DESIGN OF A COMPUTER-ASSISTED SPEECHREADING TRAINING SYSTEM
FOR JAPANESE

by

MASAKO FUJIU
B.A., International Christian University, Japan, 1982
A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

in
THE FACULTY OF GRADUATE STUDIES
SCHOOL OF AUDIOLOGY AND SPEECH SCIENCES

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
June, 1989
© Masako Fujiu, 1989
In presenting this thesis in partial fulfilment of the requirements for an advanced
degree at the University of British Columbia, I agree that the Library shall make it
freely available for reference and study. I further agree that permission for extensive
copying of this thesis for scholarly purposes may be granted by the head of my
department or by his or her representatives. It is understood that copying or
publication of this thesis for financial gain shall not be allowed without my written
permission.

Department of Audiology & Speech Sciences

The University of British Columbia
Vancouver, Canada

Date May 29, 1989
ABSTRACT

Speechreading is known to play an important role in speech perception, particularly for hearing-impaired individuals. Computer-Aided Speechreading Training (CAST) systems have been implemented for English and French, to train adults with mild-to-moderate hearing loss to improve their communicative ability with an effective use of visual speech information. Recent tests have shown that the CAST system is effective in improving speechreading performance. The purpose of this study is to design a similar system for Japanese, which will be called CAST-J.

CAST-J follows the basic principles used for the English and French versions of CAST: (1) it uses the method of discourse tracking, in which training is focused on the perception of ongoing speech; (2) it aims at training subjects to integrate visual information, auditory information, and linguistic redundancy of the message; (3) it is designed for post-lingually hearing-impaired adults. Adults with cochlear implants are also considered as a possible target population for CAST-J, as long as they use auditory information of speech as a primary source in speech perception.

The lessons of CAST-J are organized around the concept of viseme. Fifteen visemes of Japanese have been determined for this study, more specifically, five vowel visemes: \{ i \}, \{ e \}, \{ a \}, \{ o \}, and \{ u \}, seven consonant visemes: \{ p, b, m \}, \{ w, \phi \}, \{ r \}, \{ s, z, t, d, n, ts, dz \}, \{ k, g, h \}, \{ f \}, and \{ j, ç, ŋ, tʃ, dʒ, n \}, and three haku visemes: \{ N \}, \{ Q \}, and \{ R \}. The four sections of each lesson are designed (1) to review previously taught visemes; (2) to practice the recognition of a new viseme; (3) to practice new and old visemes by the tracking method; and (4) to recap the lesson. The lessons progress from easy-to-recognize visemes to difficult-to-recognize visemes. There are sixteen lessons, and two extra lessons are available for pre-and post-training comparison.
The hardware required by CAST-J includes a microcomputer, a video disc player, a high resolution video monitor, and a printer. The system allows the instructor (an audiologist) to select the appropriate speaking rate for the video (fast or slow), whether the trainee will practice with or without audio signal, and to skip if wanted the optional lessons or part of a lesson, depending upon the level of the trainee.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>viii</td>
</tr>
<tr>
<td>NOTATION</td>
<td>ix</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. SPEECHREADING IN AURAL REHABILITATION: STATE OF THE ART</td>
<td>4</td>
</tr>
<tr>
<td>2.1. Speechreading and speech perception</td>
<td>4</td>
</tr>
<tr>
<td>2.2. Speechreading: evaluation and training</td>
<td>7</td>
</tr>
<tr>
<td>2.3. Speechreading and auditory rehabilitation in Japan</td>
<td>9</td>
</tr>
<tr>
<td>2.4. Speechreading for the individuals with cochlear implants</td>
<td>10</td>
</tr>
<tr>
<td>3. TARGET POPULATION</td>
<td>13</td>
</tr>
<tr>
<td>4. VISEMES</td>
<td>17</td>
</tr>
<tr>
<td>4.1. Concept of viseme in general</td>
<td>17</td>
</tr>
<tr>
<td>4.2. Visemes for Japanese</td>
<td>22</td>
</tr>
<tr>
<td>4.2.1. Phonological and phonetic characteristics of Japanese</td>
<td>22</td>
</tr>
<tr>
<td>4.2.2. Significance of haku (mora) in Japanese</td>
<td>27</td>
</tr>
<tr>
<td>4.2.3. Determination of visemes of Japanese</td>
<td>33</td>
</tr>
<tr>
<td>5. EQUIPMENT AND INPUT/FEEDBACK MODE</td>
<td>44</td>
</tr>
<tr>
<td>5.1. Equipment</td>
<td>44</td>
</tr>
<tr>
<td>5.2. Input and feedback modes</td>
<td>45</td>
</tr>
<tr>
<td>5.2.1. Kanji, kana, and romaji</td>
<td>45</td>
</tr>
<tr>
<td>5.2.2. Input</td>
<td>47</td>
</tr>
<tr>
<td>Table</td>
<td>Legend</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td>I</td>
<td>Phonemes of Japanese and Their Allophones/Phones</td>
</tr>
<tr>
<td>II</td>
<td>List of Hakus Written in Romaji</td>
</tr>
<tr>
<td>III</td>
<td>List of Hakus Written in Phonetic Symbols</td>
</tr>
<tr>
<td>IV</td>
<td>Examples of Loanwords Widely Used in Japanese</td>
</tr>
<tr>
<td>V</td>
<td>Lesson-by-Lesson Inventory of Viseme Tokens</td>
</tr>
<tr>
<td>VI</td>
<td>Viseme Inventory for Lesson 17 and Lesson 18</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Legend</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Functional Diagram of the CAST-J System</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Writing System Used at Different Stages and Conversion Rules</td>
<td>52</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENT

I would like to express my appreciation to all those people who have contributed to this work:

- Dr. André-Pierre Benguerel, for his considerate assistance and for the fine model he has provided as a teacher.

- Noelle Lamb, for her valuable advice and assistance.

- Kathy Pichora-Fuller, for her advice as the originator of CAST.

- John Nicol, for sharing his expertise in computers.

- Yumiko Fukuda and Ikuko Seki, for their advice regarding speechreading in Japanese.

- My classmates, for their friendship.

- My family and friends in Japan, for their support.
NOTATION

In this study, some punctuation marks and print faces are used for a specific situation, which is indicated on the right of each symbol.

/ / denotes a phoneme
[ ] denotes a(n) (allo)phone
\ \ denotes a viseme
< > denotes a(n) (allo)phone
I I denotes a grapheme
<< >> denotes a(n) (allo)graph
{ } denotes a set of phonemes, visemes, etc.
" " denotes a word with a special meaning
" " denotes a quotation
@ stands for "in the context of"

**bold type** denotes emphasis of an important word

**small capitals** denotes a translation

*italics* denotes a romaji transliteration

**bold type italics** denotes a kana

! denotes separation between kana characters
CHAPTER 1

INTRODUCTION

Computer-assisted devices for the communicatively impaired have been successfully used both in clinical and educational settings. Aural rehabilitation is one of the areas which can be refined by utilizing such devices. A Computer-Aided Speechreading Training system (CAST) has been developed by Pichora-Fuller and her colleagues as a clinical and research tool to train and assess the speechreading abilities of adults with a mild-to-moderate acquired hearing loss (Pichora-Fuller and Cicchelli, 1986; Benguerel and Pichora-Fuller, 1988; Pichora-Fuller and Benguerel, 1989). At present, English and French versions of CAST have been implemented. The purpose of this study is to outline the design and the lessons of a computer-aided speechreading training programme for Japanese (CAST-J).

The CAST-J system provides an interactive environment for speechreading practice, by means of a microcomputer connected to a video disc player and a high resolution video monitor, as shown in Figure 1. The trainee views a video-recorded speaker talking on screen. He/she is given a target string (usually a phrase) that he/she has to speechread, either with or without auditory signal. Presentation of the target strings by the speaker is controlled by the computer. The trainee inputs his/her response to the computer via a keyboard. The response, echoed on screen, is analyzed by the computer, and feedback is provided by the computer on screen, in written form. If the trainee wishes to re-play the target string, or to move to another location in the lesson, he/she is able to do so by using various command keys. Although the speechreading trainer is the CAST-J system itself (both instruction and feedback are computer-controlled), the trainee still plays an active role. The data of the trainee’s
performance are recorded, tabulated and analyzed, and they are printed out for further analysis. The detail of the system will be described in Chapters 5 and 6.

