
**SOON-HIN HEW**  
**MITSURU OHKI**  
*Kyoto University*

**ABSTRACT**

This study examines the effectiveness of imagery and electronic visual feedback in facilitating students’ acquisition of Japanese pronunciation skills. The independent variables, animated graphic annotation (AGA) and immediate visual feedback (IVF) were integrated into a Japanese computer-assisted language learning (JCALL) program focused on the pronunciation of specific word pairs. Students enrolled in an elementary Japanese course in Malaysia were randomly divided into three groups and used three different versions of the program: (a) text + sound, (b) text + audio + AGA, and (c) text + audio + IVF. The results of the study showed that students in the text + audio + AGA group significantly outperformed students in the text + audio only group in pretest-posttest improvement in learning the pronunciation of (a) *akusento* ‘pitch’ words, (b) *seion* ‘voiceless’ and *dakuon* ‘voiced’ words and (c) *tanon* ‘short vowel,’ *chouon* ‘long vowel,’ and *sokuon* ‘geminate consonants’ words. Students in the text + audio + IVF group significantly outperformed students in the text + audio only group in learning *akusento* words and *seion* and *dakuon* words. There were no significant differences between the AGA and IVF groups.

**KEYWORDS**

Animated Graphic Annotation (AGA), Immediate Visual Feedback (IVF), Japanese Computer-assisted Language Learning (JCALL)

**INTRODUCTION**

In learning Japanese, students are generally taught both the written forms and sounds of 46 basic *kana* characters. Traditionally, students learn the *kana* characters and the Japanese words by listening repeatedly to audio files or a teacher. This is a time-consuming and wearisome task. In addition, traditional methods of correcting students’ pronunciation rely on subjective evaluations involving students’ recognition of their own errors in a language lab exercise that guides
them to compare their pronunciation to that of a native speaker on an audiotape (Molholt, 1988). Moreover, most Japanese textbooks or even dictionaries designed for learners do not provide information on intonation and pitch marks. Consequently, students are frequently unable to perceive intonation correctly (Soga & Matsumoto, 1978). Correct pronunciation of Japanese cannot be achieved without learning accent, intonation, tempo, and rhythm of Japanese pronunciation (Mizutani, 1990).

In this paper, we will first summarize previous research on the impact of imagery and the effectiveness of visual feedback on foreign language acquisition. We will then briefly describe the Animated Graphic Annotation (AGA) and Immediate Visual Feedback (IVF) versions of the JCALL program created for this project. Next, we will discuss a comparative study of the effectiveness of AGA and IVF conducted in Malaysia. Finally, we will conclude with a discussion of the limitations of the study, the implications of its findings, and suggestions for further research.

The Effect of Imagery on Language Learning

The effect of visual information (imagery) on the acquisition of pronunciation, particularly within a CALL environment, has not been well explored. However, the effect of visual information on text comprehension and L2 vocabulary acquisition has received wide attention. Research on L2 vocabulary acquisition has shown that foreign words associated with imagery techniques are learned more easily than those without these associations (Kellogg & Howe, 1971; Chun & Plass, 1996). Danan (1992), in a different type of study, found that subtitling helped students learn vocabulary items and attributed their improvement to the dual coding theory of Paivio (1986), which stresses the importance of providing input to both one’s verbal representational system and visual representational system. A study by Omaggio (1979) showed that the use of pictures had a significant improvement on college-level beginner’s reading comprehension in French. Oxford and Crookall (1990) stated that visual imagery helps learners package information more efficiently than just words alone because most learners are more easily able to associate new information to concepts in memory by means of meaningful visual images. A study comparing the effects of pictorial and textual glosses on incidental vocabulary growth for foreign language learners conducted by Kost, Foss, and Lenzini (1999) showed that the combination of text and picture was superior to textual gloss alone or pictorial gloss alone. A study by Snyder and Colon (1988) revealed that students provided with additional audiovisual aids performed significantly better in vocabulary retention than students without such aids. Finally, Hanley, Herron, and Cole (1995) and Al-Seghayer (2001) showed that video clips are more effective organizers than still pictures in language learning situations. Al-Seghayer suggested that video clips build a better video mental image, create greater curiosity leading to increased concentration, and embody an advantageous combination of the modalities of dynamic image, sound, and printed text.
The studies above clearly demonstrate that imagery (still or dynamic) has an impact on text comprehension and L2 vocabulary acquisition. However, there are no widely available studies describing the integration of imagery, especially animated graphic annotations, into CALL-based pronunciation exercises to help foreign language students improve both their perception and production of tone and intonation.

