愛知淑徳大学論集―人間情報学部篇　第5号　2015年3月，pp. 9-14

原著

Development of Japanese Read-sentence Database for Non-native Speakers of Japanese

Kimiko YAMAKAWA*, Shigeaki AMANO*, Mariko KONDO**

山川仁子，天野成昭，近藤眞理子

ABSTRACT

A Japanese read-sentence database was developed with 129 non-native speakers of Japanese, including Korean, Chinese, Thai, Vietnamese, French, and English speakers, and with 28 native speakers of Japanese. Two versions of Aesop's fable, "The North Wind and The Sun," were used as Japanese sentence materials for recording, and each recorded sentence of the story was registered in the database. The database will contribute to development of a computer-aided language learning system for Japanese speech education of non-native speakers. It will also contribute to scientific research that reveals characteristics of Japanese utterances by non-native speakers in terms of phonemes, prosody, and other aspects of speech.

Keywords : Non-native speaker, read sentence, Japanese

1. INTRODUCTION

Computer-aided language learning (CALL) is an effective and efficient method for language education for non-native speakers. CALL will become more popular with progress in computer technology and increasing Internet use. However, to develop a reliable and high-performance CALL system, a speech database of non-native speakers is necessary as a resource for statistical training and system parameter setting.

Speech databases of non-native speakers have been developed in several languages. For example, English across Taiwan (EAT) (Chen, Yu, & Wang, 2010) is a database for English spoken by Chinese speakers in Taiwan. English Read by Japanese (ERJ) (Minematsu, 2010) is another example database for English spoken by Japanese speakers.

As for the Japanese language, Nishina (2010) constructed a database for utterances of 141 non-native speakers of Japanese, including Chinese, Korean, Thai, Vietnamese, Malaysian, Indonesian, Arabic, Spanish, French, and English speakers. The database contains read speech of sentences, minimal-pair words, and dialogues.

We developed a read-sentence database to supply additional speech data of non-native speakers of Japanese. The database can be expected to contribute to development of a CALL system, as previous databases have done. It will also contribute to scientific research that reveals characteristics of Japanese utterances by non-native speakers in terms of phonemes, prosody, and other aspects of speech.

* Aichi Shukutoku University

** Waseda University

2. DATABASE DEVELOPMENT

2.1. Speakers

Native speakers of Korean, Chinese, Thai, Vietnamese, French, and English participated in the recording as non-native speakers of Japanese. There were 129 non-native speakers of Japanese in total. The non-native speakers could at least read Hiragana, a moraic alphabet for the Japanese language. However, the speakers had wide-ranging Japanese proficiency, because it was difficult to recruit speakers with the same Japanese proficiency across different countries. Speakers' profile information, such as Japanese skill, learning period of Japanese, living period in Japan, birthplace, and early development location, was collected using a questionnaire after a recording.

In addition to these non-native speakers of Japanese, 28 native speakers of Japanese participated in the recording. They were born and live in Tokyo metropolitan and its suburbs.

The number of speakers of each language is shown in Table 1. All speakers had also participated in the recordings of the Japanese Read-Word Database (Yamakawa, Amano, & Kondo, 2014). They were paid for recording.

2.2 Sentence materials

The IPA and the PASL-DSR versions of Aesop's fable, "The North Wind and The Sun," were used as the sentence materials for recording. Each version consisted of eight sentences.

The IPA version (Appendix 1) was derived from a Japanese phonetic transcription of "The North Wind and The Sun" shown in "Handbook of the International Phonetic Association" (International Phonetic Association, 1999). Japanese orthography given to the phonetic transcription was used as the IPA version. The IPA version was adopted, because it has been used for recording in many languages, including Japanese. The IPA version enables a comparison between non-native and native speakers of Japanese, because there are many recordings of the IPA version by native speakers of Japanese.

The PASL-DSR version (Appendix 2) was the sentence set of "The North Wind and The Sun" that Itahashi (1992) used for development of a Japanese speech database. The PASL-DSR version was used, because it has been used in research on an automatic speech recognition system for native speakers of Japanese. If speech data with the same version of the sentence set are supplied, the data could easily be used in research to develop an automatic speech recognition system for speech education for non-native speakers of Japanese.

