Journal of Pragmatics xxx (2014) xxx–xxx

journal of PRAGMATICS

www.elsevier.com/locate/pragma

Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues

Lucien Brown [a,*], Bodo Winter [b], Kaori Idemaru [a], Sven Grawunder [c]

[a] East Asian Languages and Literatures, Friendly Hall, University of Oregon, Eugene, OR 97403, USA
[b] University of California, Merced, Cognitive and Information Sciences, 5200 North Lake Road, 95340 Merced, CA, USA
[c] Max Planck Institute for Evolutionary Anthropology, Department of Linguistics, Deutscher Platz 6, 04103 Leipzig, Germany

Received 17 October 2013; received in revised form 26 February 2014; accepted 27 February 2014

Abstract

In languages such as Japanese or Korean, most research on politeness focuses on morphological and lexical honorifics. Here, we ask whether listeners can perceive the intended honorific level of Korean utterances even in the absence of explicit verbal markers, and whether these phonetic cues are available cross-linguistically. We carried out two perception experiments with Korean listeners and also English listeners with no knowledge of Korean. In Experiment 1, stimuli from multiple voices were presented at random and participants had to judge the intended honorific level of isolated stimuli. Overall accuracies were low (58% for Koreans; 53% for English listeners). In Experiment 2, we blocked the presentation of different voices and asked participants to compare honorific and non-honorific speech from the same voice. Accuracies increased to 70% for Koreans and 57% for English listeners, indicating that speech acoustics become an important cue for politeness-related meanings when listeners can compare utterances produced by the same speaker. Our work shows that politeness does not merely reside in verbal markers but is co-signaled by phonetic cues. And, because the English listeners performed above chance on Experiment 2, the results suggest that some acoustic correlates of politeness may be understood in similar ways across cultures.

Keywords: Korean; Honorifics; Politeness; Phonetics; Perception

1. Introduction

The recognition of vocal aspects of (im)politeness dates back to the earliest days of modern politeness theory. Indeed, Brown and Levinson's (1987) seminal work on politeness universals contained a short section dedicated to "phonetics and prosody," where phonetic aspects of politeness in such languages as Tzeltal, Tamil and Basque were discussed. We can thus say that politeness research has long recognized that politeness resides not just in *what* people say, but also in *how* people say something. Hence, we expect the phonetic quality of delivery to be important alongside lexical and morphological politeness formulae.

Despite this, the acoustic analysis of (im)polite language has never flourished within the pragmatics and politeness literature, including phonetic perception studies. Early studies focused mostly on anecdotal observation rather than actual

* Corresponding author. Tel.: +1 541 346 4103.
E-mail addresses: lucien@uoregon.edu (L. Brown), bodo@bodowinter.com (B. Winter), idemaru@uoregon.edu (K. Idemaru), grawunder@eva.mpg.de (S. Grawunder).

http://dx.doi.org/10.1016/j.pragma.2014.02.011
0378-2166/

phonetic measurement. For example, Brown and Levinson (1987:267) suggested that Tzeltal speakers employ high pitch[1] for negative ('deferential') politeness and creaky voice for positive ('friendly') politeness. Corum (1975) noted that palatalization might be linked to negative politeness in such languages as Basque. And Loveday (1981) observed that Japanese women, but not men, raise their pitch when speaking politely. More recently, some studies on impoliteness feature discussion of prosody, including Culpeper's (2005) analysis of how British English television quiz show hosts use intensity, pitch and pauses to create suggestions of impoliteness.

Phonetic studies that incorporate detailed acoustic analyses of polite speech are few in number and often do not address the implications of their findings for politeness research. Notably, Nadeu and Prieto (2011) showed that pitch range and concurrent facial gestures together affected politeness ratings of Catalan participants; similarly Ofuka et al. (2000) found that a final pitch rise leads to increased politeness ratings in Japanese. Campbell (2004) found that breathy phonation was used by Japanese speakers when talking to strangers (for a related study on voice quality, see Ito, 2004). Winter and Grawunder (2011, 2012) attempted to develop a more comprehensive ''phonetic profile'' of honorific speech in Korean. It was found that basically any measurable phonetic aspect covaries with honorific levels, including pitch, loudness, voice quality and speech rate, as well as the occurrence of filled pauses and breathing sounds. This study indicates that politeness affects speech acoustics to large extents, and at a fairly general level.

We believe that integrating mainstream politeness research and phonetic analysis better informs our understanding of the rich system of politeness. From the discursive approach to politeness (see e.g. Grainger, 2011:170), politeness is no longer seen as residing in individual lexical items or grammatical structures. Watts (2003:168) states categorically that ''no linguistic structures are inherently polite''. Indeed, Brown (2013) demonstrated how Korean honorific forms—perhaps the most iconic politeness markers—may be used sarcastically as markers of *im*politeness. Thus, an analysis that purely focuses on lexical and grammatical markers may in some cases be insufficient to explain how an utterance is perceived in terms of (im) politeness. Alongside the undeniable influence of contextual factors, the phonetic quality of an utterance can be expected to play an important role.

Phonetics furthermore has the potential to suggest underlying motivations for politeness phenomena cross-linguistically: Some languages such as Japanese associate high pitch with politeness (Ohara, 2001; Ofuka et al., 2000). Ohala's (1984, 1994) frequency code hypothesis explains the association between high pitch and politeness via a link between high pitch and perceived subdominance (see also Gussenhoven, 2002; Chen et al., 2004). However, in other languages such as Korean (Winter and Grawunder, 2012 – see below) and perhaps Mursi (Irvine, 1979), *low* pitch correlates with politeness-related phenomena. Winter and Grawunder (2012:812) propose that high pitch may also indicate animatedness or arousal, which may conflict with perceptions of politeness in certain languages such as Korean. These findings show how phonetic aspects of speech are tied in with cross-cultural differences in the realization of politeness. However, since studies such as Winter and Grawunder (2012) rely on production data, more work is needed to examine the role of these acoustic cues in the perception of politeness.

The current study assesses the importance of phonetics in politeness perception, focusing on Korean as a test case. Korean contains two main speech registers: an honorific register known as *contaymal* ('respect-speech') and a non-honorific register known as *panmal* ('half-speech'). Whereas honorific *contaymal* is used prototypically when addressing elders, superiors and adult strangers, non-honorific *panmal* is applied when interacting with intimate adults of equal or inferior age/rank, as well as with children.[2] There are in fact numerous sub-levels within Korean *contaymal* and *panmal*, the intricacies of which are beyond the scope of the current paper. However, this simple binary contrast is the most basic and fundamental distinction and the one which Korean speakers are the most sensitive to (Lee and Ramsey, 2000:260).[3]

Politeness is of course a complex phenomenon that has been defined in different ways (see Eelen, 2001:1–29) and that has different cross-linguistic and cross-cultural realizations. However, the use of honorifics (and other social deictic forms) to appropriately mark social relationships has long been recognized as one important mode of politeness. Ide (1989) defined it as ''discernment politeness'', which she contrasted with ''volitional politeness'' (i.e. use of verbal strategies for performing sensitive speech acts). Although the need to encode social position according to ''discernment'' is probably universal, it is

[1] For reasons of simplicity, in this paper we do not separate ''pitch'' and ''fundamental frequency (f0)'', referring to them both as ''pitch''.

[2] The usage of honorific *contaymal* may overlap to some extent with the concept of formal speech. However, although Korean speakers will use *contaymal* when speaking in formal scenes, they will also use it in more casual encounters with elders and superiors. Moreover, *contaymal* speech does not always contain other well-known markers of Korean formality, such as the formal comitative particle *kwa/wa*, which replaces the casual *hako* in formal speech (see Kim and Biber, 1994). Thus, rather than primarily signaling formality, *contaymal* indexes respect for those of superior social standing.

[3] Our description of Korean honorifics as being composed of two registers differs from traditional accounts which focus on the distinction between addressee honorifics, subject honorifics and object honorifics (e.g. Lee and Ramsey, 2000). Note also that we use the term *panmal* according to the layman sense. This usage is broader to how the term is sometimes used within Korean linguistics to refer only to the ''intimate'' speech style.

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

particularly salient in languages such as Korean and Japanese, where the indexing of social relationships is engrained in the grammar of the language (cf. use of the term ''discernment cultures'' by Watts, 1989:132–133).

Even though politeness research has turned toward seeing politeness as a discursive and contextual phenomenon, honorifics are still viewed as an important resource for negotiating politeness. An honorific *contaymal* utterance is not ''polite'' in an absolute sense, but the inclusion or exclusion of honorifics is still intrinsically connected with Korean-specific perceptions and ideologies of politeness.[4] Kim (2011) captures the ideological correlation between politeness and honorifics by stating that ''many Koreans would tell you that there is no way in Korean to express politeness without using honorifics''. Brown (2013:169) further observes that speakers who fail to use honorifics appropriately may be judged negatively against emic conceptions of politeness such as *yayuy eps-nun* 'lacking courtesy', *pelus eps-nun* 'lit. lacking [correct] habits' or *mos paywu-n* 'uneducated'.