**The Impact of Electronic Visual Feedback on the Acquisition of L2 Pronunciation**

In recent years, the effect of visual feedback on L2 pronunciation acquisition, typically a display of a model pronunciation pattern and that of a student, has received considerable attention. Computer-assisted instruction programs with visual feedback features were originally developed to help the deaf acquire more natural speech (Abberton & Fourcin, 1975; Abberton, Hu, & Fourcin, 1998) but have been increasingly developed for use in the teaching of second and foreign language pronunciation. A number of researchers (Albertson, 1982; de Bot, 1983; Weltens & de Bot, 1984; Molholt, 1988, 1990; Chun, 1989; Rochet, 1990; Stenson, Downing, Smith, & Smith, 1992; Anderson-Hsieh, 1994; Stibbard, 1996) reported positive effects of visual feedback on students' learning of intonation. Stenson et al. (1992) investigated the use of a computer-based visual display for teaching pronunciation in tutorials for international teaching assistants and found that the IBM Speech Viewer provided a valuable supplement to traditional methods of pronunciation practice. Albertson (1982) discussed the use of the Visi-Pitch and commented that Visi-Pitch, with the aid of visual feedback, can be useful in teaching intonation and rhythm patterns as well as most phonemes and allophones for L2 learners, especially older language learners. de Bot (1980, 1983) carried out a study to investigate the impact of visual feedback in Dutch students' learning of English intonation. The results showed that visual feedback combined with auditory feedback was more effective than auditory feedback alone and that the learners exposed to both audio and visual feedback tended to repeat target sentences more often than learners exposed to audio feedback only. Molholt (1988), in a study on teaching segmental phonemes, found the visual feedback was helpful in improving the English pronunciation of Chinese learners. He noted that programs with visual feedback features allow learners to visualize their pronunciation by associating the patterns on the screen with their sound production; thus, this kind of ‘biofeedback’ provided an objective measure that helped students focus their attention on the exact features that needed to be changed to improve their pronunciation.

The visual representation of the articulatory characteristics of sounds and the immediate feedback it provides allows students to constantly monitor their performance as much as they wish to improve their performance. The availability of these features in CALL programs make auditory training a more controlled activity (Rochet, 1990). Compared to the use of traditional tools such as audio-
Animated Graphic Annotations and Immediate Visual Feedback

cassettes which must be rewound for each playback and which deteriorate with repeated use, using the computer to teach tone and intonation enables students to have immediate and unlimited playback capacity (Chun, 1989). Stibbard (1996) showed that a fundamental frequency (f0) analyzer (CSL-Pitch), used by university students in Hong Kong for visual feedback on their English intonation patterns was able to play a valuable role in enhancing learners’ self-monitoring skills. As Albertson (1982), Pennington (1996), and Chun (1989) have affirmed, computer-assisted pronunciation enables users to see an acoustic analysis of their recorded utterances with functions for viewing amplitude, pitch, duration, and frequency range in real time (Lambacher, 1999). Finally, Anderson-Hsieh (1994) has shown that suprasegmentals can be effectively taught through the use of visual feedback on stress, rhythm, and intonation in real time.

The Study

The studies reviewed above demonstrate that the use of dynamic video and still pictures are more effective than instructional modes that lack imagery for teaching pronunciation. However, the efficacy of animated imagery on the acquisition of L2 pronunciation has not been extensively investigated. The few studies that do exist indicate that such feedback indeed has an impact on the acquisition of L2 pronunciation. Chul-Ho, Kawahara, Doshita, and Dantsuji (1997) worked with a computer-aided pronunciation learning system based on classification of the place and manner of articulation. Their system adopted articulatory vowel diagrams with formant frequencies (F1, F2) which were then mapped onto typical tongue positions for vowel sounds and the classification of the place and manner of articulation for consonant sounds.