Table 1　Characteristics of speakers participated in the recording

Native Language	Number of speakers			Age (year)			
	Male	Female	Total	Mean	Min.	Max.	SD
Japanese	14	14	28	25.8	20	30	3.4
Chinese	9	19	28	21.8	21	26	1.1
Vietnamese	11	13	24	21.4	20	23	0.8
French	7	15	22	23.0	20	27	2.3
Korean	2	15	17	21.1	19	25	1.6
Thai	10	19	29	20.7	18	23	1.3
English	7	2	9	29.2	19	49	10.7
Total	60	97	157	23.0	18	49	4.2

2.3 Recording equipments

Recording equipment included a microphone (ECM-999, SONY), an A/D converter (UA-25EX, Roland), and a personal computer (Dynabook SSRX2/T9L, TOSHIBA) with an extended display (SK-DTV133JW2, SK-NET, or On-Lap 1302/J, GECHIC) and a flexible keyboard (CB-KB85U02-BK, CLEVERY).

A linear PCM recorder (DR-1, TASCAM) connected to a microphone (ECM-999, SONY) provided a simultaneous backup recording.

2.4 Recording procedure

Recording was conducted in a quiet room using the computer. One of the sentences of "The North Wind and The Sun" was presented on the computer display in Japanese Kanji and Hiragana characters, as shown in Appendices 1 and 2. Recording started two seconds after a speaker pressed the start key. The speaker read the presented sentence aloud, and then pressed the stop key to finish the recording. The speaker's utterance was stored in the personal computer as a digital file with 16-bit quantization and 48-kHz sampling frequency.

The computer automatically checked the recorded sentence. When the intensity of the sentence was too low or too high, or when the beginning or end of the word was not properly recorded, the computer made an alert. In such a case, the sentence was recorded again. In addition to the checking by computer, an operator monitored the speaker's pronunciation, and if problems such as mispronunciation or hesitant pronunciation were found, the sentence was re-recorded.

All sentences were recorded individually by repeating this sequence. The recording order of the two versions of "The North Wind and The Sun" was counterbalanced between speakers. Speakers took a short break between the recordings of the two versions.

Recording venues were Sangmyung in Korea for Korean speakers, Taipei in Taiwan for Chinese speakers, Bangkok in Thailand for Thai speakers, Hanoi and Ho Chi Minh City in Vietnam for Vietnamese speakers, and Bordeaux in France for French speakers. English speakers and native speakers of Japanese were recorded in Tokyo in Japan.

3. DATABASE CONTENTS

Speech files for each sentence of the two versions of Aesop's fables, "The North Wind and The Sun" (Appendix 1 & 2), read by 129 non-native and 28 native speakers of Japanese are stored in the database in a WAV format with 16-bit quantization and 48-kHz sampling frequency.

There are 2,576 speech files (eight sentences × two versions × 161 speakers) in the database. Total size of the database is about 2 GB, corresponding to about 350 minutes of recording.

Phoneme labels have not been supplied to the speech files, but the labels will be registered in the database in future work.

4. DISCUSSION

One advantage of the database developed in this study is that recording of both Japanese natives and non-natives was conducted with the same sentence materials using the same recording equipment and procedure. This enables precise comparison of non-native speaker's utterances to the reference utterances of native speakers of Japanese. It also enables comparisons among languages of non-native speakers.

Another advantage is that, in contrast to the previous database (Nishina, 2010), recording was conducted in the home countries of the non-native speakers except for English speakers. This would be expected to provide more realistic speech data of non-native speakers' Japanese utterances that are affected by their native languages, because non-native speakers in their home countries have less opportunity than their counterparts in Japan to hear and speak Japanese and have more frequent contact with their native languages.

Use of the same sentences used in previous studies (e.g., Itahashi, 1992) of Japanese is another positive feature of the current database. This enables us to enlarge the set of Japanese native speakers' data by combining the previous and current speech recordings.

Although this database involves fewer languages than the previous database (Nishina, 2010), five out of the seven languages have more than 20 speakers, which would be sufficient for speakers of each language to extract the characteristics of Japanese utterances. For Korean and English, the number of speakers will be increased in future work.

Similarly to the previous database (Nishina, 2010 ; Itahashi, 1999), this database was developed for read speech but not for spontaneous speech. It might fail to capture some important utterance features of non-native speakers, because the read speech would have less variation than spontaneous speech. It would be preferable to use a database of spontaneous Japanese speech by non-native speakers (e.g., Usami, Kagomiya, & Sugimoto, 2004) together with the current database, to understand the exact characteristics of non-native speaker utterances.

