http://e-flt.nus.edu.sg/

Electronic Journal of Foreign Language Teaching
2011, Vol. 8, Suppl. 1, pp. 277–290

What Constitutes "Good Pronunciation" from L2 Japanese Learners' and Native Speakers' Perspectives? A Perception Study

Shunichi Ishihara
(Shunichi.Ishihara@anu.edu.au)
Australian National University, Australia

Chiharu Tsurutani
(c.tsurutani@griffith.edu.au)
Griffith University, Australia

Kimiko Tsukada
(kimiko.tsukada@mq.edu.au)
Macquarie University, Australia

Abstract

Native speakers of a language can tell whether a speaker is native or non-native just by hearing one word or phrase in the language. It is expected that L2 learners will develop the ability to detect 'good pronunciation' as they establish the prototypes of the L2 sound system. However, it is not known what contributes to their judgement of good pronunciation. Therefore, this pilot study aims to clarify the mechanism of L2 listeners' judgement of good pronunciation. In this study, we focus on the prosodic variations of timing and pitch-accent in Japanese. Four groups of informants participated in a perception experiment where they were asked to assess stimuli extracted from the recordings of L2 Japanese learners. These four groups are L2 Japanese learners at two levels of proficiency (beginner and advanced), and native Japanese speakers with and without formal teaching experience of L2 Japanese. All learner participants are native speakers of Australian English. We will report that the learners' assessment of good pronunciation is not straightforward, being different from the logical expectation that learners will behave more like native speakers as their L2 Japanese proficiency develops. We will also discuss possible explanations for the results and implications for L2 Japanese education.

1 Introduction

It has been reported that native speakers can detect foreign accents in their own mother tongue (L1) even from a single word or phrase (Flege, 1984). If this is indeed true, what contributes to native speakers' judgment on accented speech? Although there are some conflicting reports, the majority of previous studies across languages provide evidence to suggest that prosodic aspects are generally more important than segmental aspects for speech to be judged as less accented or to have good pronunciation. In other words, speech with prosodic errors is more likely to be perceived as foreign-accented speech than speech with segmental errors. This superiority of

prosody to segment has been confirmed for Japanese (Huckvale, 2006; Sato, 1995). However, 'prosody' is a very broad term which includes the rhythm, stress, pitch, accent, intonation, etc. of speech (basically, anything which is not segmental is classified as prosody). Some researchers argue that the perception of native listeners is more sensitive to temporal distortion (in other words, timing errors) than to other kinds of distortions (Tajima, Port, & Dalby, 1997; Tsurutani, 2010). Tsurutani (2010), who investigated the perception of 80 Japanese university students who reside in Japan and have not had regular contact with foreigners, found out that accuracy in timing is more important than in pitch-accent for L2 Japanese speech to be perceived as native-like speech.

"Speaking like a native speaker" is a goal of many students who are learning a foreign language. The achievement of this goal, which is very difficult in many cases, requires mastery of the various linguistic aspects of the target language. Through exposure to and practise of the target L2, the L2 learners' phonology is expected to approximate to (but may not exactly match) the sound system of the target L2, and the ability to detect the correctness of pronunciation is expected to develop as they establish prototypes for phones and prosody in the L2. Although there are a large number of studies focusing on native speakers' judgement on accented speech (e.g. Bond et al., 2003; Munro, 1995; Munro & Derwing, 1998; Tajima et al., 1997), to date, studies on non-native listeners' perception of L2 speech are scarce and heavily dominated by studies with English as the target language (cf. Holm, 2008). Consequently, we do not know how learners of Japanese, whose L1 phonological system is different from the Japanese phonological system, perceive accented Japanese, or how their perception changes as a function of their level of Japanese. Answering these questions will contribute to a better theoretical understanding of the mechanism of L2 perception. Answers to these questions will also provide pedagogically useful insight into learning and teaching Japanese, particularly because not all language teachers are native speakers of the target language. In the case of Japanese, the Japan Foundation (2009) reported that 70% of Japanese language teachers overseas are non-native speakers. Thus, it is pedagogically meaningful to understand non-native speakers' perception of foreign-accented speech in comparison to native speakers' perception, as any differences in judgment between native and non-native teachers may result in inconsistent or unfair assessment.

Thus, the current study poses the following research questions:

1. Can L2 learners acquire criteria similar to those of native listeners or is their judgement harsher or more lenient towards accented speech?
2. Is advanced learners' judgment closer to native listeners' than to beginners'?

The current study is an extension of Tsurutani (2010) in which the perception of naïve native Japanese listeners (those Japanese native speakers who are not Japanese language teachers or have never been trained as Japanese language teachers) was tested using L2 Japanese stimuli. In the current study, exactly the same stimuli as Tsurutani's were used for two groups (beginner and advanced) of English-speaking L2 Japanese learners, and the results are compared with those of the naïve native Japanese listeners given in Tsurutani (2010). The main finding of Tsurutani (2010) is that naïve native Japanese listeners put more weight on accuracy in timing than in pitch to assess the L2 Japanese speech. In other words, accuracy in timing is more important than accuracy in pitch for naïve native Japanese listeners. A more detailed summary of Tsurutani's study will be given in Section 3 immediately before the results of the current study are provided so that the readership can better understand the similarities and differences between the results of Tsurutani and those of the current study. Please note that all learner participants are native speakers of Australian English.