The distinction between *contaymal* and *panmal* is traditionally understood as residing in morphology and lexical alternations (for more detailed treatments of the honorific system, see Brown, 2011:19–58; Kim, 2011). The most important and consistent difference is that whereas *contaymal* sentences terminate with the *–(su)pnita* or *–yo* verb ending, *panmal* sentences finish with *–e* or *–ta*. The following two sentences have identical propositional content, but morphology and lexicon indicate the distinction between *contaymal* and *panmal*:

(1) a. *hyeng-nim*, *annyenghi* *cwumwu-si-eyo* [*contaymal*]
older:brother-HON peacefully sleep:HON-HON-HON
'Good night, older brother'

b. *hyeng* *cal* *ca* [*panmal*]
older:brother well sleep
'Good night, older brother'

Important as these morphological and lexical differences no doubt are, the idea that *contaymal* and *panmal* may also be phonetically distinct has been overlooked until recently. In the only study to date that provides a thorough acoustic analysis of *contaymal* and *panmal* in Seoul Korean (previously, Shin, 2005 analyzed pitch only), Winter and Grawunder (2011, 2012) collected samples of the two registers by means of a spontaneous role-playing task. When speaking *contaymal*, speakers generally spoke with lower pitch, smaller pitch range and lower pitch variability. Speakers moreover tended to speak more slowly, and in less breathy voicing indicated by lower values of jitter, shimmer, H1–H2 as well as an increase in harmonics-to-noise ratio. Winter and Grawunder (2011, 2012) interpreted this acoustic profile as indicating a less variable and more monotonous speech style that is also more tensed.

Given that Korean *contaymal* and *panmal* have reliable acoustic correlates in phonetic production studies, this paper sets out to assess the importance of these for the *perception* of honorific levels. Our research questions were:

1. Can listeners perceive the intended honorific level (*contaymal/panmal*) of a Korean utterance from phonetic information alone, that is, when morphological and lexical marking is not available?
2. Can listeners of other languages also perceive the phonetic distinction between *contaymal* and *panmal*, or is the distinction only available to Korean listeners?

We carried out two perception experiments to address these questions: Experiment 1 (see Section 3) and Experiment 2 (see Section 4). We also looked briefly at which acoustic correlates are actually used by listeners in judging the intended honorific level (Section 5).

Across these perception experiments, participants were asked to judge the honorific level of Korean utterances that were devoid of morphological honorifics. Following such cross-cultural perception experiments as Shochi et al. (2007), we recruited American English speaking participants in addition to Korean listeners to see if non-native listeners were able to perceive the intended honorific level in another language. Although we use the American English group as a means of comparison, we should point out that this paper does not specifically look at how English speakers perceive vocal politeness. Rather, American English is taken as one easily available example language for studying cross-cultural differences more generally.

If participants *are* able to differentiate *contaymal* and *panmal* from phonetic information alone, this would indicate that phonetics is an integral aspect of (non-)honorific speech and, more broadly, the communication of politeness.

[4] Note however that some early work on Korean politeness (notably Hwang, 1990) saw honorifics as belonging to the category of ''deference'' rather than ''politeness''. However, placing honorifics outside of the scope of politeness does not appear to have any basis in Korean emic concepts of politeness.

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

Furthermore, if speakers with a different linguistic background can distinguish the difference in Korean, this would point to cross-cultural similarities as to how vocal politeness is perceived.

2. Stimuli construction

Eight native Korean speakers (four male, four female, average age = 28, average stay in the U.S. = 1.5 years, range from 0.1 to 3.0 years) produced the sentences listed in Appendix A. Seven of these speakers were from the Seoul Metropolitan area, one was born in Gyeongsang but moved to Seoul aged seven. All speakers self-identified as mono-dialectal speakers of Standard Seoul Korean.

All stimulus utterances were requests, similar to the one below. Half of them were in non-honorific *panmal*; the other half were in honorific *contaymal* (as in (2)). For the *panmal* sentences, the assumed interlocutor was an intimate friend; for the *contaymal* sentences it was a professor.

(2) 1 kyoswu-**nim**
professor-HON
'*professor(HON)*'

2 cinan pen-ey **malssum**ha-**si**-n khemphyuthe phulokulaym-ul kwuha-yss-**supnita**
last time-at words:HON-do-HON-MOD computer program-ACC 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?*'

All utterances begin with a vocative address form (line 1). For the *contaymal* items such as the one above, this was *kyoswunim* 'professor(HON)'. For the *panmal* items, the participants were asked to supply the name of a real-world friend. Line 2 contains a short sentence contextualizing the request, followed by an explanation or justification for the request (line 3), and then the request head act (line 4).

In (2), all morphological and lexical honorific elements are underlined and in bold. As can be seen, the overall utterance is clearly a request in line with *contaymal* speech. It includes various honorific elements: ''deferential'' *–(su)pnita* and ''polite'' *–yo* speech style endings (lines 2, 4), the suffix *–nim* which renders address terms honorific (line 1), the subject honorific verb ending *–si–* (lines 2, 4) and the honorific noun form *malssum* 'words (of a status superior)'. However, the clause in line 3 is unmarked for honorific level. Taken by itself, there is no lexical or morphological information in this clause that indicates the honorific level, rendering it ambiguous with respect to *contaymal* and *panmal.* We asked speakers to read the full passage, but used only line 3 for the perception experiments.

In total, there were ten different ''scenarios'' (i.e. different contexts and requests), each spoken with *contaymal* and *panmal* (twenty items in total). The items were presented visually in Korean orthography in random order on a computer screen via E-Prime (Schneider et al., 2002). Each speaker read each item twice and they were permitted further productions if they were not satisfied with their performance or if they produced disfluencies. In selecting from the productions, our general policy was to use the final production as stimulus for the perception experiment. In cases where the final production included disfluencies, we used the second-to-last production. The productions were recorded in a sound-attenuated booth using a Marantz PMD670 solid-state recorder and a Shure SM10A head-mounted microphone at a sampling rate of 44.1 kHz and 16-bit quantization.

While Winter and Grawunder (2011, 2012) used role-played speech for their acoustic analysis, here, we had to use read speech because for a controlled listening experiment, honorific and non-honorific speech had to be morphologically and lexically identical. Collecting such stimuli from spontaneous conversation would have been near to impossible. However, because we were concerned that the use of read speech may neutralize some of the phonetic distinctions identified by Winter and Grawunder (2012), we took several measures to highlight the *contaymal*/*panmal* distinction for our speakers. First, we presented *contaymal* and *panmal* versions of the same scenario in pairs. Second, each text was accompanied by a picture of the imagined interlocutor so that the speakers could more easily imagine themselves in an actual communicative situation. For the professor items, we used a picture of a Korean professor which contained a number of semiotic cues for authority, elderliness and scholarliness (suit and tie, glasses, gray hair, fountain pen, books,

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

etc.). For the friend items, we asked speakers to supply a digital image of a friend ahead of the recording session. This image was incorporated into the E-Prime procedure for each speaker. A final related measure that we took was to purposefully use a non-Korean speaking research assistant. This was motivated by the concern that doing the recordings in the presence of an unfamiliar Korean-speaking assistant may create a bias toward *contaymal*.

3. Experiment 1

3.1. Methodology

Twenty native Korean listeners and 20 native Californian American English listeners participated. The Korean listeners (11 male, 9 female; average age = 27; average stay in the U.S. = 2.9 years, range from 0.2 years to 7.0 years) were students at the University of Oregon and received a small payment for their participation. Eleven were from the Seoul area and nine were from other dialect areas of South Korea. All of the nine from the provinces spoke standard Seoul Korean as a second dialect and had been heavily exposed to the Seoul standard through the education system and mass media.[5] No Korean listeners had participated in the stimuli construction. The English listeners (8 male, 12 female; average age = 21) were students at the University of California, Merced, and received extra credit for their participation. Eight of the English listeners were Spanish/English bilinguals who grew up in Central Valley, California. Four others were bilingual (2 Chinese, 1 Danish, 1 Russian). None of the participants reported having any problems with hearing, reading or eyesight.

The stimuli construction described above in Section 2 resulted in 160 stimuli (8 speakers, 10 scenarios, 2 honorific levels – *contaymal* and *panmal*). These stimuli were presented to the listeners via E-Prime. Participants listened through headphones seated in front of a computer monitor in a sound-attenuated booth. Both Korean and English listeners completed the same experiment, although the instructions were in Korean for the Korean listeners and English for the English listeners.

Participants were asked to judge whether each stimulus was spoken to ''someone above the speaker'' (*wi salam*) or to ''someone below the speaker'' (*alay salam*). The notion of someone being ''above'' or ''below'' is prevalent in Korean society and has been claimed to correlate with the use of *contaymal* and *panmal* (Yoon, 2004). We purposefully avoided using the specific words *contaymal* and *panmal* because pretesting the procedure showed that Korean participants would associate this with the presence/absence of honorific morphology. Moreover, focusing on ''above'' or ''below'' the speaker allowed us to use similar category labels for the Korean and the English version of the experiment. The ''above''/''below'' distinction was explained to the English participants through an additional oral explanation as follows: ''Korean society can be very hierarchical, and people are often thought of as being 'above' or 'below' oneself. Those 'above' include professors, workplace superiors and elders. When talking with people who are 'above', you have to be respectful and talk politely.''

