A corpus of Japanese speakers' pronunciation of American English: preliminary research

Takehiko Makino
Chuo University

1 Introduction Corpus-based study of language has been flourishing for some time now, and the results of error analyses with learners' corpus are being applied for ELT purposes, such as English learners' dictionaries. Phonetics is lagging behind in this respect. This is understandable because speech data have to be transcribed phonetically, the task which not all linguists could do and which is very time-consuming. Nonetheless, corpora of this kind is beginning to appear, for example the Corpus of Spontaneous Japanese (Maekawa 2003) and the Buckeye Speech Corpus (of American English) (Pitt et al 2007).

Learners' phonetic corpus is being constructed for Japanese (Esaki et al 2004), but I do not know of any corpora of English spoken by Japanese. Phonetics textbooks such as Makino (2005) were written with no objective data of learners' pronunciation. Informal observation (="conventional wisdom") and the consideration of differences of phonemic systems were the only sources of what I wrote in that textbook.

The purpose of the present paper is to explore the possibility of constructing a corpus of Japanese speakers' pronunciation of English. Such a corpus would certainly improve what is written in English phonetics/pronunciation training textbooks for Japanese speakers. Fortunately, a large-scale database of English read by Japanese already exists. What I should do is to examine part of it, find out what variations in pronunciation to expect and identify possible problems in constructing a corpus.

2 The Data "English Learners' Speech Database" (Minematsu et al 2002) is a collection of recordings made in view of CALL system development which consists of thirteen sets of English material read aloud by Japanese university students. There are two kinds of sets: eight sentence sets and five word sets. The former consist of sentences derived from TIMIT phonemically-balanced sentence sets, phrases to be read in a single intonational phrase, sentences with intonation markings and the sentences with stress markings. The latter mainly consist of word/phrase/compound lists. The subjects were presented with phonemic transcriptions before recording so that the target pronunciation of individual words would not vary because of their unfamiliarity.

The number of subjects was 210, and each subject read one sentence set and one word set. Thus the same sentence sets were read by twenty-six subjects and the same word sets by forty-two subjects. The total of recorded utterances is about 70,000.

That the same set were read by a number of subjects means that the whole database contains individual differences of the same sentences/phrases/words, which are expected to vary greatly from subject to subject, according to their proficiency and idiosyncrasy.

3 The Method of Study One sound file for each speaker was taken out for investigation from the phonemically-balanced sets so that we could look at as many individual variations as possible. Then each file was transcribed phonetically by listening to it and referring to the waveform, spectrogram and fundamental frequency trace made by Wavesurfer version 1.8.5. Efforts were made to make as narrow segmental transcriptions as possible, but comparable prosodic transcriptions were not possible to make because one could not decide which prosodic system the speaker in question used, or if he or she was using some "interlanguage" prosody.

4 Notable Findings At the time of writing only seventy-five out of 210 sound files had been transcribed phonetically. We expect to have all the files transcribed by the conference. So the results have not been handled statistically, as is usual in corpus studies, but we still see some notable tendencies/findings from these limited data.

4.1 Items present in the "conventional wisdom"

4.1.1 /l/-/r/ distinction

Japanese speakers are notoriously unable to distinguish pre- and intervocalic /l/s and /r/s. In the observation made here, however, many speakers had no problems in using [ɹ] for /r/. More speakers had problems for pre- and intervocalic /l/s; they tend to replace it with [ɾ], one of the usual phones for Japanese /r/, while using [ɹ] for /r/ correctly. There were more problems for non-prevocalic /r/s; almost all of them were dropped or replaced with vowels such as [ɐ]. A number of speakers substituted [ɾɬ] for non-prevocalic /l/s, but more of them managed to use [ɬ] or vocalized [ɯ]. There were a couple of instances of intervocalic /r/ deletion, where the vowels that flank it were different. Deletion also happened to intervocalic /l/s, but in this case the flanking vowels were /i, i:/ and there was an instance where /l/ between [i]s was replaced with [ʎ].

	Conventional description	This study
Prevocalic /r/	[ɾ]	more [ɹ] than [ɾ], [l]
Prevocalic /l/	[ɾ]	[ɾ], [l], [ɹ]
Intervocalic /r/	[ɾ]	more [ɹ] than [ɾ], lost
Intervocalic /l/	[ɾ]	[ɾ], [l], [ɹ], [ʎ] (between [i]s), lost
Non-prevocalic /r/	Not mentioned	Mostly lost or substituted with [ɐ]; occasionally [ɹ]
Non-prevocalic /l/	Not mentioned, but assumed to be [ɾɯ]	[ɬ], [ɾɬ]

Table 1. /r/ and /l/ in conventional description and real data in this study

4.1.2 /s/-/θ/ distinction

In conventional description, Japanese lacks phoneme /θ/ and Japanese speakers are assumed to substitute it with [s]. It is true that there is frequently no contrast between /s/ and /θ/ in Japanese speakers' English, but the actual phonetic realization does not always match the conventional wisdom. Since there is currently a tendency among Japanese speakers to pronounce Japanese /s/ with [θ], English /s/ and /θ/ also tend to be realized by [θ]. In our data there are more substitutions of [θ] for /s/ than /θ/s replaced with [s], which is the norm in the conventional description.