Besides the above research questions, which primarily focus on naïve native vs. non-native listeners, it is important to investigate if there are any differences in perception between naïve native Japanese listeners and native Japanese language teachers, as it has been reported that these two groups use different criteria for assessment (Nohara, 2008; Okamura, 1995). Thus, the following research question is also pursued:

1. Is the perception of native Japanese language teachers different from that of naïve native Japanese?

1.1 Prosodic nature of Japanese and English

It is sensible to expect that phonological differences between Japanese and English will affect the perception of foreign-accented Japanese speech by native speakers of English. In this section, the two prosodic features which the current study is most concerned with, and in which Japanese and English significantly differ – namely rhythm (or timing) and pitch-accent – will be explained.

Japanese and English are typologically different rhythmically (Tsujimura, 1996). English is a so-called stress-timed language, that is, stressed syllables appear at a roughly constant rate, and Japanese is a mora-timed language in which each mora is spoken at a roughly constant rate. Mora is a sub-syllabic unit mainly formed by an optional consonant and a vowel. Furthermore, Japanese has a phonemic contrast in short/long vowels and consonants. Thus, a minimum pair of, for example, /se.ki/ "seat" vs. /se.e.ki/ "century" (both of which are two-syllable words, but the former is a two-mora and the latter is a three-mora word, where a period stands for a moraic boundary) requires special attention by English speakers, who tend to pronounce these words in the same rhythmic pattern as both of them are two-syllable words.

Another difference between Japanese and English is that Japanese is a so-called pitch-accent language as opposed to English, which is a stress-accent language (Beckman, 1986). In Japanese, each mora in a morpheme/word is associated with a specific pitch. Thus, segmentally identical words may have different meanings with different pitch shapes (e.g. /i.ko.o/ LHH "intention", HLL "after", LHL "let's go").

As can be predicted from these prosodic characteristics of Japanese, incorrect pitch pattern and/or the duration of vowels/consonants have the potential of becoming a source of miscommunication or a cue for foreign-accented speech in Japanese.

2 Experiment

In this section, the detailed methodology employed for the current study will be explained in the order of the materials (Section 2.1), participants (Section 2.2) and the procedure of the experiments (Section 2.3). As mentioned earlier, the same stimuli as those used in Tsurutani (2010) are used in the current study.

2.1 Materials

The stimuli were extracted from recordings of Australian English speakers who had studied Japanese for approximately 160 hours at university at the time of recording. The recordings were taken from a computer exercise that was developed as a self-assessment of pronunciation, in which students were required to utter various kinds of sentences. These recordings were compiled as a speech database. Utterances that contain timing errors (e.g. errors in long/short contrast in vowels and consonants) or pitch errors (errors in pitch contour) or both, with no obvious segmental errors, were chosen, together with utterances without any obvious errors, from this large speech database. The judgment of errors was made by the second author of this paper and two other native speakers of Japanese who had been teaching Japanese for many years. When two or more out of the three agreed, the judgement was accepted. The four patterns given in Table 1 were considered.

1	Correct pitch, correct timing	PcTc
2	Incorrect pitch, correct timing	PiTc
3	Correct pitch, incorrect timing	PcTi
4	Incorrect pitch, incorrect timing	PiTi

Table 1: Four stimulus types

Using the same sentence as material would be ideal for analysis, yet it could wear listeners' concentration. Because of this, the stimuli were taken from the six sentences given in Table 2.

Note that the sentences given in Table 2 contain many words in which long vowels/consonants appear. Obviously, these long vowels/consonants need to be articulated with enough duration in contrast to corresponding short vowels/consonants. For each sentence given in Table 2, an utterance which fits each of the four patterns (see Table 1) was extracted from the database, and compiled as 24 stimuli (= 6 sentences x 4 error patterns). It is important to point out here that since all stimuli are natural utterances, they are not perfectly controlled in a mutually comparable manner, yet every effort was made to select utterances which are as comparable as possible in terms of characteristics such as speech rate and number of errors. We emphasize again that the stimuli do not contain any obvious segmental errors.

1	*shachoo-no kekkonshiki-ni okyakusan-ga sennin kita.* 社長の結婚式にお客さんが千人来た。 1000 people attended the president's wedding reception.
2	*tsugi-no jugyoo-no suugaku-wa chotto muzukashii desu.* 次の授業の数学はちょっと難しいです。 Mathematics in the next class is a bit hard.
3	*watashi-no kookoo-de isshoni shashin-o torimashoo.* 私の高校で一緒に写真をとりましょう。 Let's take a photo together at my high school.
4	*otooto-no okusan-wa ryokoo-ni ikuno-ga suki desu yo.* 弟の奥さんは旅行に行くのが好きですよ。 My younger brother's wife likes travelling.
5	*tanjoobi-ni tomodachi-kara kireena hana-o moratta.* 誕生日に友達から綺麗な花をもらった。 I received beautiful flowers from my friend on my birthday.
6	*shuumatsu-kara futarino hito-to shigoto-o suru yotee desu.* 週末から二人の人と仕事をする予定です。 Starting from this weekend I'm planning to work with 2 people.

Table 2: Sentences used as stimuli

The task of the informants is to assess the pronunciation of the independent sentences given in Table 2. In real conversation, however, we utter a sentence within a specific context. Thus, even if a sentence is syntactically correct and is delivered with perfect pronunciation, it may sound odd by itself depending on the content of the sentence without any specific context. We could have provided context for the sentences, but it would add extra complexity to the experiment and may distract the focus of the informants away from the target sentences. Instead, in order to avoid the unnaturalness arising from the independent sentences having no specific context, the target sentence was presented in writing to the informant each time before they actually listen to it (refer to Appendix).

