

Prosodic characteristics of Japanese polite speech spoken by native and non-native speakers

Chiharu Tsurutani¹, Shi Shuju², Nobuaki Minematsu²

¹ School of HLSS, Griffith University, Australia

² Dept. of Information and Communication Engineering, The University of Tokyo, Japan

C.Tsurutani@griffith.edu.au, {shishuju, mine}@gavo.t.u-tokyo.ac.jp

Abstract

We conducted experiments in the production and the perception of Japanese polite speech to explore the effectiveness of prosodic features on the perception of politeness. We recorded common polite forms, in polite and non-polite scenarios using 10 native and 10 non-native speakers, and 10 listeners assessed their level of politeness. The results showed that: 1) native speakers tend to use a lower pitch register, a narrower pitch range and a slower speech rate, 2) there is a certain type of pitch modification observed in polite speech; 3) learners seemed to use only a slower speech rate to express politeness.

Index Terms: polite speech, prosodic features, Japanese, native and non-native speech

1. Introduction

Prosody plays an important role not only in intelligibility but also in speakers' attitudes and emotions. It is a general consensus that intonation can completely change the interpretation of the utterance. Prosodic characteristics in speech conveying emotion and attitude share features across all languages. Ohala [1] argues that the use of high pitch conveys an impression of subordination and submissiveness and is considered to be non-threatening, conveying goodwill. High pitch also leads to high scores in polite, non-aggressive, friendly and other positive attitudes in perception studies [2] [3]. Nevertheless, attitudes associated with social behavior can have language specific prosodic characteristics.

In this study, politeness was chosen to investigate possible language specific requirements in Japanese speech. Japanese people are known to be polite and well-mannered in public. Learning to sound polite is particularly important for second language learners (L2) of Japanese. As Japanese has elaborate honorific systems, politeness studies and guidance for L2 learners tend to focus on morphological and lexical aspects, but the information on appropriate intonation for polite speech has not been made available to L2 learners. This study attempts to identify common prosodic features used by native speakers in polite speech and to provide some strategies useful in delivering polite speech for L2 learners of Japanese. To this end, the following steps were taken, which are used in structuring the paper: After reviewing the phonetic aspects of politeness dealt with in previous studies, an experiment was conducted to collect polite and non-polite speech from both native and non-native speakers. The politeness of their performance was judged by native listeners, and statistically analyzed to identify factors that affect the degree of politeness.

2. Phonetic aspects of politeness

The degree of politeness can be affected by non-linguistic factors such as non-verbal expressions, the contents and context of the utterance. However, after considering all these factors, there are still some phonetic characteristics which can differentiate between polite and non-polite utterances. Brown et al. [4] investigated the significance of phonetic information in Korean without morphological and lexical marking. The perceptual ability of native and non-native listeners who had no knowledge of Korean was tested on a sentence inserted in casual and honorific conversation, respectively. Both native and non-native speakers were able to identify polite speech above chance level, which implies the existence of some cross-linguistic cues for vocal politeness. They used similar cues such as pitch and vowel quality, but the ranking and importance of cues differed between the two groups, which indicates there are also features that are specific to certain languages and cultures. There are several prosodic factors that have been commonly investigated for their influence in expressing politeness in previous studies. These are discussed in the following sections.

2.1. Pitch register (average pitch)

As a general impression, Japanese female speakers are known to use high pitch for polite speech [5] [6]. However, high pitched speech is not always observed in polite speech in all language communities. Grawunder and Winter [7] found that polite speech in Korean had a lower pitch level and narrower pitch range than informal speech. Recently Idemaru, Winter and Brown [8] added new information concerning pitch register in Korean polite speech in terms of perception. They found that female Korean listeners perceived high pitch as being polite, while male Korean listeners perceived it to be informal. Their study used informants (6 females and 4 males) who resided in the US, including one male listener who was born in the US. Although it is difficult to draw a conclusion from this one perception study, it does imply that pitch register is not an absolute determiner of politeness.

2.2. Pitch range

In the study of attitudinal prosody in Mandarin Chinese, it was reported that polite speech had wider pitch, as well as higher register of pitch, than neutral speech [9]. On the other hand, Grawunder and Winter's study on Korean language reported a narrower pitch range for polite speech. The different result can be attributed to different prosodic requirements between the two languages; a tone language vs. a non-tonal language. Or it could be due to the sentence type used as stimuli, whether it is

a question or a statement. Ofuka et al. [10] investigated six Japanese male speakers' production of polite speech and found that both average pitch and pitch range did not become cues for politeness. In another study of Japanese polite speech, increase of pitch range in casual speech addressed to intimate friends was observed in the interview to the participants [5]. Using both male and female speakers and various sentence types, pitch range in Japanese polite speech needs to be reexamined further.