The number 8 key on the number pad (the ''up'' arrow key) was assigned for ''above the speaker'' and the number 2 key (the ''down'' arrow key) for ''below the speaker''. These keys were selected due to the metaphorical relationship between verticality and power (Schubert, 2005; Giessner and Schubert, 2007). After each response was registered, there was a 1 s interval before the next trial began.[6] All stimuli were presented randomly, with randomization across all levels of the design (scenarios, honorific levels, speaker voices). This meant that the experiment did not contain blocks or long sequences of the same speaker, the same scenario or the same honorific level. The entire experiment took about 15 min.

Some additional steps were taken to maximize participant performance. In the Korean version, the on-screen instructions contained explicit reference to the fact that the utterances were clauses devoid of honorific marking and that judgments would need to be made based on the sound or feeling of the utterance. Similar to our production study, the research assistant present during the experiment was non-Korean speaking, so that the Korean participants did not have an interaction requiring the use of *contaymal* or *panmal* immediately prior to the experiment. The experiment was preceded by four practice items—identical in format to those used in the test trials—so that the participants had a chance to become familiar with the trial structure and visual stimuli.

[5] Young, educated South Koreans from the provinces are invariably proficient in Seoul Korean due to the vigorous way that this ''standard'' version of the language is promulgated at the national level through the education system and mass media (King, 2006: 278). Indeed, most are fluent enough in Seoul Korean to pass as standard language speakers, reserving use of the dialect only for communicating with family and friends back in the local community.

[6] There were two different versions of Experiment 1. One was used for the first ten Korean participants and the first ten English participants and had no time limit. The other one (used for the remaining participants) had a three-second response time limit, introduced to encourage faster responses. There were no significant differences between these two versions; therefore, we analyze both together as a single experiment. Similarly, we put Korean dialect and non-dialect speakers together, as well as monolingual and bilingual English speakers, as there were no significant differences between these groups.

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

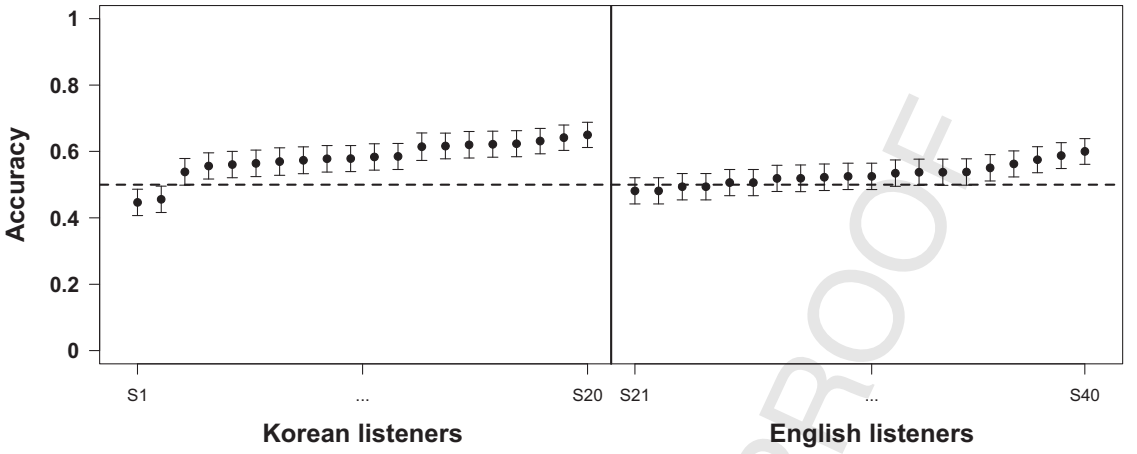

Fig. 1. Korean and English listeners' accuracies with dashed line indicating chance level performance. Each data point represents one participant. Within each language group, listeners are sorted by their accuracy value. Standard errors are computed based on accuracy proportions and do not correspond to the logistic regression reported above.

3.2. Results

We excluded 82 data points that had response times slower than 5 s (1.3% of the total data). Overall estimated accuracies were 58.1% for Korean listeners and 52.9% for English listeners (Fig. 1). A mixed logistic regression with listener, scenario and speaker voice as random effects (accounting for by-listener, by-scenario variation and by-speaker) indicates that English listeners did not perform significantly above chance (log odds: 0.12, SE = 0.076, $p = 0.12$). Korean listeners performed significantly better than English listeners (log odds: 0.21, SE = 0.051, $p < 0.0001$)[7]. Expressed in odds, Korean listeners were about 1.2 times more likely to respond correctly than English listeners.

The above analysis does take individual differences in overall accuracy levels into account, but it does not directly account for bias. For example, a given listener might be more likely to choose the *contaymal* category more often, or the *panmal* one. To this end, we analyzed sensitivity, or d' (d prime) (Green and Swets, 1966). English listeners had average sensitivities of 0.18 (significantly above 0, $t(19) = 4.04$, $p = 0.00069$) and Korean had sensitivities of 0.41 (significantly above 0, $t(19)p > 6.6$, $p < 0.0001$). Across languages, overall sensitivity was significantly different from 0 in a by-listener analysis ($t(39) = 6.8$, $p < 0.0001$), a by-scenario analysis ($t(9) = 5.\ 2$, $p < 0.001$) and a by-speaker analysis ($t(7) = 3.4$, $p < 0.01$), indicating that both Korean and English speakers showed sensitivity better than chance across all scenarios and speakers. As expected, sensitivity was significantly higher for Korean listeners than for English listeners ($t(38) = 3.6$, $p < 0.001$). While these values are above zero, typical d' values are around 2.0 (with a maximum of 6.93), indicating that sensitivity to the *contaymal/panmal* distinction was overall very low.

In the above analysis, we only look at overall accuracy or sensitivity levels. However, our design allows us to examine learning effects (whether listeners become better over the course of the experiment) and speed-accuracy trade-offs (whether listeners' performances were influenced by their response speed) by including trial order and response speed (in log ms) into the analysis. The results showed there was no effect of trial order, indicating that listeners did not become significantly better or worse during the experiment ($p = 0.79$). There was, however, a significant interaction between response speed and language. Korean listeners—more so than English listeners—tended to respond more correctly when they responded relatively faster

[7] All analyses were conducted with R (R Core Team, 2013), using mixed logistic regression and the package *lme4* (Bates et al., 2013). In all models, accuracy (correct vs. incorrect) was the dependent variable. All models include listener, scenario and speaker voice as random intercepts. For both experiments, we initially constructed models that only include the effect of language (and repetition in the case of experiment 2). These models test for overall accuracy regardless of any other variables. After this, a second model was constructed that included the control variables (secondary aspects of the design), such as trial order and log response time, which test for learning effects and trade-offs between accuracy and response speed. This model also included the interaction between language and the control variables, as well as random slopes for the effect of trial order and log RT by subject (Barr et al., 2013). A third model included language, listener gender and speaker gender, as well the two-way interaction between the two gender effects. In additional, separated analyses, we looked at effects of length of stay and dialect for the Korean participants, finding no significant differences. For our analysis of sensitivity, we used one-sample and unpaired *t*-tests, using d' values calculated separately for listeners and speaker voices. The results reported here do not differ if the full data set is analyzed, where no data is excluded because of slow RTs. Unless otherwise noted, *p*-values of the mixed logistic regression are based on Wald's *Z*. However, we only count *p*-values as significant where a likelihood ratio test has also shown that the fixed effect in question matters.

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

(log odd estimate for interaction: −0.21, SE = 0.06, $p = 0.00033$). Correct and incorrect responses were of very similar speed for English listeners (666 ms, 668 ms respectively), but different for Korean listeners (966 ms correct, 1109 ms incorrect). This trade-off between response speed and accuracy has an interesting interpretation for vocal politeness: Having more time to perceive an utterance does not necessarily help. It could be that this is due to over-thinking a choice. Maybe, Korean listeners' faster, more intuitive responses are the more accurate ones. For English listeners—who do not have any experience with Korean—, intuition did perhaps not play a role and predicted accuracies for fast and slow response times are both around 53%.

A model that includes the two-way interaction between listener gender and speaker gender did not fare significantly better than the base model considered above ($\chi^2(3) = 1.05$, $p = 0.79$). This indicates that it was not easier for listeners of a certain gender to hear a voice of the same gender. Similar to the production data by Winter and Grawunder (2011, 2012), the present perception study is characterized by relative homogeneity regarding gender differences.