These stimuli were presented in 6 blocks (one per sentence) within which the order of error patterns given in Table 1 was randomized. None of the utterances had pauses longer than 300 ms and all sounded reasonably fluent.

2.2 Participants

Three different groups participated in listening experiments. These groups are: native speakers of Japanese who are professional Japanese language teachers (LTNS); native speakers of English whose level of Japanese was judged as beginners level (BNNS); and native speakers of English whose level of Japanese was judged as advanced level (ANNS). The LTNS group consists of 11

teachers (1 male and 10 females). The BNNS group consists of 21 students (7 males and 14 females) and the ANNS group 12 students (5 males and 7 females). The teachers belonging to the LTNS group had been teaching Japanese at a tertiary institution for 2 to 20 years. The students belonging to the BNNS group had been studying Japanese for about 170 hours at university. All of the students belonging to the ANNS group had passed Level 1 of the Japanese Language Proficiency Test or were judged by the first author to possess an equivalent level of proficiency or higher.

Including the participants of Tsurutani (2010) who were naïve native speakers of Japanese who were not language teachers and had never been trained as a language teacher (NLTNS), this study concerns the perception of L2 Japanese speech with the groups of NLTNS, LTNS, BNNS and ANNS as factors. The attributes of the four groups of participants are summarised in Table 3

Native Speakers of Japanese	Naïve Japanese; non teachers	NLTNS
	Japanese language teacher	LTNS
Native speakers of Australian English learning Japanese	Beginners level	BNNS
	Advanced level	ANNS

Table 3: The attributes of the four participating groups concerned

2.3 Procedure

The 24 stimuli extracted from the large L2 Japanese database (see Section 2.1) were played to the three groups of listeners explained in Section 2.2 for their judgements on pronunciation goodness. These stimuli were played in 6 blocks for ease of comparison, as shown in Table 2. Each stimulus was played twice with a 4 second interval, and an inter block interval of 8 seconds. Before the task, 3 practice sentences were played for listeners to become accustomed to the task and the proficiency level of the L2 speakers. The participants were asked to rate the naturalness of utterances on a Likert scale ranging from 1 (not at all native like) to 7 (native like) (refer to Appendix). The listening task took about 15 minutes, including the time for instructions.

3 Findings of Tsurutani (2010)

The findings of Tsurutani (2010) will be summarised in this section. Readers should be reminded that Tsurutani's study concerns only the perception of naïve native speakers of Japanese (NLTNS). The bar plots given in Figure 1 show the mean values of the scores pooled separately for the four different stimulus types for the NLTNS group. In Figure 1, one standard deviation is also given around the mean. Although the mean difference (0.39) between the PiTc (4.69) and PcTi (4.30) types is relatively small, the relationship of PcTc > PiTc > PcTi > PiTi was statistically confirmed for the NLTNS group. It is not surprising that the PcTc received the highest average score while the PiTi the lowest. The main finding of Tsurutani (2010) is that there is a statistically significant difference between PiTc and PcTi (PiTc > PcTi). This indicates that the NLTNS group (or naïve native speakers of Japanese) puts more weight on accuracy in timing than in pitch to assess the L2 Japanese speech.

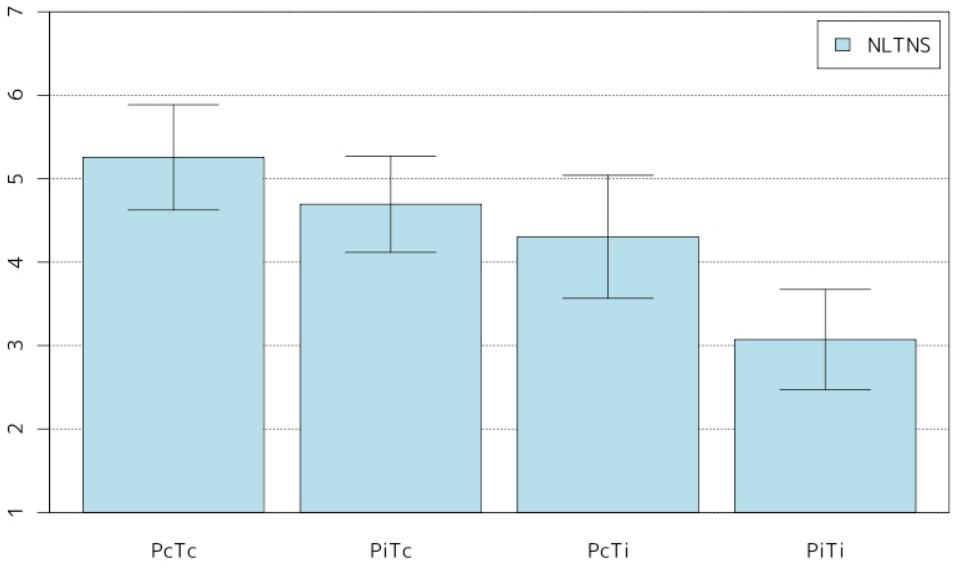

Fig. 1: Barplots showing the average scores of the PcTc, PiTc, PcTi and PiTi types for the NLTNS group. One standard deviation is plotted around the mean (Tsurutani, 2010).