2.3. Speech rate

Speech rate has been regularly discussed in polite speech [7] [10]. The general impression that slow speech sounds more polite has been proven in studies of Korean speech. On the other hand, studies into Chinese and Japanese speech reports otherwise. Wentao et al. [9] reported that polite speech in Chinese was faster. Ofuka et al. [10] did not find any conclusive tendency, apart from listeners' preference for a speech rate similar to their own, even though the rate factor showed some relevance to politeness. Clearly, speech rate is a factor to be investigated. In this study, speech rate will be measured in the form of the duration of the utterance, since it is known that listeners gain an impression of slow speech by listening to the entire speech [11].

2.4. F0 direction and duration of the final syllable of the sentence

In previous studies on Japanese politeness, prosodic information of final particles was paid attention to as a cue for politeness. Japanese has a rich inventory of sentence final particles. Interrogative sentences are expressed with "ka" without changing the word order from a declarative sentence. The intonation is normally raised for question sentences, as is the case for over 70% of the world's languages [12]. The question sentence without "ka" will require a rise in the sentence's final tone more than in unmarked question sentences, which will be the case in other languages as well. The problem is that the pitch rise in the sentence ending with the final particle "ka" does not always occur in actual speech. In Ofuka, et al.'s study [10], 2 out of 6 speakers had a fall in the F0 direction in the final vowel, and the proportion of rise and fall was half and half in Ogino and Hong's data [5]. However, in their politeness judgement test [10], native listeners preferred the unmarked sentence intonation, a final pitch rise, to a final fall. They further investigated the prosody of the final vowel of the sentence and found a great impact on politeness judgements in duration and pitch change, in particular in the last 100ms. The sentence's final prosody is considered to accurately convey the speaker's attitude or intention. In the experiment of this study, pitch register, pitch range, speech rate and the prosodic information in the final syllable of the sentence will be investigated.

3. Experiment

Stimulus sentences were prepared considering the following points: speech style, sentence type, and the type of attitudinal speech. *Desu/masu* form was chosen as it is a common polite form familiar to L2 learners at all levels of proficiency. In previous studies, only one sentence type was used, which was mostly a question [5] [10]. Whether the sentence is a question or a statement could influence pitch height and pitch range. Thus, both types of sentences were included. From the

attitudinal speech categories listed in Japanese textbooks, Request, Inquiry and Decline were selected, as they could present a different level of politeness by prosody more clearly than other functions.

3.1. Materials

Morphologically and lexically identical sentences, which are crucial for acoustic comparison, were used for polite and non-polite expressions. The target sentences followed the phrases which created either a polite setting or a non-polite setting. In the table below, the clear column, (a), shows polite versions and the shadowed column, (b), has non-polite versions.

Table 1. *Stimulus sentences*

	First phrase to extract the settings	Target sentences
1	a. <i>Kachoo</i> (Boss),	<i>Konoken ni tsuite, iken o onegai dekimasuka</i> (Can I ask your opinion about this matter?)
	b. <i>Dareka</i> (Someone),	
2	a. <i>Ie, mada Download shiteimasemode</i> (No, I have not downloaded it yet).	<i>Sonomamani shiteoite kudasai</i> (Please leave it as it is.)
	b. <i>A, hakuban wa kesanaide</i> (Don't wipe the whiteboard.)	
3	a. <i>Ie, jitsuwa.</i> (Well, actually)	<i>Shiryoo ga mitukaranai ndesuga, dokode sagashitara iideshooka.</i> (I cannot find the material, where can I find it?)
	b. <i>Kono site wa totemo wakarunikui desune.</i> (This site is hard to understand.)	
4	a. <i>Soodesune</i>	<i>Chotto yoteio mite mimasuga, muzukashiito omoimasu</i> (I will check my schedule, but think it is difficult)
	b. <i>Soodesune</i> Let me see.	
5	a. <i>Mooshiwakearimasen</i>	<i>Mata, jikaini onegai dekimasuka</i> (Could you please try again next time?)
	b. <i>Mooshiwakearimasen</i> (I am sorry.)	
6	a. <i>Shusseki shitakatta nodesuga, zannendesu.</i> (Shame, I cannot attend.)	<i>Tsuginokikaini shussekisasetekudasai.</i> (Please let me attend next time.)
	b. <i>Sumimasen</i> (I am sorry.)	