Finally, we can use likelihood ratio tests to test for random effects, allowing us to see whether there was significant variation in accuracy based on listener, scenario and speaker voice. We did this separately for the Korean and English data. For Koreans, the random effects for listener, scenario and speaker voice were all significant ($\chi^2(1) = 5.22$, $p = 0.022$; $\chi^2(1) = 9.76$, $p = 0.0058$; $\chi^2(1) = 36.44$, $p < 0.0001$). For English, none of these random effects were significant (all $p > 0.1$). This means that there was a significant amount of variance in the data that was explained by the identity of the listener, the type of scenario and the identity of the speaker—but only for Korean listeners. This is, perhaps, unsurprising, given that English listeners already performed at chance level. Another possible reason for this difference between the Korean and English participants is that Korean listeners, with their additional knowledge of Korean and better sensitivity to Korean phonetics, are more affected by small differences between different speakers and scenarios.

3.3. Discussion

English listeners performed the task basically at chance level. Korean listeners were much better, but they still performed with only about 58% accuracy. For comparison, the phonetic contrast between a purely linguistic contrast, such as voiced and voiceless stops is often perceived with more than 95% accuracy (see e.g. Röttger et al., 2014; Idemaru and Holt, 2011), although we may not expect sensitivity levels quite as high as this for a context-sensitive and socially variegated politeness distinction. The accuracy was not modulated by listener gender or speaker gender; however, for Koreans, there were significant differences between listeners, scenarios and speaker voices. Accuracy was furthermore modulated by response speed, but only for Koreans, who tended to respond more accurately when responding quickly.

Given the conservative nature of our design, the overall low accuracy values are perhaps not too surprising: Stimuli were short, decontextualized sentence fragments that would have contained only limited phonetic information. These fragments were then presented with complete randomization across scenarios, honorific levels and speakers. So, one stimulus could be honorific *contaymal*, scenario 3 and speaker 4 (male), the one immediately after that could be non-honorific *panmal*, scenario 7 and speaker 8 (female). This design was chosen so that the participants would be forced to judge honorific level based only on the phonetics of each utterance in isolation. However, as a result of this, the listeners were exposed to a lot of acoustic variability, of which only a small part was due to honorific distinctions (e.g. male speakers' voice pitch in our sample was about 100 Hz, female voice pitch about 183 Hz, much larger than reported acoustic differences for honorific levels, Winter and Grawunder, 2012). The randomized experiment design also disallowed listeners to focus on the characteristics of a speaker's voice. Acoustic correlates of *panmal* and *contaymal* are relative in nature (higher/lower pitch, etc.). Perceiving a *panmal* utterance may depend on knowing how a specific speaker normally speaks *panmal*—and how this differs from how they speak *contaymal*.

In short, Experiment 1 modeled a worst-case situation for perceiving honorific level based on acoustics alone, with everything being randomized and decontextualized. Experiment 1 essentially asked the question: ''Can listeners, without any context to compare, and without any knowledge about the speaker, detect whether a totally random and incomplete utterance fragment was spoken to someone superior or not?'' The fact that even in this context that is biased against the perception of politeness distinctions, Korean listeners were successful in detecting the difference reliably above chance might suggest that in natural discourse with rich contexts and knowledge about the speaker, perception of vocal politeness is easier. Given this, in Experiment 2, we allowed listeners to compare utterances spoken by the same speaker. We now ask a slightly different question: ''Can listeners hear which one of two utterances from the same speaker was intended to be *contaymal* or *panmal*?''

4. Experiment 2

4.1. Methodology

A total of 10 native Korean listeners (3 male, 7 female; average age = 24; average stay in the U.S. = 1.9 years, range from 0.2 years to 6.0 years) and 22 native Californian American listeners (9 male, 13 female; average age = 19)

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

participated in Experiment 2. Four of the Korean listeners came from the Seoul area and the other six from other dialect areas, with all being fluent in standard Seoul Korean. The English listeners included 11 monolinguals and 11 bilinguals (9 Spanish, 1 Hmong, 1 Punjabi). As in Experiment 1, the Korean listeners were based at the University of Oregon and the English listeners at the University of California, Merced. They were compensated with payments and extra credit respectively, as in Experiment 1. None of the participants reported having any problems with hearing, reading or eyesight. None of the listeners had participated in Experiment 1 or in the stimuli construction.

This experiment re-used the same stimuli from Experiment 1. However, two important changes were made to the procedure. First, instead of listening to each individual utterance in isolation, participants listened to utterances in pairs, with a 1 second interval between the two. Each pair was composed of utterances produced by the same speaker and in the same scenario, with one being *contaymal* and the other being *panmal*. In other words, the two utterances were identical except for the change of honorific level (*contaymal*/*panmal*). The task for the participants was to decide which of the two utterances was spoken to ''someone above''. After the auditory stimulus had finished, the numbers 1 (indicating first utterance) and 2 (second utterance) appeared on the screen so that participants could register their choices by pressing the corresponding number key.

Each of the utterance pairs was presented twice (although not successively), with the second appearance featuring the two utterances in the reverse order to the first appearance (e.g. if the first appearance of a speaker/scenario combination was *contaymal* then *panmal*, the second appearance was *panmal* then *contaymal*). For the analyses of Experiment 2, we thus included repetition as an additional fixed effect (everything else was kept the same to Experiment 1), allowing us to see whether listeners become better at the second occurrence of the same utterance pair. Participants had three seconds to respond and there was a 1 s interval between stimuli.

The second important change was that stimuli were blocked by speaker. In other words, participants heard all of the stimuli produced by one of the speakers, then all of the stimuli produced by another speaker and so forth. This ensured that phonetic distinctions produced by individual speakers to signal honorific level would not be obscured by contrasting voice characteristics of different speakers. Because of this change in the experimental design, we now also included the fixed effect within-block trial order on top of overall trial order. This allows us to test whether listeners become better throughout the entire experiment, or whether they become better throughout the course of one block where they are repeatedly exposed to the same voice.

4.2. Results

A total of 93 responses were excluded because of reaching the 3 s response time threshold (1.8% of the total data). Compared to Experiment 1, estimated accuracies were higher for both Korean and English listeners: Korean listeners perceived the intended honorific category 70.0% of the time, English listeners 58.1% (Fig. 2). A mixed logistic regression with listener, scenario and speaker as random effects and repetition as a fixed effect indicates that this performance was significantly above chance for English listeners (log odds: 0.32, SE = 0.096, $p = 0.00066$), as well as for Korean listeners, who outperformed English listeners (log odds: 0.52, SE = 0.12, $p < 0.0001$). Korean listeners were 1.7 times more likely to perceive a given stimulus correctly than English listeners. There was no effect of repetition (log odds: −0.064, SE = 0.059, $p = 0.27$), indicating that listeners did not become better or worse when listening to the same utterance pair the second time.

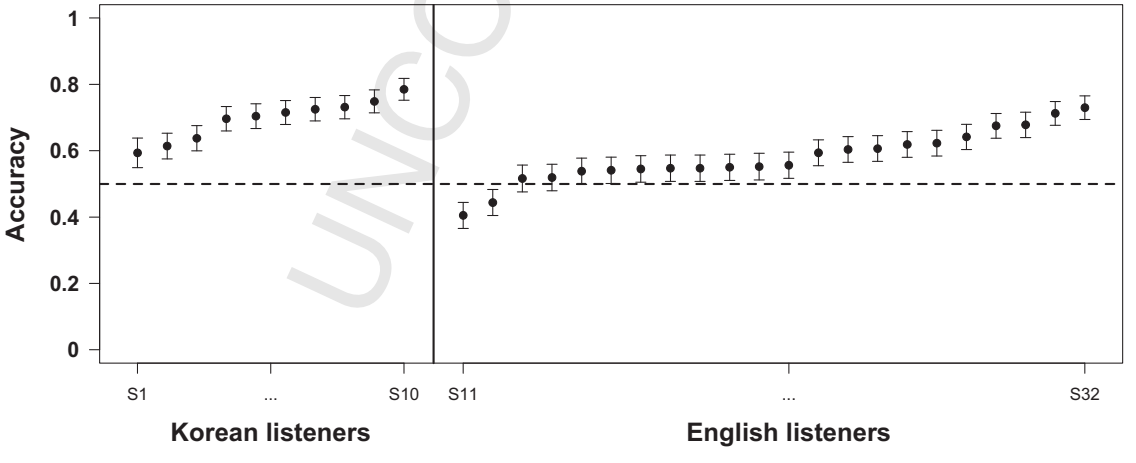

Fig. 2. Korean and English listeners' accuracies by participant. Dashed line indicates chance performance. Error bars indicate standard errors that do not correspond to the logistic regression.

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

This improved performance was also seen in increased sensitivity values (d'), which were 1.06 for Korean and 0.41 for English listeners, in both cases significantly above zero ($t(9) = 9.96$, $p < 0.0001$; $t(21) = 4.63$, $p < 0.001$). The difference in sensitivity between these two listeners groups was significant ($t(30) = 4.2$, $p < 0.001$). Collapsed across listener groups, overall sensitivity was significantly above 0 for a by-listener analysis ($t(31) = 7.06$, $p < 0.0001$), a by-scenario analysis ($t(9) = 8.8$, $p < 0.0001$) and a by-voice analysis ($t(7) = 7.9$, $p < 0.001$).