4.2 Items not present in the "conventional wisdom"

4.2.1 Predictable variations In Japanese, intervocalic obstruents of the same place of articulation are not distinguished. Thus /b/ has initial [b] and intervocalic [β~b]; /d/ has initial [d̪] and intervocalic [ð~d̪]; /z/ has initial [dz] and intervocalic [z~dz], and before /i, j/, initial [dʒ] and intervocalic [ʒ~dʒ]; /g/ has initial [g] and intervocalic [ɣ~ŋ] (Makino 2005). These have complex interrelations with English system and this is reflected in the present data. Intervocalic /b/ is very frequently pronounced [β] and so is /v/. Word-initial /ð/ is often replaced with [d̪] and also with [z, dz]. Intervocalic /dʒ/ is in most cases pronounced [ʒ]. Intervocalic /g/ is in most cases pronounced [ɣ]. These are totally predictable substitutions.

Japanese moraic /N/ is realized as nasalized version of the following sounds, and in cases where it is phrase-final, uvular [ɴ] or nasalized version of the preceding vowels. In Japanese speakers' English, this is often applied to word-final /n/s and our data support the pattern. They are pronounced [ə̃ , ʊ̃ , ĩ, ɔ̃ , õ], not to mention [m, n̪ , n, ŋ].

4.2.2 Unpredictable variations I have also found cases which are notable in that they are not predictable from the systematic difference between Japanese and English. Word-initial plosives /b, d, g/ are occasionally pronounced with fricatives [β, ð, ɣ], even when the preceding words end with a consonant (i.e., the plosives are not intervocalic). This does not match the conventional description of Japanese phonetics, but if we examine the spontaneous speech corpus of Japanese mentioned above, this might indeed be a recurrent pattern we have yet to find.

There are cases where word-final voiced plosives are realized as *voiceless* fricatives, for example, *tube* [tjʉːɸ]. Usually Japanese learners of English are advised not to add a vowel such as [ə] or [ɯ] to a final obstruents to make the word conform to the CV syllable pattern which is the norm in Japanese. Apparently this case seems to be the result of trying to follow such advice. Notably, since there are basically no word-final consonants in Japanese, there is also no phonological process which devoices and spirantizes them. We just wonder what the motivations for such substitutions are.

Also notable are the cases where voiceless plosives are realized as fricatives, even prevocalically: *make pearls* [meɪxpɑːɾɬs] and *exposure* [ɛksɸoːʒə]. Again, since there is no such phonological process in Japanese, we do not have easy explanation of how such substitutions occur.

5 Discussion and conclusion We cannot exhaustively list all the deviations of such a small part of the whole data here. This in turn proves that ELSD contains a wealth of information concerning Japanese-accented English. However, there are some problems to address before setting out to construct a corpus.

First, as we saw above, there is at present no way to objectively transcribe prosodic aspects of the utterances. It is necessary to work out some sort of "hybrid" system cover the possible variations.

Second, to construct a corpus, it is necessary to consider the structure of data presentation. Raw phonetic transcriptions will be of little use because one does not know

what he or she could find before searching. In other words, the transcriptions should be linked to something one knows is there. Should they be linked to words they correspond? Or should they be linked to phonemes they may have substituted?

In any case, a corpus-based error analysis of foreign-accented pronunciation is a promising field of study and we believe that constructing a corpus is a worthwhile project.

References

Esaki, T. et al. (2004) The Project for the Corpus of Spontaneous Japanese Spoken by Non-Native Speakers. *IEICE technical report. Speech* Vol.104, No.148.

Maekawa, K. (2003) Corpus of Spontaneous Japanese: Its design and evaluation. *Proceedings of the ISCA & IEEE Workshop on Spontaneous Speech Processing and Recognition* (SSPR2003*)*, Tokyo.

Makino, T. (2005) *Nihonjin no Tame no Eigo Onseigaku Lesson*. [Lessons in English Phonetics for Japanese Speakers.] Tokyo: Taishukan Publishing Company.

Minematsu, N. et al (2002) English speech database read by Japanese learners for CALL system development. *Proc. Int. Conf. Language Resources and Evaluation* (LREC2002), pp.896-903.

Pitt, M. A. et al. (2007) Buckeye Corpus of Conversational Speech (2nd release) [www.buckeyecorpus.osu.edu] Columbus, OH: Department of Psychology, Ohio State University (Distributor).