3.2. Participants

For a small payment, 10 native (5M, 5F) and 10 non-native (6M, 4F) speakers participated in the recording. Native speakers who speak the standard Tokyo dialect were recruited. Their age ranged from 29 to 63. All of them had spent a substantial period of time as office workers at the time of recording. Non-native speakers were students who had been studying in Japan to complete a one year exchange program. Their level of Japanese was intermediate and they were aware of the existence of Japanese polite speech. The utterances in the laboratory recording heavily rely on the participants' role-playing ability. However, in order to extract the same sentence pattern between polite and non-polite versions, we had to rely on a controlled setting, and prepared a scenario which required a polite or non-polite utterance. More than 10 participants were recruited in both groups. The recordings of the 10 best speakers from both groups whose speech clearly exhibited difference between polite and non-polite versions were chosen and used as stimuli.

3.3. Procedure

3.3.1. Methods of recording

The script was presented on the computer screen with the description of the situation and appropriate photos. The researcher operated the computer screen and reminded the speaker of the context of each utterance before s/he produced the sentence. The participants were asked to read out the stimuli sentences twice in the context described on the screen, imagining themselves in the described situation. The participants were told that the first half (1a~6a) of the list required them to be polite towards their senior, while the second half (1b~6b) required them to act as a person in a higher social position so that they could be firm and blunt.

3.3.2. Perception test

10 native listeners (5M, 5F), recruited separately from speakers of the recording, but in a similar age range participated in the perception test. The stimuli were presented to the subjects as an online perception task which was created using Praat. The subjects were asked to score the politeness of the stimulus utterance on a 5-point Likert scale as follows: <5. Very polite ~1. Not at all polite>. There were three exercise files to let the participants familiarize themselves with the task. They were allowed to listen to the stimuli for up to three times in each trial. 6 sentences were divided into two blocks, Block 1 (Short sentences 1,3,5) and Block 2 (Long sentences 2,4,6), depending on the length of the sentences, so that the length does not affect their judgement. Polite and non-polite speech by both native and non-native speakers were mixed and presented in a randomized order.

4. Results

The politeness score for native and non-native speakers' polite and non-polite speech were calculated below in Figure 1. We evaluated speech to be polite when the score was above 3.5 (dotted line in the graph)

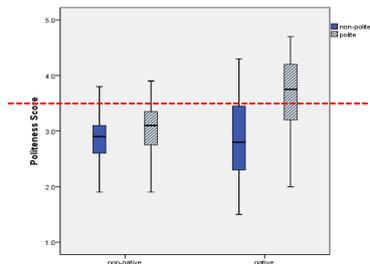


Figure 1: *Politeness scores for native and non-native speakers*

The dark lines in the boxes indicate the mean scores for each category. As found in previous studies, the majority (70% in this study) of polite speech by native speakers was judged as being polite, while only 20% of polite speech by non-native speakers was judged as being polite. We would like to find out what kind of prosodic features contribute to this difference. In the following section, the acoustic measurement for 6 sentences between native and non-native speakers is compared.

4.1. Pitch register

A Z-score was used to normalize the difference between male and female voices. The scores from the 6 sentences were calculated separately. The dotted lines are questions ending

with “ka” and the solid lines are statements and polite imperatives (please do). The left and right ends of the bars in the graph indicate the scores for polite and non-polite speech respectively.

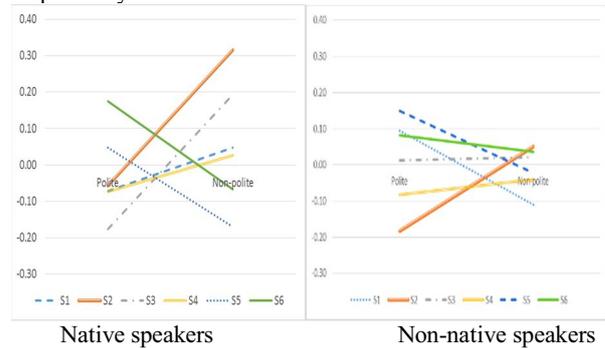


Figure 2: *Average pitch of 10 speakers (Z-score)*

In native speakers' production, a lower pitch was used for polite speech in 4 sentences, and Sentences 5 and 6 show the opposite pattern. In this data, however, statistical significance was not observed. The general idea that polite speech has a high pitch was not supported by this group of native speakers. No particular pattern was found in non-native speakers' production.