Collapsing the data across two listener groups, we tested whether adding main effects of overall trial order (1–160) or within-block trial order (1–20 for each speaker voice) had any effect, and, similar to Experiment 1, we also tested whether response speed played a role. There were no effects of overall trial order or trial order within speaker voice, indicating that listeners did not necessarily get better over the course of the experiment, nor did they get better over repeated exposure to the same voice.

Experiment 2 replicated the interaction between language and response speed found in Experiment 1. Again, Korean listeners, but not English listeners, were more accurate for relatively faster responses (log odds: −0.37, SE = 0.097, $p = 0.00014$). Correct and incorrect responses were of very similar speed for English listeners (633 ms, 645 ms respectively), but different for Korean listeners (630 ms correct, 781 ms incorrect).

Similar to Experiment 1, there were no listener gender/speaker gender interaction and main effects. And, similar to Experiment 1, for Koreans, there was significant variability by listener, scenario and speaker voice ($\chi^2(1) = 5.26$, $p = 0.022$; $\chi^2(1) = 19.21$, $p < 0.0001$; $\chi^2(1) = 27.89$, $p < 0.0001$). These random effects were less important in the English data, where only the effect of the listener became significant ($\chi^2(1) = 35.3$, $p < 0.0001$).

4.3. Discussion

The improved listening conditions provided in Experiment 2 resulted in improved performance. Korean listener accuracy changed from 58% in Experiment 1 to 70% in Experiment 2. English listeners improved from 53% to 58%. The English listeners now performed reliably above chance level. These improved accuracies confirm our suspicions that Experiment 1 was too conservative and that total randomization obscured phonetic cues for honorific level, particularly for the Korean listeners. The improved results in Experiment 2 demonstrate that these phonetic cues become more reliable when listeners can compare different utterances made by each individual speaker. The absence of any within-block trial order effect suggests that the difference in task is what drives the difference in accuracy between Experiment 1 and Experiment 2. Because listeners did not become better as a function of repeated exposure to the same voice, it seems that the difference in blocking/randomization was not the critical difference between the two experiments. Rather, the direct comparison of *contaymal* and *panmal* that was now available seems to be the crucial difference. Compared to Experiment 1, the accuracy rates for Korean listeners now at 70% and the sensitivity levels at 1.06, acoustics now appear as a more important cue.

It is noteworthy that the other patterns from Experiment 1 have been replicated in this study: First, there were no gender effects; male and female participants behaved very homogenously, and so did male and female voices. Second, Koreans were better when responding more quickly. Third, for Koreans, differences between scenarios and speaker voices mattered more than for English listeners. This shows that Koreans, presumably because they have more knowledge about the language and more sensitivity to small phonetic detail of the stimuli, are better at tapping into subtle differences between speakers and items. The exact nature of these differences will need to be explored in future studies.

Comparing the data across Experiment 1 and 2, another noteworthy pattern becomes apparent: The fact that listeners in Figs. 1 and 2 can be ordered by performance without any big gaps between listener groups suggests that there is a continuity of performance. There is no evidence for a big split between ''performers'' and ''non-performers.'' This perhaps suggests that across listeners, similar decision strategies were used, and similar cues were being paid attention to.

5. Which acoustic cues were used?

In this section, we explore what cues were relevant to listener's choices, using random forests[8] (Breiman, 2001), implemented by the *party* R package (Hothorn et al., 2006; Strobl et al., 2007, 2008). For a discussion of these techniques in the context of linguistics and sociolinguistics, see Tagliamonte and Baayen (2012).

[8] Random forests is a data mining technique used for classification. It is a so-called ''ensemble method'' because a multitude of decision trees is constructed (500 in this case). Each tree takes a set of variables and sees which variable best splits the data along a criterion (in this case, *contaymal/panmal* in Experiment 1, ''accuracy'' in Experiment 2). Each tree is built on a random subset of variables and data. The final classification is based on the overall ensemble of trees.

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

To get an impression as to how much information is 'in' the acoustics of the speech production data, we constructed a random forest that used different acoustic variables to classify whether an utterance is *panmal* or *contaymal*. Variables considered were based on Winter and Grawunder (2012) and measured as in that study: average pitch (as measured by fundamental frequency in Hz), pitch variability (as measured by fundamental frequency standard deviations), pitch range, loudness (as measured by median intensity), loudness range, local jitter and shimmer (two measures of vocal perturbation, see Winter and Grawunder, 2012), H1–H2 (a measure of spectral slope that is associated with the breathiness of the voice) and harmonics-to-noise ratio (a measure associated with the breathiness of the voice). With these measures, the random forest algorithm correctly classified 85.4% of the stimuli into the honorific levels that were intended by the speaker. This can be taken as an upper bound or 'ground truth' to be used in comparison to the human performance in Experiments 1 and 2. From the perspective that a sophisticated data mining algorithm is able to classify 85.4% of the stimuli based on speech acoustics alone, the Korean listeners in Experiment 2 do very well to achieve 70% accuracy.

Then, we used random forests to predict listener choices in Experiment 1 and listener accuracy in Experiment 2 based on the acoustic variables. For Experiment 1, we used raw acoustic measures. For Experiment 2 we had to use a different approach: As each listener response to be predicted is an accuracy value that is based on a perceptual comparison between two stimuli, we used the difference score of each acoustic variable (acoustic values of *contaymal* minus *panmal*) as a predictor of listener performance. Thus, the acoustic variable used to predict accuracy incorporates both members of an utterance pair, as was the case for the listeners in this task.

For Experiment 1, the procedure was able to predict Korean listeners' choices (*contaymal* vs. *panmal*) 67% of the time, and English listeners' choices 64% of the time. For Experiment 2, the procedure was able to predict the accuracy of Korean responses correctly 73% of the time, and of English responses only 59% of the time. Interestingly, the random forest was much better at predicting accurate responses in Experiment 2 (Koreans: 92%, English: 81%) than inaccurate responses (Koreans: 30%, English: 30%). This seems to suggest that accurate listener responses actually depended on the speech acoustics, but incorrect responses seemed to be due to other factors that cannot be predicted from our acoustic variables. Presumably, incorrect responses are due to such factors as listener-specific biases with respect to particular scenarios and speaker voices, or inattention at particular trials.

Fig. 3 shows the ''variable importance'' calculated based on the random forest analysis (similar to beta coefficient in regression), showing how much a given acoustic dimension was important for correct classification. The acoustic variables are ranked from top to bottom by importance. As can be seen, different acoustic variables are important for Korean and English listeners. In Experiment 1, Korean honorific level choices could be best predicted by H1–H2 (spectral slope), followed by another potentially breathiness-related parameter, the harmonics-to-noise ratio (HNR). English choices were best predicted by intensity, a measure of loudness. HNR and pitch are ranked similarly for Korean and English listeners. It is noteworthy that variable importances were overall much lower in Experiment 2, which suggests that in the direct comparison task, listeners based their choices less straightforwardly on the acoustic measures that we took into account. Listeners clearly had more information available to them in Experiment 2, otherwise they would not have performed better than in Experiment 1. However, even though they fully based their decision on speech acoustics (because they had no helpful cues from honorific markers), that decision was not entirely predictable based on the phonetic features identified in previous research (Winter and Grawunder, 2012) and applied here. Future research will need to see if further phonetic differences exist between *panmal* and *contaymal*, including segmental distinctions and intonation contours.

6. General discussion

Our experiments showed that both Korean and English listeners can perceive the intended honorific level of a Korean utterance from phonetic cues alone. The accuracy was generally low in a strenuous listening condition, with only the Korean listeners performing above chance (58%). However, when participants were given the chance to listen to and compare multiple utterances by the same speaker, the Korean listeners reached higher accuracy (70%) and the English listeners also performed above chance (58%). Although an accuracy of 70% does not seem high when measured against the perception of linguistic contrasts such as the distinction between voiced and unvoiced stops (see Section 3.3), it is high considering that politeness is a more context-sensitive and socially variegated phenomenon. If we compare this accuracy to the perception of something that is much more concretely grounded in measurable physical characteristics, such as vocal speaker height estimation, listeners in our experiments do seem to fare pretty well: Arsikere et al. (2012) report correlations between their height estimation algorithm and actual body size of around $r = 0.5$, which, they ascertain, is the limit of accuracy with which speaker height can be perceived. From this perspective, our accuracies in Experiment 2 are not too low at all, particularly given that we were using read speech stimuli. Since honorific levels are typically triggered by interactional factors such as the identity of the interlocutor and the setting, controlled laboratory stimuli may well underestimate acoustic differences between honorific levels in spontaneous speech. This suggests that speech acoustics might be a meaningful subsidiary cue for politeness-related phenomena in natural discourse.