In order to examine the effectiveness of imagery on L2 pronunciation acquisition, Hew and Ohki (2001) conducted a study with 112 students of Japanese using a Japanese computer-assisted language learning (JCALL) program with AGA. The researchers reported that students who used the JCALL program with AGA showed improvement in listening (but not pronunciation) over those who worked with the program without AGA. This result may have been partly due to the relatively small number of words tested (three pairs of akusento words, four pairs of seion and dakuon words, and three pairs of tanon, chouon, and sokuon words).

Pronunciation has always been a difficult aspect of foreign language learning. Learning to identify and produce the differences between L2 sound contrasts can be a difficult task for learners (Lambacher, 1999). A study carried out by Komaki, Yamada, Tajima, and Choi (2000) demonstrated that Korean and Japanese learners of English had difficulty differentiating among differences in English that were not distinctive in their native language.

Malaysian learners also have similar difficulties in learning Japanese. In order to shed light on these difficulties, we added AGA and IVF features to a
Soon-Hin Hew and Mitsuru Ohki

JCALL program to help students learn the pronunciation of akusento words; seion and dakuon words; and chouon, tanon, and sokuon words.

Japanese word pairs in the akusento category have the same written form except for their pitch pattern such as 橋 (hashi ‘bridge’) and 竹 (hashi ‘chopsticks’). Both ‘bridge and ‘chopsticks’ have the same kana written form, はし, but different accents. If a speaker pronounces hashi with the high-low pitch pattern, the word will be interpreted as ‘chopsticks’ without a context by a speaker of standard Japanese.

The terms seion and dakuon refer to written representation of voiceless and voiced segments, respectively. These terms are supplemented by the term handakuon (explosive). Words containing a vowel that is lengthened are called chouon (long vowel), whereas words containing a vowel that is not lengthened are called tanon (short vowel).

The term sokuon refers to long consonants or, more technically, geminate consonants. Japanese is a mora-based language in which the mora is a unit of pronunciation corresponding to a single vowel, a consonant-vowel combination, the initial beat of a double (geminate) consonant, the second beat of a lengthened vowel, or the nasal consonant N (ん, ン). Each mora corresponds to one beat in the regular staccato rhythm that characterizes the pronunciation of Japanese words or sentences and takes approximately the same amount of time to pronounce as other segments. Therefore, in Japanese, a long consonant is phonetically a single segment such as kk in kokka (national anthem) (Natsuko, 1997).

The pairs of words selected for use in the study described here were checked for authenticity by native Japanese speakers (see Figures 1-3).

Figure 1
Sample Word Pair for akusento (pitch) (アクセント)

Figure 2
Sample Word Pairs for seion (voiceless), dakuon (voiced), and handakuon (explosive)（清音 & 半）濁音）
Animated Graphic Annotations and Immediate Visual Feedback

Figure 3
Sample Word Pairs for tanon (short vowel), chouon (long vowel), and sokuon (geminate consonants) (短音，長音及促音)

As can be seen in Figures 1-3, AGAs are animated graphic symbols designed to help Japanese learners to acquire listening and pronunciation skills. The set of symbols includes , , , , and . For the akusento words in Figure 1, represents low pitch in ame ‘sweets,’ and represents high pitch in ame ‘rain.’ In Figure 2, symbolizes seion (voiceless), symbolizes dakuon (voiced), and symbolizes handakuon (explosive). In the first display in Figure 3, stands for tanon (short vowel) and chouon (long vowel). Finally, in the second display in Figure 3, represents a single consonant, and represents sokuon (geminate consonant). The small dark section at the end of the first symbol for sokuon illustrates the “pause” at the end of the first part of the geminate consonant, and the space between the two illustrates the fact that no sound occurs at this point.

The IVF feature is a user-friendly sound processing tool developed by the researchers for use in teaching pronunciation. It was created using the Tcl/Tk scripting language (see ftp://ftp.scriptics.com/pub/tcl; http://dev.scriptics.com) and an audio programming toolkit (see http://www.speech.kth.se/snack). As in other electronic visual feedback programs, IVF consists of two components: a native sound display (see Figure 4) and a learner sound display (see Figure 5).