4.2. Pitch range

In the graph depicting the native speakers, the polite versions have a narrower pitch range than the non-polite versions in most sentences. The difference between two versions was statistically significant ($t(59)=5.497, p=0.00$). On the other hand, non-native speakers did not show any consistent patterns.

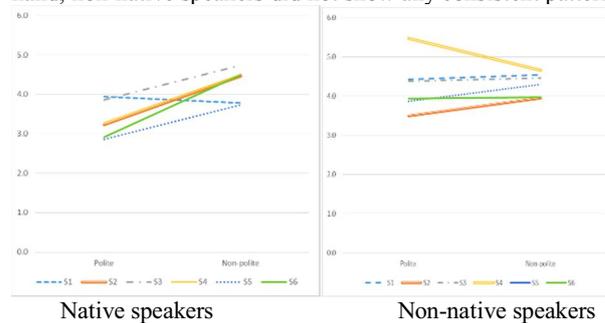


Figure 3: *Pitch range (Z-score)*

4.3. Speech rate

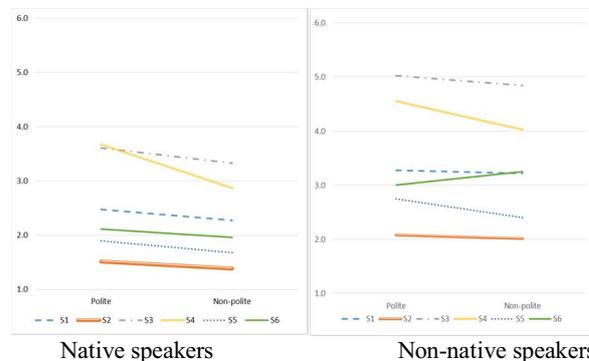


Figure 4: *Duration of speech (including pauses)*

It has been reported that listeners perceive the speech rate as an overall impression of a sentence including pauses. The duration of the speech instead of the speech rate was used as

an index to measure the perception of slow speech. We can see that slower speech is perceived as being more polite than faster speech ($t(59) = -6.067, p=0.00$). This is the same result as found in a study of Korean polite speech by native speakers [7]. It would be interesting to find out whether this comes from cultural similarities between the two languages.

4.4. F0 direction and duration of the final syllable of the sentence

The acoustic measurement in sentence final prosody was looked at in comparison with politeness scores, together with the acoustic measurement of the entire utterance.

4.4.1. The pitch movement in the final particle

There were three sentence final endings, “ka”, “sai”, “mas”. In “sai”, “mas”, pitch was mostly lowered. In the case of the Wh-question sentence, Sentence 3, “ka” has a definite pitch fall. “ka” in yes-no questions, Sentences 1 and 5 showed a contrastive difference between native and non-native speech. The pitch pattern, LH was frequently observed in native speakers’ production as shown below;

Table 2: Pitch movement in “ka”

Pitch pattern	Native	Non-native
LH	11	5
H	3	6
L	4	9

The use of LH in “ka” instead of a simple rise must have given a polite impression, as it was observed in many utterances with high scores, which inevitably made the final syllable longer (see Table 3). Native speakers’ ka is mostly of 2 mora length, while non-natives’ “ka” was generally shorter than two mora. Table 3: Duration of final particle “ka” in Sentence 1 (Duration = Number of mora, Score=politeness score)

Natives				Non-natives			
Score	Gen	No. mora	Pitch pattern	Score	Gen	No. mora	Pitch pattern
4.7	F	2.0	H	3.8	M	1.6	H
4.7	F	1.9	LH	3.7	M	1.1	H
4.2	F	2.4	LH	3.4	F	2.5	LH
4.0	M	1.9	LH	3.4	M	2.0	L
3.9	F	2.6	LH	3.3	M	1.5	L
3.9	M	2.4	LH	3.3	M	1.5	H
3.5	F	2.2	H	3.1	F	2.0	H
2.8	M	1.7	L	2.9	F	1.5	L
2.7	M	2.7	LH	2.5	F	1.4	LH
2.2	M	1.8	H	2.5	M	1.8	L

4.4.2. Delay of pitch rise

In several places in native speakers’ polite speech, a delay of pitch rise after initial lowering was observed. This change of pitch pattern was followed by LH in “ka” and seems to contribute to a narrowing pitch range in polite speech.

e.g. Sentences 1/5 (correct pattern)

Onegai dekimasu ka (LHHH HHHL H?)