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

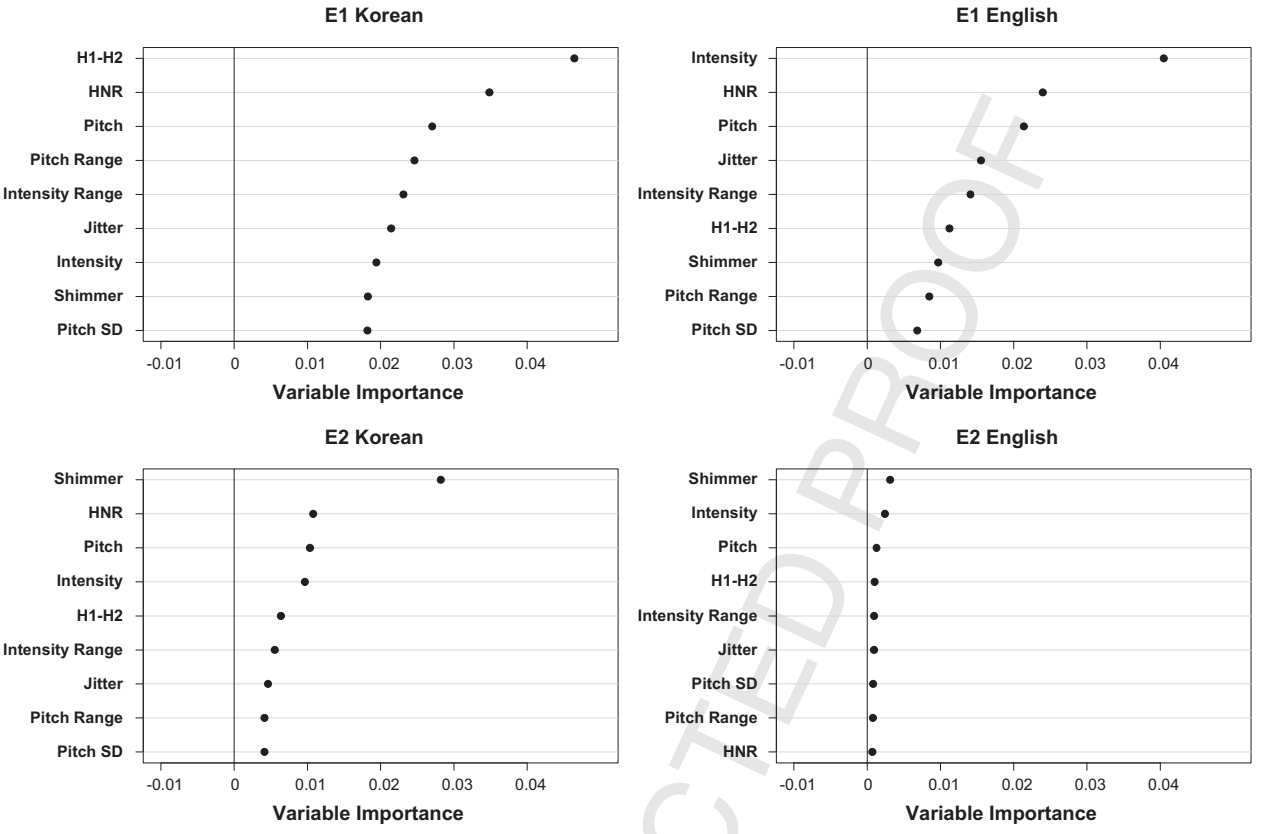

Fig. 3. Variable importance measure generated by random forests for Experiment 1 (top row) and Experiment 2 (bottom row) for Korean listeners (left column) and English listeners (right column).

The improved performance on Experiment 2 shows that interpreting vocal politeness cues heavily relies on comparing utterances with different politeness intentions. In other words, these cues are intrinsically relative in nature. Our claim that phonetic politeness cues work in this way finds support in previous studies on voice perception. Kreiman and Sidtis (2011), for instance, investigated the perception of speaker physical size based on their voice. When participants heard speaker voices in isolation, accuracy rates were around 60%. But when they were given two voices together and asked to indicate which speaker is larger, accuracies increased to around 80%. These results are similar to the accuracy differences reported in this paper for Korean listeners (58%, 70%).

Our experiments have interesting interpretations for the cross-cultural perception of vocal politeness. First, it is noteworthy that English listeners performed above chance at all (in Experiment 2). One could easily expect English listeners to perform at random in judging honorific levels in Korean, a language our English listeners had no experience with. The fact that they did better than chance suggests that there must be some cues for vocal politeness that are common to Korean and English. This is also suggested by the fact that Korean and English listeners paid attention to similar cues in Experiment 1 (see Section 5).

Second, the fact that English listeners were worse than Korean listeners indicates that, even though they use a set of similar cues, the way they use the cues may be different. This was supported by the finding that ranking and importance of the cues were different across Korean and English groups. For example, for Korean listeners, phonation characteristics (H1–H2 and HNR) were ranked highest, whereas for English listeners intensity was ranked highest. These findings indicate that while there may be phonetic features that cross-linguistically signal politeness-related meanings, there are also features that are specific to certain languages and cultures. These findings are consistent with those of prior studies. For example, Shochi et al. (2007) showed that Japanese politeness was often misperceived by the French and American hearers as arrogance or irritation. Further research including listeners of various languages is needed to further investigate which phonetic politeness markers are cross-linguistically shared and which are language specific.

Regarding the comparison of Korean and English speakers in this paper, we note that on top of cross-cultural differences in the acoustic expression of honorific levels, concepts of politeness and ideologies may be different as well. Our experiments focused only on one mode of politeness—discernment politeness. This has been described as more salient in ‘‘discernment

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

cultures'' such as Korea and Japan than in Western cultures, where more stress is placed on strategic or ''volitional'' modes of politeness (Watts, 1989:132–133). Indeed, in a comparison of Korean and American politeness, Koo (1995) found that American participants had little or no awareness of the need to vary speech according to social relationships. Future research will need to explore different modes of politeness and in a broader range of cultural settings.

Finally, what conclusions can we draw from our experiments about the importance of phonetic cues in relation to morphological and lexical honorific markers (or other verbal politeness forms)? Although our research design did not explicitly compare the relative weightings of phonetic and morphological/lexical cues, the accuracy rates of 58% and 70% can still provide some clues. Although these accuracies may not be low in the context of similar phonetic perception experiments (see above), they do not seem to be high enough to suggest that phonetic cues could be the primary means by which Korean speakers distinguish *contaymal* from *panmal* in most conversational contexts. Lee and Ramsey (2000:260) and Yoon (2004:191) note that Korean speakers are extremely sensitive to the *contaymal/panmal* distinction. As an illustration, Koo (1995:17–23) remarks that a student may be expelled from school for failing to use honorifics toward a teacher. For such a sensitive distinction, speakers most probably depend primarily on the more reliable and salient morphological and lexical markers, at least in normative contexts.

Besides, there are more general reasons to doubt whether phonetic cues can be as powerful markers of politeness-related phenomena as morphological and lexical forms. Although no linguistic feature (be that a word, grammatical structure or phonetic cue) is (im)polite in an absolute sense, certain lexical and morphological forms are certainly strongly biased toward a polite or impolite interpretation, even if used by strangers and/or in isolation. As noted by Culpeper (2005:41), ''one has to work quite hard to imagine contexts in which 'you fucking cunt' would not be considered impolite''. Similarly, conventionalized indirect requests in English (''Can you pass the salt?'') are normatively perceived as polite (see Blum-Kulka, 1987), at least when uttered in their prototypical contexts. The same can be said for Korean honorific forms. When applied according to socially normative convention (i.e. in contexts in which they are normally expected), it is difficult to imagine a situation in which they could be interpreted as impolite. It seems unlikely to us that vocal patterns could maintain such close mappings with politeness or impoliteness.

This is particularly the case given that there is no one phonetic aspect that is specialized for politeness. Rather, politeness is communicated by converging phonetic cues (pitch, loudness, voice quality, speech rate), all of which are capable of co-signaling other diverse social meanings (cf. Bryant and Fox Tree, 2005). Low pitch, for example, may be associated with politeness-related meanings in Korean, but can also be associated with masculinity and dominance in other contexts (see e.g. Puts et al., 2007). In real world interactions, phonetic cues will most commonly combine with other cues to communicate politeness-related social meanings. In addition to morphological and lexical markers, these other cues may include facial expressions and gestures (Nadeu and Prieto, 2011)—other areas of politeness as yet under-researched. Politeness is thus communicated in a multi-modal way through the overlaying of various, complementary cues.

In many cases, language users may rely primarily on the stronger verbal cues for politeness. However, it is also important to note that redundancy of cues is crucial in real-world linguistic communication, in which the signal may be less than perfect due to various factors such as background noise and casual articulation (Miller, 1951; Winter and Christiansen, 2012). It is thus possible that phonetic cues signaling politeness may be playing an important secondary role in real-world conversations. There is evidence that listeners across cultures are sensitive to phonetics (e.g. pitch) as a signal of emotion layered on top of the propositional content (e.g. Pell et al., 2009).

