A Computer Program for Pronunciation Training and Assessment in Japanese Language Classrooms – Experimental Use of 'Pronunciation Check'

CHIHARU TSURUTANI, *Griffith University, Australia*

E-learning is today an ever more popular mode for second language learning. However, the relative newness of using a computer program for pronunciation training and assessment (CAPTA) in language classrooms presents the need to explore possibilities and limitations of this mode to achieve the most effective outcomes possible. This paper illustrates the process to develop CAPTA from the perspective of educators and discuss the effectiveness and limitations of CAPTA in the current technological environment.

Introduction

Globalization of commerce and education makes speaking a second language an important skill in workplaces worldwide. In response to this demand, educators have employed a range of teaching methods and materials that focus on general language communication skills but these generally neglect teaching pronunciation.[1] However, an accurate or at least comprehensible level of pronunciation is a foundation of effective verbal communication.[2] In the classroom, individual teachers rarely have the capacity or time to consistently check and correct each student's pronunciation, especially in large classes. Even when teachers are able to point out learners' pronunciation errors, learners may not know how to correct these errors, and teachers may not have the requisite knowledge of phonetics and phonology to advise on correction. For adult learners in particular, correction of their pronunciation in front of others may even discourage them from continuing their study.

Yet if pronunciation is corrected by computer on an individual basis, learners do not have to worry about awkward repetition in front of others and can practice pronunciation privately, as many times as they like.[3] Since improving pronunciation tends to require more time and devotion than other areas of language learning, as first language transfer (traces of the first language in a second-language learner's production) is most evident in phonetics and phonology,[4] a computer program for pronunciation training and assessment (CAPTA) is an ideal educational tool to supplement the shortcomings in current methods of foreign language teaching and learning. To fulfill this pedagogical need, the researcher developed a CAPTA software for second language learners of

---

[1]Celce-Murcia, Brinton and Goodwin, *Teaching Pronunciation.*
[2]Jenkins, 'Research in Teaching Pronunciation and Intonation'; Raux and Kawahara, 'Optimizing Computer-Assisted Pronunciation Instruction'.
[3]Kaltenböck, 'Learner Autonomy'.
[4]Norris and Ortega, 'Effectiveness of L2 Instruction'; Lado, *Linguistics across Cultures.*

Japanese together with a team of engineers and trialed it in a Japanese language course. In this paper, the production and the experimental use of the CAPTA software in the language course will be reported and the implication of the results discussed.

Use of CAPTA in Language Education

For the last couple of decades, various computer assisted language learning (CALL) programs have been developed and become widely available. However, progress in development of CALL programs for pronunciation has been relatively limited by comparison with development of reading and listening materials, due to language educators' lack of knowledge of speech technology. Assessment by humans is to some extent subjective in any area of teaching, and this is particularly so in the assessment of pronunciation. CAPTA has been found to be a great aid for teachers, especially for non-native teachers of the target language.[5] Segments of CAPTA programs have been developed in engineering laboratories for experimental purposes,[6] but they have not satisfied pedagogical conditions for classroom use.[7] Software currently available for pronunciation training takes one of the following forms:

- 'Playback and comparison' The user compares his/her production through listening, using a model for correct pronunciation provided by the system;
- 'Visible speech aids' Comparison of spectrograms[8] (or other visible acoustic displays) of a native speaker's model with a comparable physical display of the learner's utterance (e.g. 'Accent Lab', released 1997; 'Vowel Target', marketed 1998–2005).

In these forms, the judgment and interpretation of learners' performance have to rely on learners themselves, although some software can provide feedback in a limited manner, such as indicating good or not good.[9] These computer programs have not been able to provide information on why learners achieved certain scores and what learners need to do to improve their scores, in a systematic way. Besides, without knowledge of acoustic phonetics it is difficult to understand the display screen of spectrograms. Thus, a computer program that can assess and give feedback on specific aspects of learners' pronunciation has long been sought in this area.