Figure 4
Sample Native Speaker Pronunciation of ame (あめ)
Figure 5
Sample Student Pronunciation of *ame* (ー)

The JCALL program described here allows students to listen to the pronunciation of a native speaker and simultaneously view a representation of that pronunciation on the screen. Students click on the record button to record their voice and then view their recording in visual form immediately after clicking on the stop button. They can objectively evaluate their production errors by visually comparing their pronunciation with that of a native speaker. They can play both the model recording and their own recording as many times as they wish.

The study described here was designed to answer two research questions.

1. Do animated graphic annotations (AGAs) and immediate visual feedback (IVF) facilitate elementary Japanese language learners’ pronunciation acquisition of *akusento*, *seion* and *dakuon*, and *tanon*, *chouon*, and *sokuon*?

2. If so, which of the two, AGA or IVF, is more effective in facilitating students’ acquisition of these pronunciation patterns?

**EXPERIMENT**

**Participants**

One hundred and thirty-two students (39 male and 93 female) who were taking a Japanese level 1 course participated in the study. The study was conducted in the Faculty of Modern Languages and Communication at the Universiti Putra Malaysia (UPM) and in the Center for Modern Languages and Communication at the Multimedia University (MMU). The students had studied Japanese an average of 2.64 months before the start of the experiment. The most experienced student had studied Japanese for 26 months (she had taken a Japanese course before she joined the Level 1 course at the university), and the least experienced student had studied the language for only 1 month. Of the 132 participants, 129 (97.73%) had studied Japanese language for only two to three
months at the time the research was conducted. One hundred and six students (80.30%) reported that they were native Chinese speakers, 18 students (3.64%) reported that they were Malay speakers, and 8 (6.06%) declared that they spoke Tamil or other languages such as English. Ninety-five students (71.97%) had taken English language as their second language.

**Procedure**

Three versions of a JCALL program with identical content (same text displays and audio files) were prepared for the project: version one contained text + audio only, version two contained text + audio + AGA, and version three text + audio + IVF (see sample lesson content for each version in Appendix A). The participants were randomly assigned to three groups (see Table 1): students in Group 1 used version one (text + audio), students in Group 2 used version two (text + audio + AGA), and students in Group 3 used version three (text + audio + IVF).

Table 1

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPM</td>
<td>21</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>MMU</td>
<td>23</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td>44</td>
<td>44</td>
</tr>
</tbody>
</table>

All participants took a Japanese pronunciation pretest consisting of 15 Japanese word pairs (see complete test in Appendix B): 5 word pairs for *akusento*, 4 word pairs for *seion* and *dakuon*, and 6 word pairs for *tanon*, *chouon*, and *sokuon*. Students were given 5-10 minutes to read the word pairs and then recorded their pronunciation of the word pairs on DAT Walkman recorders (Sony Model TCD-D8).

After completing the pretest, the appropriate version of the JCALL program was briefly introduced to the students in each group: version one to the students in Group 1, version two to the students in Group 2, and version three to the students in Group 3. The students in each group then used the relevant version of the program to practice their pronunciation of the Japanese word pairs.

Students in Group 1 (the control group) used the audio files on the computer in much the same way that students practice pronunciation exercises with traditional audiotapes. Students in Group 2 clicked on the text, viewed the AGA displayed above the word and listened to the audio file. Students in Group 3 clicked on the text and then the play button to listen to the audio file. Then, they clicked on the record button to record their voice. When they clicked on the stop button, they saw the wave form and pitch contour display of their recording.
Students in all groups listened to the word pairs and recorded their voice to compare their pronunciation to that of the native speaker. They were given 30-45 minutes to complete all the exercises and were able to listen to each word pair and record their voice several times if they wished.

After the students finished the exercises, they completed a pronunciation posttest containing the same word pairs as on the pretest. Also as in the pretest, they recorded their voice on DAT tape. After completing the posttest, the students were asked to complete a questionnaire in which they provided demographic data (e.g., age, sex, race, L1, and L2) and then rated their ability to differentiate between the individual words in the word pairs on a 5-point scale ranging from 1 (very sure) to 5 (not sure) (see questionnaire in Appendix C).

**Data Analysis**

The students’ pronunciation of the word pairs on both the pretest and posttest was evaluated by 10 native speakers of Japanese on a 5-point scale ranging from 1 (very bad) to 5 (very good) (see evaluation form in Appendix D). Most of the evaluators were from the Kanto Area, the area where standard Japanese originated. Each evaluator was given evaluation forms and the CDs on which the students’ pronunciation files were stored. The evaluators’ scores were then submitted to analysis.