Based on the result, Tsurutani argues the superiority of timing to pitch for the judgement of L2 Japanese speech. She also points out that the inferiority of pitch can be accounted for in light of the variations of pitch patterns observed across Japanese dialects and generations. That is, since there are various patterns observed even across native speakers of Japanese depending on their dialects and generations, Japanese native speakers are more tolerant of pitch errors than timing errors.

4 Results

In this section, the results of the current study are given. In Section 4.1, the experimental result of the LTNS group will be presented, particularly in comparison to the NLTNS group. In Section 4.2, the results of the BNNS and ANNS groups will be given. In Section 4.3, the overall results of the four groups will be compared, focusing on the four types of stimuli (PcTc, PiTc, PcTi and PiTi).

4.1 Native Japanese listeners: NLTNS vs. LTNS

In this subsection, the results of the LTNS group will be given in comparison to that of the NLTNS given in Section 3. Following Figure 1, the mean values of the scores calculated separately for the four different stimulus types are plotted in Figure 2 for the LTNS together with the results of the NLTNS group. The scores given by the LTNS group were separately submitted to a one way repeated ANOVA with the stimulus types (PcTc, PiTc, PcTi and PiTi) as a factor, followed by Post hoc Tukey HSD tests with a confidence level of 95%.

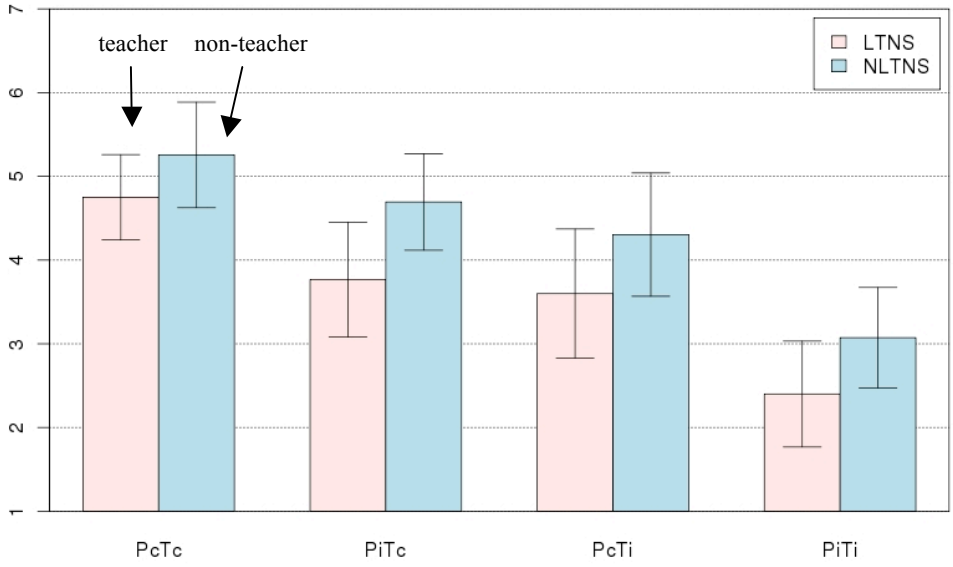

Fig. 2: Barplots showing the average scores of the PcTc, PiTc, PcTi and PiTi types, separately plotted for the LTNS and NLTNS groups. One standard deviation is plotted around the mean.

In perceptions of the PiTc and PcTi types, the difference between the teacher group (LTNS) and the non-teacher group (NLTNS) is evident. The NLTNS group gave a better score to PiTc than PcTi, putting more weight on the correctness in timing than in pitch, whereas the same difference was not observed for the LTNS group (average scores: PiTc = 3.76; PcTi = 3.60) ($p=0.898$).

Another significant difference between the LTNS and NLTNS groups is that a higher score is generally given by the latter group than the former group, regardless of the stimulus types [two way repeated ANOVA, $F(1,2155) = 64.126$, $p<0.001$]. In other words, the native Japanese language teachers (LTNS) assessed the pronunciation more critically than the naïve native speakers of Japanese (NLTNS).

4.2 Non-native Japanese listeners: BNNS vs. ANNS

In this subsection, the results of the non-native groups (BNNS and ANNS) will be presented. The presentation procedure of the results is identical to that of the NLTNS and LTNS groups in Section 3 and Section 4.1, respectively.

The results of the BNNS and ANNS groups are plotted in Figure 3. The scores given by the BNNS group and the ANNS group were statistically analysed separately using a one way repeated ANOVA and the Tukey HSD tests.

For the PiTc, PcTi and PiTi types, both the BNNS and ANNS groups showed the statistically significant relationship of PiTc > PcTi > PiTi [ANNS: $F(3,284) = 20,198$, $p<0.001$; BNNS: $F(3,500) = 56.26$, $p<0.001$]. A two way repeated ANOVA using the stimulus type and the two groups of students (ANNS and BNNS) as factors confirmed the relationship of PiTc > PcTi > PiTi [$F(3,787) = 74.198$, $p<0.001$]. That is, like the NLTNS group, the PiTc type was given better scores than the PcTi type by both of the non-native groups. This indicates that the non-native groups also put more weight on accuracy in timing than in pitch for the perception of L2 accented Japanese.