Development of the CAPTA Program for Japanese Language

In 2006 collaboration between engineers and language educators, funded by Panasonic Foundation, enabled the initial development of CAPTA software called Pronunciation Check, which uses Automatic Speech Recognition (ASR).[10] Pronunciation Check is a self-assessment program for Japanese pronunciation which detects learners' errors and

[5]Kim, 'Automatic Speech Recognition'.
[6]Furui, Iwano and Hirose, 'Onsei ninshiki o mochiita'; Hirose, Gendrin and Minematsu, 'A Pronunciation Training System'.
[7]O'Brien, 'Teaching Pronunciation and Intonation'.
[8]A continuous analytic display of acoustic properties of sound over time.
[9]Auralog, *Tell Me More English* [computer software]; Bungalow Software, *Vowel Target* [computer software]. See Miwa, Sasaki and Tanno, 'Japanese Spoken Language Learning System'.
[10]Tsurutani, Yamauchi and Minematsu, 'Development of a Program'.

provides scores and feedback to improve learners' pronunciation. This initial
development was for an assessment tool for Japanese language in Australia. The
current model has the following functions:

1) *Detection of Pronunciation Error*
   The program detects typical pronunciation errors of non-native speakers of
   Japanese (segmental errors only).
2) *Assessment of Pronunciation*
   Learners read a Japanese sentence that appears on the screen. The program
   provides scores for each phrase in the sentence and an overall score for the whole
   sentence. A response to learners' question is also provided.
3) *Feedback*
   The program teaches ways to improve pronunciation by explaining how to
   produce sounds correctly.
4) *An Editorial System*
   An editorial system allows teachers to modify test words or sentences as well as
   mark criteria according to students' levels.
5) *Monitoring*
   Since the server stores access information such as recording date, time, perfor-
   mance score and error types, the participant and his/her teacher can view previous
   performance and compare it with his/her current trial.

These pedagogical considerations make the program unique and attractive to users. An
editorial system is particularly innovative and gives teachers freedom to create sentences
they wish to test. In the pilot survey conducted in 2006, the assessment component was
tested and favorable responses to a questionnaire were received from 14 participants,
commenting on the program's usefulness for self-study.

Errors to be Detected by the Program

The ASR system employs the pattern of information matching between acoustic input
and language models stored in the computer through the Hidden Markov Model
(HMM).[11] For this acoustic information matching, a large number of acoustic models
are normally required to recognize the pronounced utterance accurately. Alternatively,
by refining the language models in the system we can reduce variations of possible
utterances from the predicted word sequence of acoustic input and detect the actual
utterance. The detection of error patterns plays a role in this stage of the program. The
list of error patterns was created drawing from research on the second language speech of
English speakers[12] and the researcher's experience as a language teacher. The patterns
were ranked in order from the most frequent error to the least frequent error. To test the
reliability of the list, 1,200 sentences in Japanese language spoken by 30 English-speaking
learners were transcribed and examined. The original list (1–16) covered 90 percent of
learners' errors. Of the remaining 10 percent, errors that occurred more than a few times
were added to the list (17–19), and the remaining errors were discarded since they were
idiosyncratic and were no more than slips of the tongue. (If such errors occur, the

---

[11]A statistical model to extract model parameters to perform pattern recognition.
[12]Tsurutani, 'Acquisition of Word Prosody'; Tsurutani, 'Speech Rate'.

The Experimental Use of Pronunciation Check

In Semester 1 of 2007, Pronunciation Check was introduced to 52 students (15 Chinese or Korean native speakers and 37 English native speakers) in a second year Japanese course as part of continuous assessment at Griffith University in Brisbane, Australia. The students had completed 160 hours of class study in Japanese at university at the stage when the task was introduced, but some had also studied Japanese at high school before beginning university study.

Exercises Created by Pronunciation Check

Students were instructed to practice the three exercises below during private study hours, with three percent of the assessment requirements for this subject granted automatically to students who completed the exercises required for the mark.

1) Ten sentences for beginners – for practice at least once in March and in May;
2) LAL2032 (the subject code at Griffith University) – practice after studying each lesson;
3) Interview for exchange program – an interactive speaking exercise, try any time.

While Exercises 1 and 2 followed the pattern of reading a sentence that appears on the screen, Exercise 3 ('Interview for Exchange Program') involved a role play and simulated an interview session that students might undertake before going to Japan. Exercise 2 contained sentence patterns from the Japanese-language subject (LAL2032) the students were taking. The third exercise had more interactive features than the first two exercises, and required students to listen to questions.