This study includes a discussion of the current role of speechreading in aural rehabilitation, a survey of the linguistic characteristics of Japanese, particularly where relevant to speechreading, and a set of lessons. The actual video-recording of the lessons and training of the subjects are not included in the present study.
Figure 1. Functional Diagram of the CAST System.
CHAPTER 2

SPEECHREADING IN AURAL REHABILITATION: STATE OF THE ART

2.1. Speechreading and speech perception

Speechreading, a term broader than lipreading, refers to the "recognition of a speaker's words by watching his lips, facial expressions, gestures, etc., as well as using closure, rules of language, contextual cues, etc." (Sheeley, 1985, p. 1087). Besides the movements of the visible articulators (e.g. the range of jaw opening, the shape of the upper and lower lips, the appearance of the teeth, lingual movements (Fukuda, 1987), the receiver integrates various environmental and contextual cues.

In general, speechreading itself is not sufficient for the interpretation of speech. However, it is now regarded as playing an important role even in normal speech perception, where it is integrated with auditory information (Massaro, 1987). The McGurk effect is a good example of this integration: people rely on visual cues even when auditory information is intact (McGurk and MacDonald, 1976). Nevertheless, speechreading has a greater importance in disambiguating speech when auditory information is degraded (Sumby and Pollack, 1954). Middelweerd and Plomp (1987) examined the effects of speechreading on speech-reception thresholds in noise. They found an improvement of approximately 4 dB, both in normal hearing subjects and in presbycusic subjects. Speechreading and auditory information are known to work synergistically in speech perception by the hearing impaired. Manner of articulation and voicing contrast are known to be better perceived through the auditory channel, while place of articulation fares better through the visual channel. (Erber, 1972; Binnie, Jackson and Montgomery, 1976; Fukuda, Sakamoto and Kuroki, 1976; Fukuda and Hiki, 1977). Information available through speechreading, however, is not sufficient when
place of articulation is in the middle or back of the oral cavity (McGurk and MacDonald, 1976).

Recent studies in normal hearing subjects show that speechreading can be improved when supplemented with various types of auditory information such as presentation of the acoustic signal, low-pass or band-pass filtered version of the signal (Breeuwer and Plomp, 1984), selected formant information (Breeuwer and Plomp, 1985), and/or voice fundamental frequency ($F_0$) information (Breeuwer and Plomp, 1986; Boothroyd, Hnath-Chisolm, Hanin and Kishon-Rabin, 1988). In the study of hearing impaired subjects, Grant (1987) reported that improvement of speechreading was observed with presentation of $F_0$.

Presentation of speech information to the cochlea, as an electrical rather than an acoustic signal (e.g. in cochlear implants), is another technique used to enhance speech perception. The significance of the use of cochlear implants in conjunction with speechreading will be discussed separately in Section 2.4.

As pointed out by Massaro (1987), speechreading has a great potential in enhancing speech perception when it is combined not only with auditory information but also with other sensory information. One way of compensating for the limitations of speechreading is the manual presentation of speech cues. Manual systems developed in the past include the Mouth-Hand System developed by Forechhammer in Denmark (Børrild, 1967), Cued Speech originated by Cornett (1970), and Q-Code proposed by Fant (1970) in Sweden. Another means of supplementing speech perception is visualization of voicing and manner of articulation by pulsing lights: the speechreader wears glasses which have several small lamps on the lenses (Upton, 1968), or the speaker wears glasses to which a ring-shaped frame with lamps is attached (Traunmüller, 1976, cited in Fukuda, 1987). Both devices provide the less visible part of speech information (i.e. voicing and manner of articulation) after it has been extracted from the speech signal by an electrical circuit.