→ LLHH LLHL LH?

The correlation coefficient with politeness scores for all acoustic measurements was also calculated. In native speakers’ production, duration of utterance had the highest correlation ($r=0.557$), followed by duration of sentence final syllable ($r=0.346$) and pitch range ($r=0.232$). In the non-native speech, only duration of utterance showed a correlation with politeness, scoring ($r=0.212$). This means that polite

speech was longer, but pitch range was not effectively changed to express politeness in non-native production

5. Conclusions

Prosodic characteristics of Japanese polite speech found in this study are summarized in the following five points; (1) Polite speech does not always have high pitch. Although there was no statistically significant difference, a lower pitch register than non-polite speech was often observed in native speech. (2) Pitch range was narrowed in most polite speech, as found in Korean polite speech. Japanese polite speech could be prosodically very close to Korean polite speech. (3) Slower speech was perceived as being polite. (4) LH Pitch rise at the end of yes-no question sentences was often observed. (5) Delay of pitch rise was observed in the second mora of the word with flat pitch pattern. Non-native speakers only had the characteristic (3) in their polite speech, which means that the other characteristics will be the points for them to pay attention to. The lack of these characteristics is probably the reason for their politeness scores being lower than those of native speakers. Further study is required to identify prosodic factors which affect the impression of the politeness.

6. Acknowledgement

This study was funded by HAKUHO Foundation fellowship, Japan.

7. References

- [1] Ohala, J.J. (1984). An ethological perspective on common cross-language utilization of F₀ in voice. *Phonetica*, 41, 1-16.
- [2] Nadeu, M., & Prieto, P. (2011). Pitch range, gestural information, and perceived politeness in Catalan. *Pragmatics*, 43(3), 841-854.
- [3] Menzes, C., Erickson, D., & Franks, C. (2010). Comparison between linguistic and affective perception of sad and happy – A cross-linguistic study. *Proc. Speech Prosody*, Chicago, IL.
- [4] Brown, L., Winter, B., Idemaru, K. & Grawunder, S. (2014). Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues, *J.of Pragmatics Vol. 66*, 45-60.
- [5] Ogino, T, and Hong, M. (1992). Nihongo onsei no teineisa ni kansuru kenkyuu (A study on politeness in Japanese speech) In Kunihiro, T. (ed.) The state of the art and analysis of Japanese intonation, Ministry of Education, Tokyo, 215-258.
- [6] Ohara, Y. (2001) Finding one’s voice in Japanese: a study of the pitch levels of L2 users, In A. Pavlenko., A. Brackledge, I. Pilller, M. Teutsch-Dwyer (Eds.), *Multilingualism, Second Language Learning, and Gender*. Mouton de Gruyter, New York, 231-254.
- [7] Grawunder, S., & Winter, B. (2010). Acoustic correlates of politeness: prosodic and voice quality measures in polite and informal speech of Korean and German speakers. *Proc. Speech Prosody*, Chicago, IL.
- [8] Idemaru, K., Winter, B. and Brown, L. (2015) The role of pitch in perceiving politeness in Korean, *ICPhS*.
- [9] Gu, W., Zhang, T., & Fujisaki, H. (2011). Prosodic analysis and perception of Mandarin utterances conveying attitudes, *Proc. INTERSPEECH*, Florence, Italy, 1069-1072.
- [10] Ofuka, E., McKeown, D., Waterman, M., & Roarch, P. (2000). Prosodic cues for rated politeness in Japanese speech. *Speech Communication*, 32, 199-217.
- [11] Kagomiya, T. Yamazumi, Maki, Y. and Maekawa, K. (2008) Factors that affect global perceived speaking rate of spontaneous speech *Journal of the Phonetic Society of Japan*, 12 (1), 54-62.
- [12] Bolinger, D. L. (1978). Intonation across languages, In J.P. Greenberg, C.A. Ferguson, and E.A. Moravcsik (eds.), *Universals of human language*. Vol.2 *Phonology*. Stanford University Press.