**RESULTS**

**Pretest-posttest Improvement**

The pronunciation scores of students in all three groups improved from the pretest to the posttest. The overall mean score for all groups was 2.53 on the pretest and 3.02 on the posttest. Table 2 and Figures 6-8 list the mean scores of participants in the three groups by test section (akusento; seion and dakuon; and tanon, chouon, and sokuon). In the individual categories, the mean score for students in Group 1 (the control group) increased by .49 for akusento, .25 for seion and dakuon, and .50 for tanon, chouon, and sokuon. The improvement in mean scores for students in Group 2 and Group 3 exceeded those of Group 1.
Animated Graphic Annotations and Immediate Visual Feedback

Table 2
Pretest-posttest Mean Scores by Group and Test Section

<table>
<thead>
<tr>
<th></th>
<th>Section 1 Akusento</th>
<th>Section 2 Scion and dakuon</th>
<th>Section 3 Tanon, chouon, and sokouon</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest Group 1 (text + audio only)</td>
<td>2.14</td>
<td>3.25</td>
<td>2.36</td>
<td>2.52</td>
</tr>
<tr>
<td>Group 2 (text + audio + AGA)</td>
<td>2.33</td>
<td>3.23</td>
<td>2.42</td>
<td>2.61</td>
</tr>
<tr>
<td>Group 3 (text + audio + IVF)</td>
<td>2.14</td>
<td>3.12</td>
<td>2.28</td>
<td>2.46</td>
</tr>
<tr>
<td>Posttest Group 1 (text + audio only)</td>
<td>2.63</td>
<td>3.50</td>
<td>2.86</td>
<td>2.95</td>
</tr>
<tr>
<td>Group 2 (text + audio + AGA)</td>
<td>2.90</td>
<td>3.59</td>
<td>2.95</td>
<td>3.11</td>
</tr>
<tr>
<td>Group 3 (text + audio + IVF)</td>
<td>2.86</td>
<td>3.48</td>
<td>2.79</td>
<td>3.00</td>
</tr>
</tbody>
</table>

Figure 6
Pretest-posttest Means for Group 1 by Test Section

![Graph showing pretest-posttest means for Group 1 by test section]
Soon-Hin Hew and Mitsuru Ohki

Figure 7
Pretest-posttest Means for Group 2 by Test Section

![Graph showing pretest-posttest means for Group 2 by test section.]

Figure 8
Pretest-posttest Means for Group 3 by Test Section

![Graph showing pretest-posttest means for Group 3 by test section.]

An independent sample t test was applied to the improvement in mean scores from the pretest to the posttest and showed a significant difference for section 1 between Group 3 and Group 1 ($t(44) = 2.10, p < .05$) but not between Group 3 and Group 2. Analysis of variance was applied to the improvement in mean scores for section 2 and showed significant differences between Group 3 and Group 1 ($F(1,43) = 3.43, p < .05$) and between Group 2 and Group 1 ($F(1,43) = 2.61, p < .05$) but not between Group 3 and Group 2. Analysis of variance was also applied to the improvement in mean scores for section 3 and showed a significant difference between Group 2 and Group 1 ($F(1,43) = 1.92, p < .05$) but not between Group 3 and Group 1 nor between Group 3 and Group 2.
**Animated Graphic Annotations and Immediate Visual Feedback**

**Questionnaire Results**

Analysis of students’ ratings of their ability to distinguish between words in the word pairs showed that their ratings improved dramatically after using the JCALL program (see Table 3).