The predicted error patterns and learners' linguistic levels were taken into consideration for the production of sentences used in these topics. Errors programmed for detection in the sentences are indicated after each phrase by numbers from the error list. This task is called '*Option shitei*' (Option Indication) and the system will detect only the indicated error. These option numbers or dashes which indicate long vowels are for programming and do not appear on the actual display as shown in Figure 2. The 10 sentences for beginners follow:
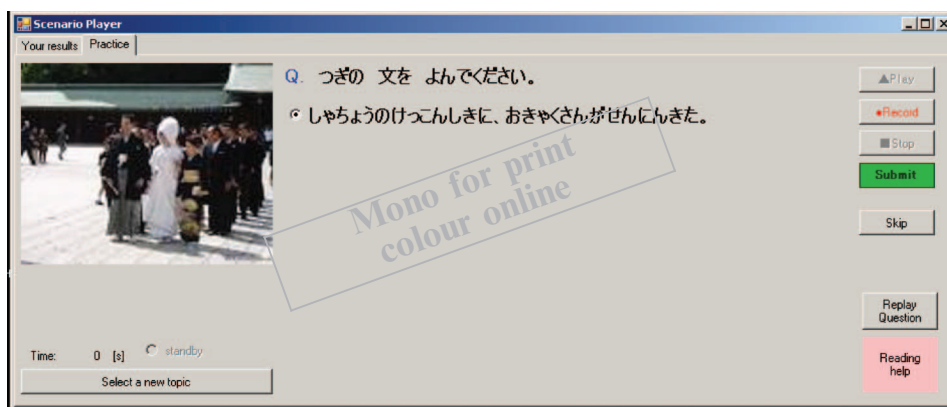

Figure 2. Computer Screen Displaying Exercise.

1. しゃちょー(1)の(2,4,5,6) けっこんしきに(2,3,9) おきゃく(1)さんが (2,5,6) せんにん(10) きた(1,2,9)

2. おばさんの(2) よーすを(2,4,19) びょーいんで(4,5,6,17,12) きーて (4,19,9) ください(12,13,16,17)。

3. わたし(1)の(2) こーこーで(4,17,19)、いっしょに(2,3,18,5,6) しゃしん(1)を(2,5,6,10) とりましょー(4,5,6,7)

4. しゅーまつ(1)から(4,5,6,7,8)、ふたりの(7,11) ひと(1)と(2) しごと(1) を(2,12) する(2,7) よてーです(2,4,9,13,12)

5. ゆーびんきょく(1)で(2,4,5,6,17,12) ひゃくえんの(5,6,10) きってを (2,3,18,9) ごまい(2,12,13) かった(3,18)

6. おとーとの(2,4,19) おくさん(1)は(2,9) 、りょこー(1)に(2,4,5,6,7) いく(1)のが(2) すき(2,12)ですよ(12,15)

7. たんじょーびに(2,4,5,6,12) ともだち(1)から(2,7,17,12) きれーな(2,4,7,9) はなを(2) もらった(3,18,7,9)

8. がくせー(1)が(4,12)、だいがく(1)の(2,12,13) としょかん(1)で(2,5,6) べんきょー(4,5,6,12,19) して(1)います(14)。

9. つぎの(2,8,12) じゅぎょーの(2,4,5,6,12) すーがく(1)は(2,4,12,9) ちょっと(3,18,5,6) むずかしーです(4,12,14)

10 ぼく(1)は(2,12)、らいねん(7,13) にほんへ(2,10) りゅーがく (1,2,4,5,6,7,12) しようと(4,15,19) おもいます(14,15,9)。

1. One thousand people attended the president's wedding reception.
2. Please check how my aunt is at the hospital.
3. Let's take a photo together at my high school.
4. From this weekend I'm planning to work with two people.
5. I bought five 100-yen stamps at the post office.
6. My younger brother's wife likes travel.
7. I received beautiful flowers from my friend on my birthday.
8. Students are studying at the university library.
9. The maths in the next class is a bit hard.
10. I'm thinking of going to Japan to study next year.

The button 'reading help' provides English translation or *hiragana* reading and can be used particularly for sentences with many *kanji* (Chinese characters). The version for beginners contains few *kanji* so that users will not have difficulty in reading.