Although phonetic cues for (im)politeness thus normally occupy this secondary role, there may be specific contexts in which their role supersedes that of morphological and lexical forms. This may particularly be the case when the verbal forms are used outside of their normal contexts and their normative values as (im)politeness markers become unreliable. In Korean, when honorific *contaymal* is used between close friends, this may simply index formality (Hatfield and Hahn, 2011:1310). However, it can also be interpreted as sarcasm (Brown, 2013). Likewise, use of *panmal* toward an elder can constitute an attempt to sound child-like—a strategy that some speakers may apply when apologizing (Lee, 1996:216–220). However, it can also represent deliberate and scornful impoliteness. We would expect phonetics to be playing an important role in differentiating these different affective meanings, although the details of this need to be empirically established. In the case of sarcastic usage, previous research suggests that phonetics is important in determining whether the sarcastic utterance is intended to be offensive or whether it simply constitutes ''banter'' or ''mock impoliteness'' (see Bousfield, 2008; also Bryant and Fox Tree, 2005 for the role of phonetics in sarcasm).

7. Conclusion

The findings in this paper have important implications for politeness research. Until now, research on politeness-related phenomena has focused overwhelmingly on lexical and grammatical forms. This particularly applies to research on Korean and Japanese, where description of morphological and lexical honorifics has attracted a wealth of previous research. Our paper shows that phonetics also plays an important role in the communication of honorific levels. This is even the case in a language (i.e. Korean) where politeness-related meanings are heavily encoded in the grammar of each sentence.

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

Based on these findings, previous descriptions of Korean (and perhaps Japanese) honorifics may need to be re-thought. Such descriptions frequently suggest that the honorific level of an utterance relies solely on the selection of morphological and lexical forms, with verb endings carrying much of the functional load. It is presumed that a sentence can be altered in honorific level just by changing the verb ending. Lee and Ramsey (2000:260), for example, note how a *panmal* sentence can be rendered into the ''polite'' speech style simply by adding *–yo* to the end, describing *–yo* as a ''simple switching device''. The current study shows that such descriptions overlook the importance of phonetic cues in *panmal/contaymal* shifting. These cues occur throughout the utterance, including parts that are otherwise unmarked for honorification. Rather than relying just on verb ending alternations, our study confirms that the honorific level of an utterance is communicated at a more global level, and that multiple cues should be taken into account.

Finally, there is an applied dimension to these findings as well. Our results may have important implications for the way that politeness phenomena are taught in L2 contexts. In the case of Korean, *panmal*/*contaymal* shifting is frequently taught as a process of adding or deleting verb endings. Similarly, German or French textbooks may teach the intimate/formal speech distinction simply through description of pronoun substitutions and associated morphology. In English classes, learners may be taught that a request can be made ''polite'' by adding ''please'' or selecting indirect over direct verbal strategies. Learners taught in such a fashion may miss the importance of attending to and producing *phonetic* politeness cues. As a consequence, their attempts to use ''polite words'' may fail if the utterance as a whole does not ''sound'' polite. The acquisitional and pedagogical implications of our paper need to be explored in future research.

Our paper highlights the need for further research into the phonetics of politeness, within the remit of pragmatics and politeness research. In addition to being limited to one language (i.e. Korean), our research only considered one mode of politeness (''discernment politeness'')—and then only normative applications of this mode—and relied entirely on read-speech stimuli. More research will thus need to be carried out in order to assess the role of phonetic cues in the communication of different modes of politeness in a wide range of languages and in naturally occurring conversational settings. As discussed above, rather than working in isolation, phonetic cues overlap with lexical and morphological markers as well as other paralinguistic non-verbal properties of communication. Going forward, the study of phonetic politeness cues should thus take place within the wider paradigm of a multi-modal approach to politeness research.

Acknowledgments

We thank all those who helped make this paper happen. At the University of Oregon, we thank Heather Cates and Seong Park for helping run the experiments and recruiting participants. We also thank Keunyoung Lee for her help and input during initial trials of the experiment. At the University of California, Merced we thank Alfonso Montoya Jr. for helping run the experiment and Teenie Matlock for providing lab space. Lastly, we thank our Korean-speaking participants at Oregon and our Californian English speaking participants at Merced.

Appendix A. Stimuli list

Parts of the longer utterances that were used as the stimuli for the perception phase are underlined.

	English translation	*Contaymal*	*Panmal*
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 주실 수 있을까요?	친구야, 지난 번에 말한 책 있잖아. 그 책이 도서관에 없는 거 같은데 빌려 줄 수 있어?
2	We were going to meet this afternoon (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 his telephone number?	교수님, 제가 김 교수님의 수업을 듣습니다. 그런데 오늘 김 교수님하고 통화해야 할 것 같은데 혹시 김 교수님의 전화 번호를 아세요?	친구야, 내가 김 교수님 수업 듣잖아. 그런데 오늘 김 교수님하고 통화해야 할 것 같은데 혹시 김 교수님 전화 번호 알아?

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

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?	교수님, 지난 주에 모든 학생들에게 이메일을 보내셨다고 하셨잖아요. 그런데 그 이메일을 못 받은 것 같은데 죄송하지만 다시 한 번 보내 주실 수 있나요?	친구야, 지난 주에 친구들한테 다 이메일을 보냈다고 했잖아. 그런데 그 이메일을 못 받은 것 같은데 미안한데 다시 한 번 보내 줄 수 있어?
5	We are having MT this weekend. 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 가는 길 좀 알려 주세요.	친구야, 이번 주말에 엠티 가잖아. 그런데 엠티 장소를 정확히 몰라서 그러는데 가는 길 좀 알려 줘.
6	Last class/study group meeting you used the word 'parametric' a lot. I don't really know what this word means; can you explain it again?	교수님, 지난 수업에 "파라메트릭"이라는 말을 많이 사용하셨잖아요. 그 단어가 무슨 말인지 잘 모르겠는데 다시 한 번만 더 설명해 주세요?	친구야, 지난 스터디 때 "파라메트릭"이라는 말을 많이 했잖아. 그 단어가 무슨 말인지 잘 모르겠는데 다시 한 번만 더 설명해 줘.
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?	교수님, 지난 번에 말씀하신 웹사이트 있잖아요. 그런데 그 사이트를 못 찾아서 그러는데 그 웹사이트 URL 한 번 더 가르쳐 주시겠어요?	친구야, 지난 번에 말한 웹사이트 있잖아. 그런데 그 사이트를 못 찾아서 그러는데 그 웹사이트 URL 한 번만 더 가르쳐 줄래?
8	You said that that book is in the library. But I can't find that book; can you tell me where it is?	교수님, 지난 번에 말씀하신 논문집이 도서관에 있다고 하셨잖아요. 그런데 그 책을 못 찾아서 그러는데 그 책이 어디 있는지 가르쳐 주시겠어요?	친구야, 지난 번에 말한 논문집이 도서관에 있다고 했잖아. 그런데 그 책을 못 찾아서 그러는데 그 책이 어디 있는지 가르쳐 줄래?
9	I've bought that computer program you mentioned last time. But the instructions are difficult; can you teach me how to use it?	교수님, 지난번에 말씀하신 컴퓨터 프로그램을 구했습니다. 그런데 사용법이 어려워서 그러는데 바쁘시겠지만 잠깐만 가르쳐 주실 수 있나요?	친구야, 지난번에 말한 컴퓨터 프로그램을 구했어. 그런데 사용법이 어려워서 그러는데 바쁘겠지만 잠깐만 가르쳐 줄 수 있어?
10	I will go to the library and fetch that book. But my bag is a bit heavy; can I leave my bag here?	교수님, 제가 도서관에 가서 그 책을 가져다 드릴게요. 그런데 가방이 좀 무거워서 그러는데 여기에 가방 좀 놓고 갔다 와도 될까요?	친구야, 내가 도서관에 가서 그 책 갖다 줄게. 그런데 가방이 좀 무거워서 그러는데 여기에 가방 좀 놓고 갔다 와도 괜찮아?

References

Arsikere, Harish, Leung, Gary K., Lulich, Steven M., Alwan, Abeer, 2012. Automatic estimation of the first three subglottal resonances from adults' speech signals with application to speaker height estimation. Speech Commun. 55, 51–70.

Barr, Dale J., Levy, Roger, Scheepers, Christoph, Tily, Harry J., 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. J. Memory Language 68 (3), 255–278.

Bates, Douglas, Maechler, Martin, Bolker, Benjamin, 2013. lme4: linear mixed-effects models using S4 classes. In: R Package Version 0.999999-2.

Blum-Kulka, Shoshana, 1987. Indirectness and politeness in requests: same or different? J. Pragma. 11, 131–146.

Bousfield, Derek, 2008. Impoliteness in Interaction. John Benjamins, Amsterdam/New York.

Breiman, Leo, 2001. Random forests. Mach. Learn. 45, 5–32.

Brown, Lucien, 2011. Korean Honorifics and Politeness in Second Language Learning. John Benjamins, Amsterdam/New York.

Brown, Lucien, 2013. "Mind your own esteemed business": sarcastic honorifics use and impoliteness in Korean TV dramas. J. Polite. Res. 9 (2), 159–186.