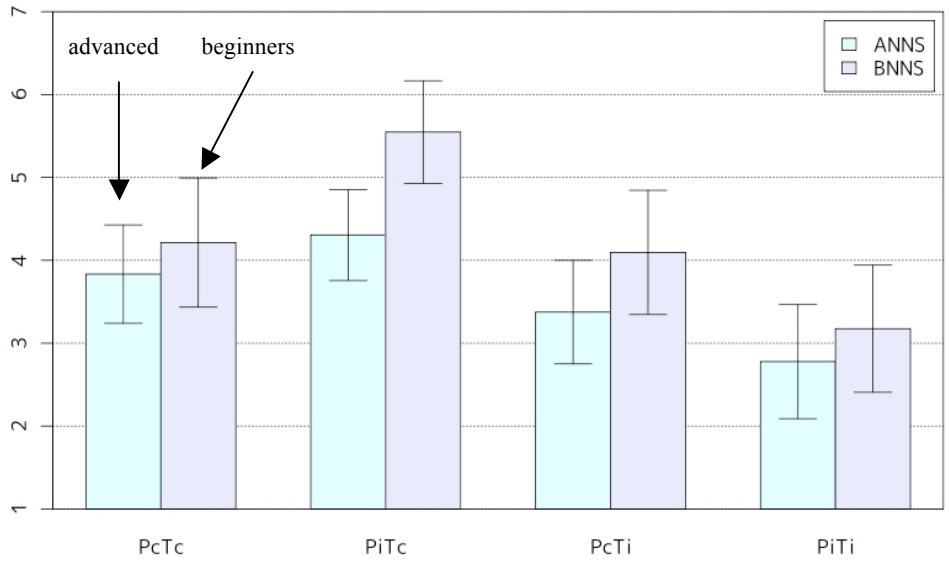

Fig. 3: Barplots showing the average scores of the PcTc, PiTc, PcTi and PiTi types, separately plotted for the ANNS and BNNS groups. One standard deviation is plotted around the mean.

An interesting observation which can be made from Figure 3 is that in comparison to the native speaker groups, both of the non-native groups very harshly assessed the PcTc type, which should get the highest average score. Thus, no significant difference was observed between PcTc and PiTc for BNNS [$p=0.100$], and the PiTc type (5.54) was given a significantly higher score than the PcTc type (4.21) by BNNS [$p<0.001$].

A significant difference between the ANNS and BNNS groups is also evident in that the ANNS group gave significantly lower scores for the stimuli than the BNNS group [two way repeated ANOVA, $F(3,787) = 46.459$, $p<0.001$], which is very similar to the relationship between the LTNS and NLTNS groups. That is, the judgement of the ANNS group is generally more critical than the BNNS group.

4.3 Overall comparisons

In this subsection, the scores given by the four groups in question will be scrutinized mainly according to the stimulus types. The average scores of the four stimulus types are plotted all together, but separately for the four groups of participants, in Figure 4. Thus, Figure 4 is essentially the combination of Figures 2 and 3.

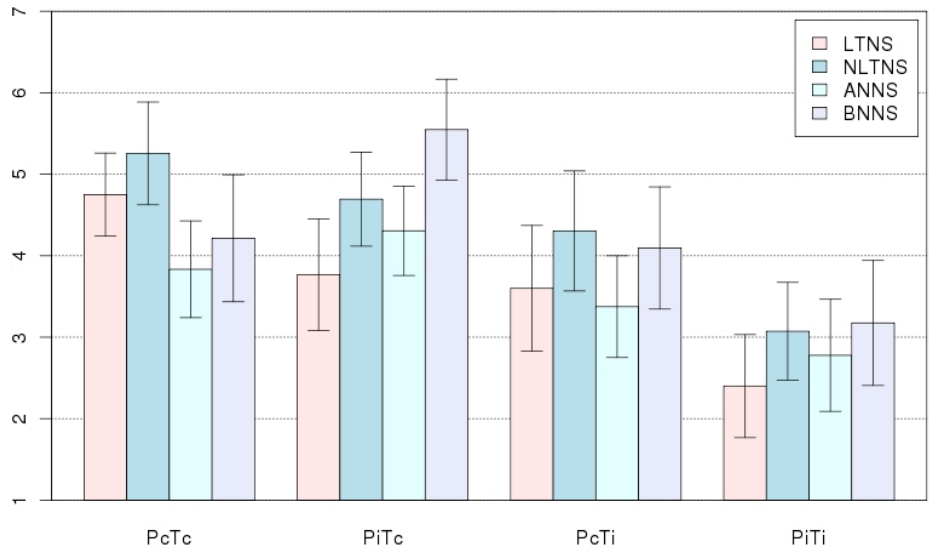

Fig. 4: Barplots showing the average scores of the PcTc, PiTc, PcTi and PiTi types, separately plotted for all of the four groups. One standard deviation is plotted around the mean.

By referring to Figure 4, discussions will progress in the order of PcTc (Section 4.3.1), PiTc (Section 4.3.2), PcTi (Section 4.3.3) and PiTi (Section 4.3.4).

4.3.1 Pitch: correct; Timing: correct

It was pointed out in Section 4.2 that the PcTc type was far more harshly marked by the non-native groups (BNNS and ANNS) compared to the native groups (NLTNS and LTNS). This difference can be clearly seen in Figure 4. It was statistically proven that the scores given by the non-native groups (average scores: BNNS = 4.21; ANNS = 3.83) were significantly lower than those given by the native groups (average scores: LTNS = 4.75; NLTNS = 5.25) for the PcTc type [one-way repeated ANOVA, $F(3,734) = 40.6$, $p<0.001$]. The PcTc type which should be assessed with the highest score out of the four stimulus types was not given the highest score by the non-native groups.