Most students used the university language laboratory for recording their readings, although the program was made accessible from home computers. Only students who were confident with downloading software and had appropriate recording equipment used their home computer.

Students' Performance and Evaluation

Students' performances varied. Of the 52 students enrolled in the subject, 23 completed the task and were therefore awarded the three percent automatically for their grade, while 24 completed exercises related to this Japanese-language subject but did not finish some part of the first exercise and therefore were awarded 1.5 to two percent depending on how much of the task they completed. Five students did not participate in the program. Students who did not participate or try their best may have thought that the three percent grade was not a sufficiently significant portion of the assessment to warrant the effort. However, the fact that the majority of students completed the exercise for LAL2032 indicates that they are more comfortable articulating words and sentences familiar to them. It implies that an exercise directly related to their study has more appeal to them even though all the exercises are equally at their proficiency level. Even though the third exercise had the most entertaining features, it was taken up mainly by students who had studied Japanese prior to university and were more confident in speaking.

The following questionnaire was used for evaluating the Pronunciation Check program.

1) Was the program easy to use?
      1-1. Recording
      1-2. Viewing your results
2) Was the program fun to use?
3) Did the program clearly indicate weak points in your pronunciation?
4) Did the program explain well what needs to be done to improve your pronunciation?
5) What other functions do you think the program should have?
6) Do you think the program is useful for your study?
7) Which do you prefer for practicing pronunciation with, a teacher or a computer? Please provide the reason.

The answers, indicated on a Likert scale from 1 (good) to 5 (poor), were averaged and summarized depending on students' marks indicating the amount of the task they completed.

As expected, students who did not complete the task considered it not useful for their study. The proportion of Chinese and Korean speakers in each group was 36 percent (9/23) in the top group, 22 percent (4/18) in the second group and 17 percent (1/5) in the third group respectively. The five students who chose not to undertake the task were all native speakers of English who had studied Japanese longer than two years. It seems that students' language background does not really affect their participation and performance. Overall, in their assessment of the program participants gave high scores

TABLE 2. Participants' Responses to Program Evaluation.

Marks	Number of students	1) 1	1) 2	2)	3)	4)	6)
3	23	2.30	1.78	2.00	2.26	2.52	2.00
2~1.5	18	2.00	2.00	2.00	2.44	2.78	2.17
1~0.5	6	3.16	3.00	2.83	3.33	3.67	3.67
Average		2.30	2.00	2.10	2.50	2.70	2.30

on both its usefulness and usability. However, they indicated strong preference to practice with humans rather than with a computer program. Their responses to Question 7, 'Which do you prefer for practicing pronunciation with, a teacher or a computer?' can be summarized as follows:

- Teacher – 32
  A teacher is real, personal and has the ability to listen, think and explain.
  Preference is for more face to face interaction.
  Computers are just cool and fun to use [but can't provide feedback like a teacher can].
- Computer – 2
  Hard to find time with a teacher.
- Both or either – 12
  Using computer means you don't get embarrassed.
  Computer allows us to attempt many times.
  Computer is good to consolidate our speaking after class, but teacher has ability to explain complex errors.
- No response – 1

A machine certainly cannot compete with a human in flexibility of communication. But it is also true that people cannot precisely assess the pronunciation of their speaking partner and correct it on the spot while conducting a conversation. Recognizing these strengths and weaknesses, it appears appropriate to acknowledge that machines and humans should both be able to contribute to practice in pronunciation.

The students' comments included requests to provide on computer programs other tasks such as listening, grammar check and *kanji* exercises. It is certain that computer exercises have potential to broaden the horizon of language teaching for the new generation of learners.

Effectiveness

Students were told to try the first exercise, '10 sentences for beginners', at the beginning and at the end of the semester. Of the 23 students who followed the instruction, 15 (eight English speakers, seven Korean/Chinese speakers) had a complete set of 10 sentences. Scores for these students who completed the first exercise twice are plotted in Figure 3 against their overall semester grade (scored out of $1000 = 100 \times 10$ sentences).