Table 3  
Students’ Ratings of Their Ability to Distinguish between Words in the Word Pairs

<table>
<thead>
<tr>
<th>Categories of words</th>
<th>Before training</th>
<th>After training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S, VS N NVS NS</td>
<td>S, VS N NVS NS</td>
</tr>
<tr>
<td>Akusento words</td>
<td>3 24 105</td>
<td>86 42 4</td>
</tr>
<tr>
<td>Seion and dakuon words</td>
<td>32 34 66</td>
<td>100 27 5</td>
</tr>
<tr>
<td>Tanon, chouon, and sokuon words</td>
<td>18 37 77</td>
<td>89 38 5</td>
</tr>
</tbody>
</table>

Three out of the 132 students (2.27%) reported that they were sure or very sure they could distinguish between akusento words before using the program, while 86 out of 132 (65.15%) stated that they were sure or very sure they could distinguish between these words after using the program. For the seion and dakuon words, 32 out of 132 (24.24%) were able to distinguish between words before using the program, and 100 out of 132 (75.76%) after using the program. For tanon, chouon and sokuon, 18 out of 132 (13.64%) could distinguish between words before and 89 out of 132 (67.42%) after.

Individual students’ comments corroborated these increases and mentioned specific features of the program that they found particularly helpful.

- Good for beginners and for pronunciation practices.2
- Good for self pace learning.
- The graphic annotations really helped users in understanding the pronunciation. I found this especially true for the tanon, chouon, and sokuon parts.
- IVF really helped us in learning accent.
- It gave me the knowledge on how to catch the proper pronunciation when looking at the visual feedback.
- The AGA symbols really assisted me in learning tanon, sokuon, and chouon.
- Should have a section in which the user can key in the word s/he doesn’t know how to pronounce.
- Mouth movement displays would help me.
- The program should test whether the participant pronounces correctly.
- After we pronounced the word, the computer corrected our mistakes.

**DISCUSSION**

Students in Group 2 (text + audio + AGA) improved their scores from pretest to posttest significantly more than students in Group 1 (text + audio) in all three categories: akusento, seion and dakuon, and tanon, chouon, and sokuon. Students in Group 3 (text + audio + IVF) performed significantly better than those
in Group 1 in the akusento category and in the seion and dakuon category. However, there were no substantial differences in performance between students in Group 2 and Group 3, that is, students generally seemed to benefit equally from the presence of AGA and IVF.

It is clear from the data that the AGA and IVF features had an impact on the learning of Japanese pronunciation. Students who used the program with the AGA feature performed better than Group 1 because the visual image provided them with external information to help them pronounce the words. In the IVF setting, the visual display enabled students to see their sound production in the form of amplitude, pitch, duration, and frequency range. Students in this group repeated words more often and made more effort to correct their pronunciation.

This study’s results support the generative theory of multimedia learning (Mayer, 1997), drawing on Wittrock’s (1990) generative theory and Paivio’s (1986) dual coding theory. Mayer’s theory assumes that learners of a second or foreign language have two separate verbal systems (L1 and L2) and a common imagery system. This theory also suggests that the presentation of explanations in words aided by meaningful visual images or illustrations is effective because this kind of presentation helps learners to build visual and verbal cues for retrieving stored information from memory. The animated imagery in the JCALL program (the AGA symbols) enabled learners of Japanese pronunciation to construct and coordinate visual and verbal representations (printed words, images, and speech sounds) of the same material and helped them to recall pronunciation patterns better than those who worked with printed words and speech sounds only.

Part of this study was based on Hew and Ohki (2001), but the current study differs from the previous study in four ways. First, the participants in the current study were students studying Japanese at the beginning level only versus students from various levels in Hew and Ohki (2001). Second, the number of word pairs tested (five pairs of akusento words, four pairs of seion and dakuon words, and six pairs of tanon, chouon, and sokuon words) and the number of students involved in the current study were higher than those in the previous study. Third, another independent variable, IVF, was added to the current study to determine which form of visual representation, AGA versus IVF, is more effective in helping learners of Japanese. Finally, a larger number of evaluators scored students’ pronunciation in this study than in the earlier study.

CONCLUSION
This study explored the effectiveness of animated graphic annotations and immediate visual feedback for improving Malaysian learners’ pronunciation of Japanese words. The results showed that both animated graphic annotations and immediate visual feedback helped the learners improve their pronunciation, but to differing degrees. These results are not surprising because our ongoing investigation of cognitive styles and abilities indicates that there is no single mode of learning that is helpful to all learners. Learners learn most successfully...
Animated Graphic Annotations and Immediate Visual Feedback

with the type of information that is best suited to their cognitive style (Knight, 1994). Among the possible factors that may explain the results reported here are that animated graphic images help learners build mental images enabling them to associate verbal representations with meaningful visual images and thereby learn more effectively. In addition, immediate visual feedback providing information on pitch, duration, and frequency range with respect to specific errors in learners’ pronunciation tended to encourage the learners to repeat the target words more often and make a greater effort to correct their pronunciation. We hope that these results are viewed as a starting point for other researchers to further explore the use of graphic annotations and visual feedback, separately or together, in CALL programs.