4.3.2 Pitch: incorrect; Timing: correct

A significant difference between the four groups has been confirmed (LTNS < ANNS < NLTNS < BNNS) [one-way repeated ANOVA, $F(3,734) = 36.70$, $p<0.001$, Tukey HSD, $p<0.454$]. A larger difference in scores between LTNS and BNNS shows that the BNNS group (5.54) assessed the PiTc type far more leniently than the LTNS group (3.76). That is, the non-native speaking beginners gave a significantly higher score for the PiTc type than the native teachers.

4.3.3 Pitch: correct; Timing: incorrect

As far as the average scores are concerned, it can be seen from Figure 4 that the NLTNS group (average score: 4.30) and the BNNS group (average score: 4.09) show very similar average scores for the PcTi type, and the LTNS group (average score: 3.60) and the ANNS group (average score: 3.37) marked the same type, i.e. PcTi, in a very similar way, but the latter groups (LTNS and ANNS) are more rigorous than the former groups (NLTNS and BNNS). Yet, statistically speaking only the following relationships were significantly confirmed: NLTNS > LTNS; NLTNS > ANNS;

BNNS > ANNS [p<0.002]. In other words, 1) the naïve native speakers scored the PcTi type significantly higher than the native teachers and the advanced learners, and 2) the beginner learners marked the same type significantly higher than the advanced learners.

4.3.4 Pitch: incorrect; Timing: incorrect

The difference in scores between the four groups is smallest for the PiTi type (average scores: NLTNS = 3.07; LTNS = 2.40; ANNS = 2.77; BNNS = 3.17). Thus, the only significant difference identified between the four groups is NLTNS > LTNS and BNNS > LTNS [p<0.001]. That means that the native teachers were more critical of the PiTi type than the naïve native speakers and the non native speaking beginners.

5 Summary of the results

The following is a summary of the results given in Section 4.

1. Except for LTNS, correctness in timing was weighted more than correctness in pitch;
2. The LTNS group assessed the L2 Japanese pronunciation more rigorously than the NLTNS group;
3. The ANNS group assessed the L2 Japanese pronunciation more rigorously than the BNNS group;
4. The non-native listeners (BNNS and ANNS) marked the correct stimulus type (PcTc) more harshly than the native listeners (NLTNS and LTNS); and
5. The BNNS group marked the sentences with incorrect pitch (PiTc) very leniently.

In the following section, each of the above findings will be discussed. Whenever necessary and relevant, implications for Japanese language teaching/learning will be pointed out as well.

6 Discussion

6.1 Except by LTNS, correctness in timing was weighted more than correctness in pitch

It is interesting to observe that not only the ANNS group but also the BNNS group employ similar perceptual cues to the NLTNS group for the judgement of L2 Japanese speech. There is a considerable amount of research in which weighting differences amongst prosodic features have been compared for foreign-accented speech (Bannert, 1995; Almberg & Husby, 2000; Kamiyama, 2004; Trofimovich & Baker, 2006; Gut, 2007). Although no conclusion has yet been reached, the investigations with a large number of subjects suggest that durational aspects influence the degree of foreign accent more than intonational aspects (Trofimovich & Baker, 2006; Gut, 2007). However, Holm (2008) argues with her own experimental results that the role of intonation vs. duration may vary according to language pairing.

The superiority of timing to pitch for the BNNS and ANNS groups may be specific to speakers of Australian English, in which duration is known to play a more important role than in other varieties of English (Harrington & Cassidy, 1994; Cox, 2006). As far as the speakers of Australian English are concerned, regardless of the level of their Japanese, they put more weight on timing than pitch like the NLTNS group.

Unlike the other groups, the result of the LTNS group indicates that Japanese native language teachers treated the timing and pitch errors equally. It has been reported that teachers and non-teachers utilise different criteria for assessing the performance of L2 learners (Hadden, 1991; Nohara, 2008; Okamura, 1995). Non-teachers tend to assess L2 speakers' utterances holistically, relying on their subjective impressions, whereas teachers' assessments are more based on detailed sets of linguistic criteria. Furthermore, based on training and experience, language teachers are well aware of and can expect the sorts of errors which native speakers of English are prone to. Hence, language teachers can (or at least are trained to) identify specific errors, and consequently provide detailed feedback. It can be logically expected, therefore, that this characteristic of lan-

guage teachers would lead to a quantitative/objective assessment rather than a qualitative/subjective assessment. Thus, it can be presumed that due to the characteristics of language teachers, one error is treated by the LTNS group as one error, regardless of the type of error.

6.2 The LTNS group assessed the L2 Japanese pronunciation more rigorously than the NLTNS group

Many studies acknowledged that teachers tend to assess learners' L2 performance more critically than non-teachers (Hadden, 1991; Nohara, 2008; Okamura, 1995). The results of our study conform to those of the previous studies.

The current study revealed that there were also differences between language teachers and non-teachers in the perception of the prosodic aspects (timing and pitch) of L2 Japanese speech. Language teachers need to be aware of these differences in perception of learners' speech between teachers and non-teachers, as learners eventually need to start interacting with naïve native speakers in the real world.