In their second attempt (at the end of semester) 10 students improved their score while five students had a lower score than their attempt at the beginning of semester. It is hard to judge whether any change in score is due to students' language practice
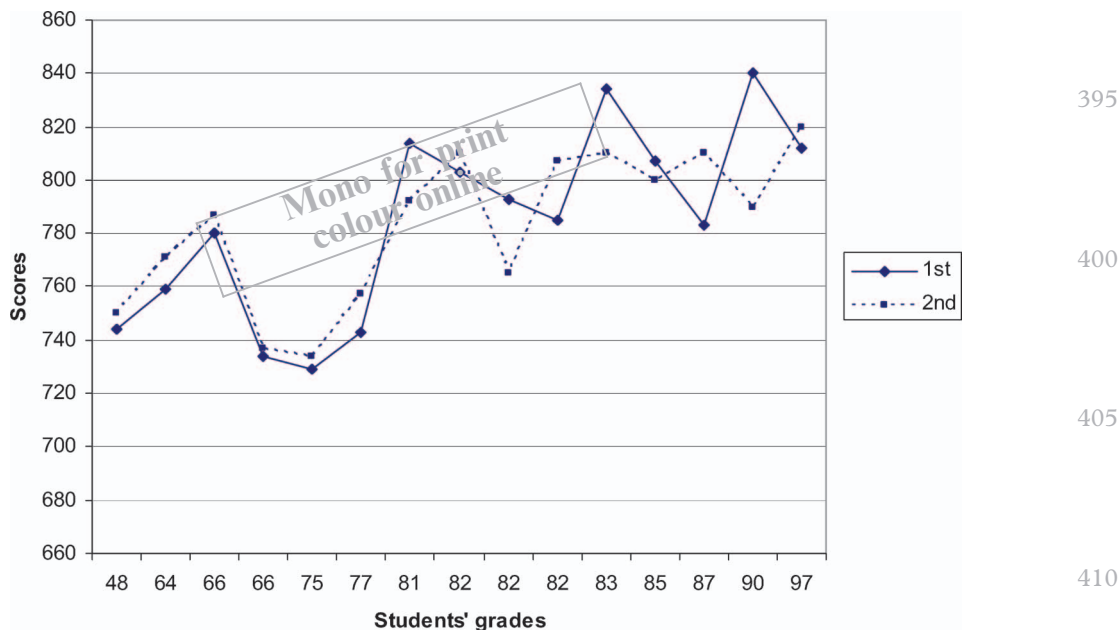
FIGURE 3. Participants' Scores and Their Final Grades.

or to improvement in their ability through taking the language course. Since the total amount of time students used the program was no more than five hours for this exercise, it would be unrealistic to expect that the CAPTA program contributed to improvement of their pronunciation. However, it certainly increased the amount of time they spent studying Japanese. Although the number of participants is small, there is some correlation ($r = 0.69$) between learners' academic results and the scores they obtained from the program (high achievers' scores are higher than the scores of students with low grades.) Many were found to be highly motivated students and fairly confident speakers of Japanese. Further examination of a large number of participants is required to determine whether or not the correlation is merely related to their level of motivation.

Limitations of the CAPTA Program

Capturing running speech is a difficult task for computers due to coarticulation, individual speakers' differences, recording conditions and some other technical factors. ASR, which uses the Hidden Markov Model, cannot reach 100 percent accuracy at its present stage of development. Here we need to rely on the advancement of technology. It is also desirable to be able to assess prosodic features of learners' speech, as duration, is often affected by pitch patterns.[13] Some attempt to assess pitch patterns has been made,[14] but we will need to allow more time for speech engineering researchers to achieve their goal.

---

[13]Tsurutani, 'Acquisition of Word Prosody'.
[14]Hirose, 'Speech Prosody and CALL'.

Educators can certainly make good use of CAPTA programs and should be encouraged to integrate them effectively into language classrooms. However, occasional malfunctions and connection difficulties can discourage users from trying out the program. Upon implementation administrators of a CAPTA need to make sure the network setting of their institution is reliable.

Summary

This paper has reported on an experimental use of a CAPTA program and has discussed the program's effectiveness and limitations. The CAPTA program Pronunciation Check is useful for self-study of pronunciation by compensating for the limitations of human conversation partners as correctors of pronunciation.

Although the development of CAPTA programs has not reached the stage where they can be used as assessment tools, survey results reported in this paper show the potential to use CAPTA programs as indicators of proficiency level. Based on the results of the participants' evaluation and the capacity of the current CAPTA program, it is considered that the program effectively serves the following purposes:

- practicing a difficult sound on your own
- monitoring your own pronunciation
- familiarizing yourself with new sentence patterns.