There were some limitations to this study. First, the use of a posttest immediately following the use of a single program for 30-45 minutes measured only short-term results. It would be interesting to investigate long-term results after more extensive use of such programs. Second, this kind of study should be replicated with students at various levels of proficiency and L1 backgrounds in various kinds of academic settings. Third, additional animated graphic symbols should be created and evaluated for their effectiveness in helping students improve their pronunciation in Japanese and other languages. Finally, we plan to integrate speech recognition technology with animated graphic annotations and immediate visual feedback in JCALL programs for use in future research projects. A recent report aimed at integrating speech recognition technology into a pronunciation training system shows that speech recognition technology can be used to considerable advantage to improve the quality and detail of feedback given to students (Menzel, Herron, Morton, Pezzota, Bonaventura, & Howarth, 2001). We believe that the combination of AGA, IVF, and speech recognition technology will go far in enhancing pronunciation training in second/foreign language learning.

NOTES

1 The overall means were computed by dividing the sum of the means of the three sections of the test by 3 (e.g., \(\frac{2.95 + 3.11 + 3.00}{3} = 3.02\)).

2 Students’ grammatical errors have been corrected.

REFERENCES


Animated Graphic Annotations and Immediate Visual Feedback


APPENDIX A

Sample Lesson Content

Group 1 (text + audio)

**Akusento section**

**Seion and dakuon section**

Group 2 (text + audio + AGA)

**Akusento section**

**Seion and dakuon section**

**Tanon, chouon, and sokuon section**

**Tanon, chouon, and sokuon section**
Animated Graphic Annotations and Immediate Visual Feedback

Group 3 (text + audio + IVF)

Akusento section

Native pronunciation display

Learner pronunciation display

Seion and dakuon section

Native pronunciation display

Learner pronunciation display

Tanon, chouon, and sokuon section

Native pronunciation display

Learner pronunciation display

Native pronunciation display

Learner pronunciation display
APPENDIX B

Pronunciation Test
Instructions: Please pronounce the following words clearly

Section 1: **akusento** (アクセント)
1. A. かみ (神) ‘God’  B. かみ (紙) ‘paper’
2. A. はし (橋) ‘bridge’  B. はし (箸) ‘chopsticks’
3. A. おもい (思い) ‘thought’  B. おもい (重い) ‘heavy’
4. A. あめ (雨) ‘rain’  B. あめ (飴) ‘sweets’
5. A. もも (桃) ‘peach’  B. もも (股) ‘thigh’

Section 2: **seion** (清音) and **dakuon** (濁音)
1. A. にし (西) ‘west’  B. にじ (虹) ‘rainbow’
2. A. きん (金) ‘gold’  B. ぎん (銀) ‘silver’
3. A. てんき (天気) ‘weather’  B. てんき (電気) ‘electricity’

Section 3: **tanon** (短音), **chouon** (長音), and **sokuon** (促音)
1. A. なす (茄子) ‘eggplant’  B. なあず (ナース) ‘nurse’
2. A. いん (印) ‘seal’  B. いいん (医院) ‘medical clinic’
3. A. びる (ビル) ‘building’  B. びいる (ビール) ‘beer’
5. A. にし (西) ‘west’  B. にっし (日誌) ‘journal’

APPENDIX C

QUESTIONNAIRE

Section 1: Background
1. Name:
2. Gender: Male ( ) Female ( )
3. Race: Malay ( ) Chinese ( ) Indian ( ) others ( )
4. Native language:
   Malay language ( )
   Chinese language ( )
   Tamil language ( )
   English language ( )
   Others ( )
Section 2: Japanese Language studies

1. How long have you studied Japanese? (       ) year(s) (       ) month(s)

2. How many hours do you take Japanese language classes in a week? (       ) hour/week

3. What kind of media have you used in learning Japanese language?
   A. Textbook (       )
   B. Audio tape (       )
   C. Video tape (       )
   D. CD-ROM/CALL courseware (       )
   E. Internet (       )
   F. Others (       )