Brown, Penelope, Levinson, Stephen, 1987. Politeness: Some Universals in Language Usage. Cambridge University Press, Cambridge.

Bryant, Gregory A., Fox Tree, Jean E., 2005. Is there an ironic tone of voice? Language Speech 48 (3), 257–277.

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

Campbell, Nick, 2004. Accounting for voice-quality variation. In: Proceedings of the Second International Conference on Speech Prosody, Nara, pp. 217–220.
Chen, Aoju, Gussenhoven, Carlos, Rietveld, Toni, 2004. Language-specificity in the perception of paralinguistic intonational meaning. Language Speech 47 (4), 311–349.
Corum, Claudia, 1975. Basque, particles and baby-talk: a case for pragmatics. In: Proceedings of the First Annual Meeting of the Berkeley Linguistics Society, pp. 90–99.
Culpeper, Jonathan, 2005. Impoliteness and entertainment in the television quiz show: *The Weakest Link*. J. Polite. Res. 1, 35–72.
Eelen, Gino, 2001. A Critique of Politeness Theories. Manchester, St. Jerome.
Giessner, Steffen R., Schubert, Thomas W., 2007. High in the hierarchy: how vertical location and judgments of leaders' power are interrelated. Org. Behav. Hum. Decision Process. 104 (1), 30–44.
Grainger, Karen, 2011. 'First order' and 'second order' politeness: institutional and intercultural contexts. In: Linguistics Politeness Research Group (Eds.), Discursive Approaches to Politeness. Walter de Gruyter, Berlin/Boston, pp. 167–188.
Green, David M., Swets, John A., 1966. Signal detection theory and psychophysics, vol. 1. Wiley, New York.
Gussenhoven, Carlos, 2002. Intonation and interpretation: phonetics and phonology. In: Bel, B., Marlien, I. (Eds.), Proceedings of Speech Prosody. Université de Provence, Aix-en-Provence, pp. 47–57.
Hatfield, Hunter, Hahn, Jee-Won, 2011. What Korean apologies require of politeness theory. J. Pragma. 43 (5), 1303–1317.
Hothorn, Torsten, Buehlmann, Peter, Dudoit, Sandrine, Molinaro, Annette, Van Der Laan, Mark, 2006. Survival ensembles. Biostatistics 7 (3), 355–373.
Hwang, Juck-ryoon, 1990. 'Deference' versus 'politeness' in Korean speech. Int. J. Sociol. Language 82, 41–55 (Aspects of Korean Sociolinguistics).
Ide, Sachiko, 1989. Formal forms and discernment. Multilingua 8 (2), 223–248.
Idemaru, Kaori, Holt, Lori L., 2011. Word recognition reflects dimension-based statistical learning. J. Exp. Psychol.: Hum. Percept. Perform. 37 (6), 1939–1956.
Irvine, Judith T., 1979. Formality and informality in communicative events. Am. Anthropol. 81, 773–790.
Ito, Mika, 2004. Politeness and voice quality – the alternative method to measure aspiration noise. In: Proceedings of the Second International Conference on Speech Prosody, Nara, pp. 213–216.
Kim, Alan, 2011. Politeness in Korea. In: Kádár, D., Mills, S. (Eds.), Politeness in East Asia. Cambridge University Press, Cambridge, pp. 176–207.
Kim, Young-Jin, Biber, Douglas, 1994. A corpus-based analysis of register variation in Korean. In: Biber, D., Finegan, E. (Eds.), Sociolinguistic Perspectives on Register. Oxford University Press, New York, pp. 157–181.
King, Ross, 2006. Dialectal variation in Korean. In: Sohn, H. (Ed.), Korean Language in Culture and Society. University of Hawaii, Honolulu, HI, pp. 264–280.
Koo, Ja-sook, 1995. Politeness Theory: Universality and Specificity (PhD dissertation). Harvard University
Kreiman, Jody, Sidtis, Diana, 2011. Physical Characteristics and the Voice: Can We Hear What a Speaker Looks Like? Wiley, New York.
Lee, Chang-soo, 1996. Variation in the Use of Korean Honorific Verbal Endings: An Interactional Sociolinguistic Study (PhD dissertation). Boston University
Lee, Iksop, Ramsey, S. Robert, 2000. The Korean Language. SUNY, New York.
Loveday, Leo, 1981. Pitch, politeness and sexual role: an exploratory investigation into the pitch correlates of English and Japanese politeness formulae. Language Speech 24, 71–89.
Miller, George A., 1951. Language and Communication. McGraw-Hill, New York.
Nadeu, Marianna, Prieto, Pilar, 2011. Pitch range, gestural information, and perceived politeness in Catalan. J. Pragma. 43, 841–854.
Ofuka, Etsuko, McKeown, J. Denis, Waterman, Mitch, Roach, Peter, 2000. Prosodic cues for rated politeness in Japanese speech. Speech Commun. 32, 199–217.
Ohala, John J., 1984. An ethological perspective on common cross-language utilization of F0 of voice. Phonetica 41, 1–16.
Ohala, John J., 1994. The frequency code underlies the sound symbolic use of voice pitch. In: Hinton, L., Nichols, J., Ohala, J.J. (Eds.), Sound Symbolism. Cambridge University Press, Cambridge, pp. 325–347.
Ohara, Yumiko, 2001. Finding one's voice in Japanese: a study of the pitch levels of L2 users. In: Pavlenko, A., Brackledge, A., Piller, I., Teutsch-Dwyer, M. (Eds.), Multilingualism, Second Language Learning, and Gender. Mouton de Gruyter, New York, pp. 231–254.
Pell, Marc D., Paulmann, Silke, Dara, Chinar, Alasseri, Areej, Kotz, Sonja A., 2009. Factors in the recognition of vocally expressed emotions: a comparison of four languages. J. Phonetics 37 (4), 417–435.
Puts, David A., Hodges, Carolyn R., Cárdenas, Rodrigo A., Gaulin, Steven J., 2007. Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. Evol. Hum. Behav. 28 (5), 340–344.
R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/
Q2 Röttger, Timo, Winter, Bodo, Grawunder, Sven, Kirby, James, and Grice, Martine, (2014). Assessing the incomplete neutralization of final devoicing in German, submitted for publication.
Schneider, Walter, Eschman, Amy, Zuccolotto, Anthony, 2002. E-Prime Reference Guide. Psychology Software Tools, Inc., Pittsburgh, PA.
Schubert, Thomas W., 2005. Your highness: vertical positions as perceptual symbols of power. J. Person. Soc. Psychol. 89 (1), 1–21.
Shin, Suin, 2005. Grammaticalization of Politeness: A Contrastive Study of German, English and Korean (PhD dissertation). University of California, Berkeley
Shochi, Takaaki, Aubergé, Véronique, Rilliard, Albert, 2007. Cross-listening of Japanese, English and French social affect: about universals, false friends and unknown attitudes. In: Trouvain, J., Barry, W.J. (Eds.), Proceedings of the 16th International Congress of Phonetic Sciences, pp. 2097–2100.
Strobl, Carolin, Boulesteix, Anne-Laure, Zeileis, Achim, Hothorn, Torsten, 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinform. 8 (25). http://www.biomedcentral.com/1471-2105/8/25

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011

Strobl, Carolin, Boulesteix, Anne-Laure, Kneib, Thomas, Augustin, Thomas, Zeileis, Achim, 2008. Conditional variable importance for random forests. BMC Bioinform. 9 (307). http://www.biomedcentral.com/1471-2105/9/307

Tagliamonte, Sali A., Baayen, R. Harald, 2012. Models, forests, and trees of York English: was/were variation as a case study for statistical practice. Language Variation Change 24 (2), 135–178.

Watts, Richard, 1989. Relevance and relational work: linguistic politeness as politic behavior. Multilingua 8 (2), 131–166.

Watts, Richard, 2003. Politeness. Cambridge University Press, Cambridge.

Winter, Bodo, Christiansen, Morten H., 2012. Robustness as a design feature of speech communication. In: Scott-Phillips, T.C., Tamariz, M., Cartmill, E.A., Hurford, J.R. (Eds.), Proceedings of the 9th International Conference on the Evolution of Language. World Scientific, New Jersey, pp. 384–391.

Winter, Bodo, Grawunder, Sven, 2011. The polite voice in Korean: searching for acoustic correlates of contaymal and panmal. In: Sohn, H., Cook, H.M., O'Grady, W., Serafim, L.A., Cheon, S. (Eds.), Japanese/Korean Linguistics 19. CSLI, Stanford, pp. 419–431.

Winter, Bodo, Grawunder, Sven, 2012. The phonetic profile of Korean formality. J. Phonetics 40, 808–815.

Yoon, Kyung-joo, 2004. Not just words: Korean social modes and the use of honorifics. Intercultural Pragma. 1 (2), 189–210.

Please cite this article in press as: Brown, L., et al. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, Journal of Pragmatics (2014), http://dx.doi.org/10.1016/j.pragma.2014.02.011