To cope with ever-increasing demands for cost-effective education, some aspects of labor-intensive language teaching will need to rely on technological advancements in the foreign-language classroom of the twenty-first century. Pronunciation training and assessment is one area where the introduction of CALL programs is considered to be effective. Further work is required for pedagogical and technological improvement of the CAPTA program to respond to the demand for better computer programs to assist in foreign-language teaching.

References

Auralog, Tell Me More English [computer software]. London: Auralog, 2000.

Bungalow Software, Vowel Target [computer software], 1998–2005. Available at http://www. bungalowsoftware.com.

Celce-Murcia, Marianne, Donna M. Brinton and Janet M. Goodwin, *Teaching Pronunciation*. Cambridge: Cambridge University Press, 1996.

Furui Sadaoki, Iwano Kōji and Hirose Sachio, 'Onsei ninshiki o mochiita Nihongo jiritsu gakushū system no kenkyū [Research on a Japanese Self-Learning System Using Speech Recognition]'. Tokyo Institute of Technology, 2002; http://www.furui.cs.titech.ac.jp/~iwano/papers/tokuteiA_01.pdf.

Hirose, Keikichi, Frédéric Gendrin and Nobuaki Minematsu, 'A Pronunciation Training System for Japanese Lexical Accents with Corrective Feedback in Learner's Voice', *Proc. EUROSPEECH*, 2003, 3149–3152.

Hirose Keikichi, 'Onsei no inritsu to CALL [Speech Prosody and CALL]', *Journal of the Phonetics Society of Japan*, 9:2 (2005), 38–46.

Jenkins, Jennifer, 'Research in Teaching Pronunciation and Intonation', *Annual Review of Applied Linguistics* 24 (2004): 109–125.

Kaltenböck, Gunther, 'Learner Autonomy: A Guiding Principle in Designing a CD-ROM for Intonation Practice', *ReCALL*, 13:2 (2001): 179–190.

Kim, In-Seok, 'Automatic Speech Recognition: Reliability and Pedagogical Implications for Teaching Pronunciation', *Educational Technology and Society*, 9:1 (2006): 322–334.

Lado, Robert, *Linguistics across Cultures: Applied Linguistics for Language Teachers.* Ann Arbor: University of Michigan Press, 1957.

Miwa, Jouji, Hiroshi Sasaki and Kazunori Tanno, 'Japanese Spoken Language Learning System Using Java Information Technology', *ICSLIP 2000*, III, 578–581.http://www.isca-speech.org/archive/icslp_2000/i00_3578.html.

Norris, J. and L., Ortega, 'Effectiveness of L2 instruction: A Research Synthesis and Quantitative Meta-Analysis', *Language Learning* 50:3 (2000): 417–528.

O'Brien, Mary Grantham, 'Teaching Pronunciation and Intonation with Computer Technology', in Ducate and Arnold eds, *Calling on CALL: From Theory and Research to New Directions in Foreign Language Teaching.* San Marcos: Computer Assisted Language Instruction Consortium, 2006, 127–148.

Raux, Antoine and Tatsuya Kawahara, 'Optimizing Computer-Assisted Pronunciation Instruction by Selecting Relevant Training Topics', *InSTIL* 2002, http://www.cs.cmu.edu/∼antoine/papers/instil2002.pdf.

Tsurutani, Chiharu, 'Acquisition of Word Prosody by Second Language Learners – A Study of the Acquisition of Japanese Prosodic Features by Speakers of English', unpublished PhD thesis, University of Queensland, 2001.

Tsurutani, Chiharu, 'Speech Rate and the Perception of Geminate Consonants and Long Vowels: A Study of English-speaking Learners of Japanese', *Journal of Japanese Language Teaching*, 119 (2003): 51–64.

Tsurutani, Chiharu, Yutaka Yamauchi and Nobuaki Minematsu, 'Development of a Program for Self Assessment of Japanese Pronunciation by English Learners', Interspeech 2006-ICSLP Ninth International Conference on Spoken Language Processing, 841–844.

②

495

500

505

510

515

520

525

530

535