4. Can you differentiate between the pronunciation of the *akusento* words (アクセント) in standard Japanese?
   Examples:
   a) あめ (雨) ‘sweets’ あめ (雨) ‘rain’
   b) はし (橋) ‘chopsticks’ はし (橋) ‘bridge’
   c) かみ (神) ‘paper’ かみ (神) ‘God’

   Before Training:
   Very sure ( ) Sure ( ) Neutral ( ) Not very sure ( ) Not sure ( )

   After Training:
   Very sure ( ) Sure ( ) Neutral ( ) Not very sure ( ) Not sure ( )
5. Can you differentiate among the pronunciation of *tanon*, *chouon*, and *sokuon* words *(短音, 長音 & 促音)*?

**Examples:**

a) おばさん ‘aunt’ おばあさん ‘grandmother’
b) ビル ‘building’ ビール ‘beer’
c) こうか ‘school song’ こっか ‘national anthem’

**Before Training:**

- Very sure ( )
- Sure ( )
- Neutral ( )
- Not very sure ( )
- Not sure ( )

**After Training:**

- Very sure ( )
- Sure ( )
- Neutral ( )
- Not very sure ( )
- Not sure ( )

6. Can you differentiate between the pronunciation of *seion* *(清音)* and *dakuon* *(濁音)* words?

**Examples:**

a) きん ‘gold’ ぎん ‘silver’
b) てんき ‘weather’ てんき ‘electricity’
c) かんだん ‘simple’ がんだん ‘New Year’s Day’

**Before Training:**

- Very sure ( )
- Sure ( )
- Neutral ( )
- Not very sure ( )
- Not sure ( )

**After Training:**

- Very sure ( )
- Sure ( )
- Neutral ( )
- Not very sure ( )
- Not sure ( )

7. Comments (if any):

*Terima kasih*

Thank you very much

*ありがとうございます*
APPENDIX D

Evaluation Form
1 = very bad, 2 = bad, 3 = average, 4 = good, 5 = very good

<table>
<thead>
<tr>
<th>Pretest Q1</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posttest Q16</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q17</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q18</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q19</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q5</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q20</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q21</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q7</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q22</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q8</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q23</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q9</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q24</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q10</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q25</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q11</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q26</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q12</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q27</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q13</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q28</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q14</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q29</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Pretest Q15</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Posttest Q30</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
AUTHORS’ BIODATA

Soon-Hin Hew received his B.Sc. (1995) in the Faculty of Science, Mathematics, and Computers and his M.Sc. (1997) in the Faculty of Technology and Information Science at the National University of Malaysia (UKM). In July, 1997, he joined Faculty of Creative Multimedia at the Multimedia University as a Lecturer. He was a member of the Japan Association for Language and Teaching (JALT) and the CALL SIG. He is currently a Ph.D. candidate in the Department of Information and Cognition, Graduate School of Human and Environmental Studies, Kyoto University, Japan. His research interests are in CAI, CALL instructional design, and applied technology in education and multimedia.

Mitsuru Ohki received his M.A. degree in Linguistics from the Osaka University of Foreign Studies in 1972 and his Ph.D. in the Department of Linguistics, University of Paris in 1978. He is currently professor in the Faculty of Integrated Human Studies and the Graduate School of Human and Environmental Studies at Kyoto University, Japan. His research interests include French linguistics, French education, and CALL.

AUTHORS’ ADDRESSES

Soon-Hin Hew
Temporary address:
Department of Information and Cognition
Graduate School of Human and Environmental Studies
Kyoto University
Yoshida-Nihonmatsu-cho
Sakyo-ku, Kyoto-shi,
Kyoto 606-8501
Japan
Email: my711010@hotmail.com

Permanent address:
Faculty of Creative Multimedia (FCM)
Multimedia University (MMU)
63100 Cyberjaya,
Selangor, Malaysia.

Mitsuru Ohki
Department of Information and Cognition
Graduate School of Human and Environmental Studies
Kyoto University
Yoshida-Nihonmatsu-cho
Sakyo-ku, Kyoto-shi,
Kyoto 606-8501
Japan