6.3 The ANNS group assessed L2 Japanese pronunciation more rigorously than the BNNS group

Fayer et al. (1987) reported that non-native speakers of a target language assess others more critically than native speakers. The current study showed that the advanced learners (ANNS) were more critical than the beginner learners (BNNS). In fact, a two-way repeated ANOVA with the Tukey HSD tests shows that there are no significant differences between the BNNS and NLTNS groups and between the ANNS and LTNS groups [$p>0.769$]. That is, the perception of the ANNS group is as critical as that of the LTNS group. This result shows that the relationship between the learners' perceptions and their proficiency levels is not straightforward, being different from the logical expectation that learners will behave more like naïve native speakers as their L2 Japanese proficiency develops. When their proficiency level increases, learners seem to have a more critical ear for incorrect speech, which presumably helps them improve their L2 speech. Thus, the judgement of the learners becomes more like that of teachers as the level of proficiency improves. Inversely speaking, it indicates that the judgement of naïve native speakers is very tolerant to foreign-accented speech.

6.4 The non-native listeners (BNNS and ANNS) marked the correct stimulus type (PcTc) more harshly than the native listeners (NLTNS and LTNS)

It is very difficult to find a sensible reason for this from possible perceptual differences between the non-native and native groups. An explanation we can moot is the actual stimuli we used for the PcTc type. All PcTc type stimuli were read by male learners, and thus their f0 ranges (their average range = 75.2 Hz) were narrower than the average f0 range of the other utterances (= 120.0 Hz). In fact, they admittedly sound very monotonous. This monotony in pitch may have adversely affected the perception of the non-native listeners. If this paralinguistic characteristic of the PcTc stimuli resulted in the difference between the non-native and native groups, it can be said that learners' perception, regardless of the level of their proficiency, is subject to the influence arising from paralinguistic differences of speech, causing misjudgement of good pronunciation. The point that specifically caught our attention is the fact that even the advanced students (whose level of proficiency is as high as Level 1 of the Japanese Language Proficiency Test or equivalent) underscored the PcTc type. This implies that although they are trained as teachers, non-native teachers of Japanese are prone to inaccurately assessing learners' pronunciation (particularly good pronunciation). This is of grave concern from a pedagogical point of view.

Although paralinguistic factors may not be the sole cause of misjudgement by L2 listeners, it is indeed the case that the model pronunciation L2 learners listen to in their classroom quite often sounds confident, perky and cheerful. It would be beneficial for them to listen to sample speech by

many different types of speakers and teachers in order to help them train their perception and familiarise themselves to a wide variety of speech samples.

6.5 The BNNS group marked the sentences with incorrect pitch (PiTc) very leniently

This result implies that the BNNS group put far less importance on correct pitch for good pronunciation than the other groups. In other words, they neglect correct pitch possibly because their mastery of pitch-accent has not reached the same level as the ANNS group. This suggests that more practice on pitch-accent may be necessary in the early stages of learning Japanese.

7 Conclusions and future studies

The following are the research questions we asked at the beginning of this study:

1. Can L2 learners acquire criteria similar to those of native listeners or is their judgment harsher or more lenient towards accented speech?
2. Is advanced learners' judgement closer to native listeners' than to beginners'?
3. Is the perception of naïve native Japanese different from that of native Japanese language teachers?

We found the following through the perception experiments:

1. Like naïve native listeners, non-native Japanese listeners were more sensitive to timing information than pitch information;
2. Advanced learners were more critical of L2 performance than beginners;
3. Advanced learners' judgement is closer to native teachers than to naïve native listeners;
4. Beginners are less sensitive to pitch errors than advanced learners;
5. Unlike naïve native listeners, native teachers are equally sensitive to pitch and timing information; and
6. Teachers are more critical than naïve native listeners.

In this study, we examined the perception of Japanese learners whose L1 is Australian English. We observed that, despite their incomplete mastery of L2 phonology, even beginners put more weight on timing than pitch information in order to make perceptual judgements. This may mean that timing information is generally more salient, accessible and, hence, easier to acquire than pitch information in Japanese. In order to gain a better understanding of learners' access to timing vs. pitch information in Japanese, it would be necessary to examine the perception of Japanese learners from different L1 backgrounds.

The findings of the present study provide useful information for L2 teaching and speech science, particularly for characterising and synthesising L2 speech. Since the stimuli were not controlled in all linguistic and non-linguistic aspects, there is a possibility that some extra-linguistic aspects of the stimuli, such as voice quality, speech rate and fluency, may have influenced the judgement of the listeners. Speech synthesis will enable more rigorous selection of the stimuli.

References

Almberg, J., & Husby, O. (2000). The relevance of some acoustic parameters for the perception of a foreign accent. In A. James & J. Leather (Eds.), *New sounds 2000* (pp. 1–10). University of Klagenfurt.

Bannert, R. (1995). Intelligibility and acceptability in foreign accented Swedish: the effects of rhythmical and tonal features. *PHONUM*, *3*, 7–29.

Beckman, M.E. (1986). *Stress and non-stress accent.* Foris Publications: Dordrecht.

Bond, Z.S., Stockmal,, V., & Markus, D. (2003). Sentence durations and accentedness judgments. *Journal of the Acoustical Society of America*, *113*, 2330–2334.

Cox, F. (2006). The acoustic characteristics of /hVd/ vowels in the speech of some Australian teenagers. *Australian Journal of Linguistics*, *26*, 147–179.

Fayer, J.M., & Krasinski, E.K. (1987). Native and non-native judgments of intelligibility and irritation. *Language*, *37*, 311–326.