5. ACKNOWLEDGMENT

This study was supported by JSPS KAKENHI Grant Numbers 22320081, 24652087, 25284080, and 26370464 and by a special research grant (2013-2014) and a cooperative research grant (2013-2014) from Aichi-Shukutoku University. Part of this study was supported by the NINJAL core collaborative research project "Foundations of Corpus Japanese Linguistics."

We would like to thank Professor Kyung-Ja Ryoo and Lecturer Kazuya Iihoshi of Sangmyung University, Professor Luong Chi Mai of Vietnam Academy of Science and Technology, Professors Nguyen Thu Huong and Nguyen Tien Luc of Vietnam National University - Ho Chi Minh City, Professor Chiu-yu Tseng of Institute of Linguistics, Academia Sinica, Professors Chang-Ho Lin and Meng-Ling. Hsu of Ming Chuan University, Professor Rong-Kuan Shen of Shih Hsin University, Professor Yaw-Huei Maa of Tamkang University, Dr. Chatchawarn Hansakunbuntheung of National Electronics and Computer Technology Center, Professor Kanokwan Atchariyachanvanich of King Mongkut's Institute of Technology Ladkrabang, Professor Nagul Cooharojananone of Chulalongkorn University, and Dr. Takaaki Shochi of University of Bordeaux for their assistance in the utterance recordings.

6. REFERENCES

Chen, S., Yu, H., & Wang, Y, (2010). The EAT (English across Taiwan) Corpus. In Itahashi, S., & Tseng, C. (Eds.), *Computer processing of Asian spoken languages*, Consideration Books, Los Angeles, pp. 155-158.

International Phonetic Association (1999). *Handbook of the International Phonetic Association*, Cambridge University Press.

Itahashi, S. (1992). Development of a Japanese speech database for research in speech information processing. *The 1991 report*

of scientific research fund (in Japanese).

Minematsu, N. (2010). Development of ERJ (English Read by Japanese) database for CALL research. In Itahashi, S., & Tseng, C. (Eds.), *Computer processing of Asian spoken languages*, Consideration Books, Los Angeles, pp. 151–154.

Nishina, K. (2010). Construction of speech database for second language learning of Japanese. In Itahashi, S., & Tseng, C. (Eds.), *Computer processing of Asian spoken languages*, Consideration Books, Los Angeles, pp. 147–150.

Usami, H., Kagomiya, T., & Sugimoto, S. (2004). Design of a database of Japanese learner's dialogue in Japanese and Japanese learner's native language. *IEICE*, SP2004-24, 29–34 (in Japanese).

Yamakawa, K., Amano, S., & Kondo, M. (2014). Development of Japanese read-word database for non-native speakers of Japanese. *Proceeding of the 17th Oriental COCOSDA*, 65–70.

7. APPENDIX 1 : IPA VERSION

kti001 : ある時(とき)、北風(きたかぜ)と太陽(たいよう)が力(ちから)くらべをしました。
kti002 : 旅人(たびびと)の外套(がいとう)を脱(ぬ)がせたほうが勝(か)ちということに決(き)めて、まず北風(きたかぜ)から始(はじ)めました。
kti003 : 北風(きたかぜ)は、「なに、ひとまくりにして見(み)せよう」と、激(はげ)しく吹(ふ)きたてました。
kti004 : すると旅人(たびびと)は、北風(きたかぜ)が吹(ふ)けば吹(ふ)くほど外套(がいとう)をしっかりと体(からだ)にくっつけました。
kti005 : 今度(こんど)は太陽(たいよう)の番(ばん)になりました。
kti006 : 太陽(たいよう)は雲(くも)のあいだから、優(やさ)しい顔(かお)を出(だ)して暖(あたた)かな光(ひかり)を送(おく)りました。
kti007 : 旅人(たびびと)は段々(だんだん)よい心(こころ)もちになって、しまいには外套(がいとう)を脱(ぬ)ぎました。
kti008 : そこで北風(きたかぜ)の負(ま)けになりました。

IPA transcriptions of these sentences are shown below.