Flege, J.E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, *76*, 692–707.
Gut, U. (2007). Foreign accent. In C. Müller, (Ed.), *Speaker classification* (pp. 75–87). Heidelberg: Springer.
Hadden, B.L. (1991). Teacher and nonteacher perceptions of second-language communication. *Language Learning*, *41*, 1–24.
Harrington, J., & Cassidy, S. (1994). Dynamic and target theories of vowel classification: Evidence from monophthongs and diphthongs in Australian English. *Language and Speech*, *37*, 357–373.
Holm, S. (2008). *Intonational and durational contributions to the perception of foreign-accented Norwegian: an experimental phonetic investigation* (PhD. dissertation). Norwegian University of Science and Technology, Norway.
Huckvale, M. (2006). The new accent technologies: recognition, measurement and manipulation of accented speech. In P. Zhang, T.-W. Xie, S. Lin, J.-H. Xie, A.C. Fang & J. Xu (Eds.), *Research and application of digitized Chinese teaching and learning*. (pp. 28–37). Beijing: Language and Culture Press.
Japan Foundation. (2009). *Survey report on Japanese-language education abroad 2006: present condition of overseas Japanese-language education (summary)*. Retrieved from http://www.jpf.go.jp/e/japanese/survey/result/img/2006all.pdf.
Kamiyama, T. (2004). Perception of foreign accentedness in L2 prosody and segments: L2 Japanese speakers learning L2 French. *Speech Prosody 2004*, 721–724.
Munro, M.J. (1995). Nonsegmental factors in foreign accent. *Studies in Second Language Acquisition*, *17*(1), 17–34.
Munro, M.J., & Derwing, T.M. (1998). The effects of speaking rate on listener evaluations of native and foreign accented speech. *Language Learning*, *48*(2), 159–182.
Nohara, Y. (2008). Factors that determine explicitness of speech. *Ochanomizu University Journal of the Graduate School of Humanities & Sciences*, *11*, 165–174.
Okamura, A. (1995). Teachers’ and nonteachers’ perception of elementary learners’ spoken Japanese. *The Modern Language Journal*, *79*, 29–40.
Sato, T. (1995). A comparison of phonemes and prosody in the evaluation of spoken Japanese. *Japanese Language Education around the Globe*, *5*, 139–154.
Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English, *Journal of Phonetics*, *25*, 1–24.
Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, *28*(1), 1–30.
Tsujimura, N. (1996). *An introduction to Japanese linguistics*. Blackwell Publishers.
Tsurutani, C. (2010). Foreign accent matters when timing is wrong. In *Proceedings of Interspeech 2010*, 1854–1857.

Appendix

Name Age M / F
I have been studying Japanese for (1.5 /2 / 3 /4 / /years).
Have you been to Japan? Yes / No How long? ( )
My mother tongue is ( )

<<Task>>
You will hear the following Japanese sentences. Judge the level of speakers by circling the number on the scale.

しゃちょう けっこんしき きゃく せんにんき
1) 社 長 の 結 婚 式 に、お 客 さんが 千 人 来 た。
(One thousand people attended the president’s wedding reception.)
つぎ じゅぎょう すうがく むずか
2) □ □ □ □ □ □ □ □□□□□ □ □□□□□
(Mathematics in the next class is a bit hard.)
わたし こうこう いっしょ しゃしん
3) 私 の 高 校 で 一 緒 に 写 真 をとりましょう。
(Let’s take a photo together at my high school.)
おとうと おく りょこう い す
4) 弟 の 奥 さんは、旅 行 に行くのが好きですよ。

(My younger brother's wife likes travelling.)
たんじょうび ともだち きれい はな
5) □ □ □ □ □ □ □□□ □□ □ □□□□□□
(I received beautiful flowers from my friend on my birthday.)
しゅうまつ ふたり ひと しごと よてい
6) 週 末 から、二人の 人と仕事をする予定です。
(From this weekend I'm planning to work with two new people.)

Not at all native-like　　Native-like
まったくネイティブではない　　ネイティブレベル

1――――2――――3――――4――――5――――6――――7

ようす びょういん き
0. おばさんの様子を 病 院 で聞いてください。
Practice 1-3.
1――――2――――3――――4――――5――――6――――7

しゃちょう けっこんしき きゃく せんにんき
1－4. 社 長 の結 婚 式 に、お 客 さんが千 人 来た。
(One thousand people attended the president's wedding reception.)

1――――2――――3――――4――――5――――6――――7

つぎ じゅぎょう すうがく むずか
5－8. 次 の 授 業 の数 学 はちょっと □ □□です。
(Mathematics in the next class is a bit hard.)

1――――2――――3――――4――――5――――6――――7

わたし こうこう いっしょ しゃしん
9－12. 私 の高 校 で一 緒 に写 真 をとりましょう。
(Let's take a photo together at my high school.)

1――――2――――3――――4――――5――――6――――7

おとうと おく りょこう い す
13－16. 弟 の奥 さんは、旅 行 に行くのが好きですよ。
(My younger brother's wife likes travelling.)

1――――2――――3――――4――――5――――6――――7

たんじょうび ともだち きれい はな
17－20. 誕 生 日 に友 達 から □ □□花 をもらった。
(I received beautiful flowers from my friend on my birthday.)

1――――2――――3――――4――――5――――6――――7

しゅうまつ ふたり ひと しごと よてい
21－24. 週 末 から、二人の 人と仕事をする予定です。
(From this weekend I'm planning to work with two new people.)

1――――2――――3――――4――――5――――6――――7