kti001 : aɾɯtokʲi̥, kʲi̥takaze to taijoː ga tɕi̥kaɾakɯɾabe o ɕimaɕi̥ta.
kti002 : tabʲibʲito no gaitoː o nɯgaseta hoː ga katɕi̥ to juː koto ɲi kʲimete, mazɯ, kʲi̥takaze kaɾa haʑimemaɕi̥ta.
kti003 : kʲi̥takaze wa, naɲi, hi̥tomakɯɾʲi ɲi ɕi̥te mʲisejoː to, hageɕi̥kɯ ɸɯ̥kʲitatemaɕi̥ta.
kti004 : sɯɾɯto tabʲibʲito wa, kʲi̥takaze ga ɸɯ̥keba ɸɯ̥kɯhodo gaitoː o ɕi̥kːaɾʲito kaɾada ɲi kɯ̥ːtsɯkemaɕi̥ta.
kti005 : kondo wa taijoː no baɲ ɲi naɾʲimaɕi̥ta.
kti006 : taijoː wa kɯmo no aida kaɾa jasaɕiː kao.o daɕi̥te, atatakana ɕi̥kaɾʲi o okɯɾʲimaɕi̥ta.
kti007 : tabʲibʲito wa dandan joi kokoɾomotɕi ɲi natːe, ɕimai ɲi wa gaitoː o nɯgimaɕi̥ta.
kti008 : sokode kʲi̥takaze no make ɲi naɾʲimaɕi̥ta.

8. APPENDIX 2 : PASL-DSR VERSION

ktp001 : あるとき、北風(きたかぜ)と太陽(たいよう)が力(ちから)くらべをすることになりました。
ktp002 :「あそこを歩(ある)いていく、旅人(たびびと)の外套(がいとう)を脱(ぬ)がせた方(ほう)を勝(か)ちにしよう。」と決(き)めて、まず北風(きたかぜ)から始(はじ)めました。
ktp003 : 北風(きたかぜ)は、自信満々(じしんまんまん)に「ひとまくりしてみせよう。」といい、激(はげ)しく

吹(ふ)き立(た)てました。

ktp004 : ところが、風(かぜ)が激(はげ)しく吹(ふ)けば吹(ふ)くほど旅人(たびびと)は、外套(がいとう)をしっかりと押(お)さえてしまいました。

ktp005 : 今度(こんど)は太陽(たいよう)の番(ばん)になりました。

ktp006 : 太陽(たいよう)が雲(くも)の間(あいだ)から暖(あたた)かな光(ひかり)を送(おく)ると、旅人(たびびと)は汗(あせ)をかき始(はじ)めました。

ktp007 : そこで、太陽(たいよう)がもっと光(ひかり)を強(つよ)くすると、旅人(たびびと)は、「暑(あつ)くなってきたなあ。」といって外套(がいとう)を脱(ぬ)ぎました。

ktp008 : こうして、この力(ちから)くらべは太陽(たいよう)の勝(か)ちになりました。

IPA transcriptions of these sentences are shown below.

ktp001 : aɾɯtokʲi kʲitakaze to taijoː ga tɕikaɾakɯɾabe o sɯɾɯkoto ɲi naɾʲimaɕita.

ktp002 : asoko.o aɾɯite ikɯ, tabʲibʲito no gaitoː o nɯgaseta hoː o katɕi ɲi ɕijoː to kʲimete, mazɯ, kʲitakaze kaɾa haʑimemaɕita.

ktp003 : kʲitakaze wa, dʑiɕimːamːaɲ ɲi, çitomakɯɾʲi ɕite mʲisejoː to i.i, hageɕikɯ ɸɯkʲitatemaɕita.

ktp004 : tokoɾoga, kaze ga hageɕikɯ ɸɯkeba ɸɯkɯhodo tabʲibʲito wa gaitoː o ɕikːaɾʲito osaete ɕimaimaɕita.

ktp005 : kondo wa taijoː no baɲ ɲi naɾʲimaɕita.

ktp006 : taijoː ga kɯmo no aida kaɾa atatakana çikaɾʲi o okɯɾɯ to tabʲibʲito wa ase o kakʲihaʑimemaɕita.

ktp007 : sokode, taijoː ga motːo çikaɾʲi o tsɯjokɯsɯɾɯto, tabʲibʲito wa, atsɯkɯnatːe kʲitanaː to itːe gaitoː o nɯgimaɕita.

ktp008 : koːɕite kono tɕikaɾakɯɾabe wa taijoː no katɕi ɲi naɾʲimaɕita.