Tactile presentation of speech information has also been investigated by some
researchers. Speech information which has been made available through tactile cues includes F₀ (Hanin, Boothroyd and Hnath-Chisolm, 1988; Hnath-Chisolm and Kishon-Rabin, 1988), resonance, envelope, and voicing contrast (De Filippo, 1984). In the studies of Hanin et al. (1988) and of Hnath-Chisolm and Kishon-Rabin (1988), the authors used two types of tactile display for F₀: a single-channel (temporal) display and a multichannel (spatial) display. In the former case, F₀ is coded as rate of vibratory stimulation and is presented to the pad of the finger. In the latter case, F₀ is coded as a locus of stimulation and as a frequency of vibratory stimulation, and it is presented to the skin of the forearm. Hnath-Chisolm and Kishon-Rabin (1988) examined the effects of the two different types of display on the perception of initial consonant voicing, final consonant voicing, pitch change, and word stress in normal hearing subjects. The results indicate that the tactile presentation of F₀ enhances the perception of speech pattern contrasts. The multichannel display is reported to be more effective than the single-channel display, at least for the perceptual contrast of pitch rise versus pitch fall. Hanin et al. (1988) examined the effects of the tactile presentation of F₀ on the perception of sentences, both in normal hearing and hearing impaired subjects. Only the multichannel display was used for the hearing impaired, while both displays were used for the normal hearing subjects. Perception of sentences improved in all three conditions.

De Filippo (1984) evaluated the spatial presentation of speech information (resonance, voicing, manner of articulation) to fingers and the temporal presentation of speech information (including spectral information) to the hand, the wrist, and the skin below the sternum. The subjects were adults who were congenitally deaf and adults with normal hearing who were made artificially deaf. All the subjects demonstrated an improvement of speech perception when speechreading and tactile information were combined. Although no significant difference was found between the two types of tactile presentation, De Filippo suggests that the temporal display with spectral information seems more efficient than the spatial display.
2.2. Speechreading: evaluation and training

Currently, several tests for speechreading evaluation are available for English. Those tests evaluate the perception of vowels and consonants, words, sentences, and connected speech. The perception of vowels and consonants is evaluated in the lipreading screening test designed by Binnie et al. (1976). Tests of word perception include the use of PB words (Binnie, 1976), subtests of the Craig Lipreading Inventory developed by Craig (1964), and the Diagnostic Test of Speechreading, by Myklebust and Neyhus (1971). Sentence tests include the one developed by Utley (1946), the Craig Lipreading Inventory (1964), the Diagnostic Test of Speechreading (Myklebust and Neyhus, 1971), the CID Everyday Sentences Test (Jeffers and Barley, 1971; Johnson, 1976), the Jacobs Test (Sims and Jacobs, 1976), and the Lipreading Discrimination Test (Bement, Wallber, De Filippo, Bochner and Garrison, 1988). Comprehension of connected speech is most frequently assessed by the 'discourse tracking' described by De Filippo and Scott (1978). In addition, standardization of a test for cochlear implant recipients has been implemented by Spitzer, Leder, Milner, Flevaris-Phillips and Giolas (1987).

Discourse tracking can be used for speechreading training as well. It consists of a procedure in which a 'receiver' (the trainee) repeats groups of words that are read or said by the 'speaker' (the trainer). In this procedure, the receiver is provided with immediate feedback by the speaker. The reception of ongoing speech is measured by calculating the number of words correctly repeated per minute. This method is one of the most popular ones, and applications to various populations of hearing impaired subjects are found in the literature (Lesner and Kricos, 1987; Owens and Raggio, 1987).

Fenn and Smith (1987) suggested some modifications to the scoring system of the tracking method, claiming that despite the potential that the method has, the traditional tracking method has not been fully used for clinical purposes, due to some practical problems. They added a measure of fluency calculated as follows: the number of words correctly repeated per
minute is corrected by deducting 'penalty points', the number of unsuccessful attempts the subject made in identifying a word. When the subject fails to identify a word, the speaker is allowed to repeat the word a maximum of two times. If the subject does not succeed after these two trials, the word is written out and tracking proceeds. It has been shown by Fenn and Smith (1987) that the subjects are unlikely to identify the word correctly after more than two repetitions. A similar result has been shown by Gagné and Wyllie (1989): simple repetition of a test word did not help the subjects to identify the word. Therefore, besides being a valid parameter for assessing speechreading ability, this new scoring method can reduce the frustration of a poor speechreader who frequently has to spend so much time for one word.

Van Uden (1983) proposes the use of visual monitoring of the learner's own speech via video replay for speechreading training: the subject sees himself speaking on the video playback and lipreads himself. In Van Uden's study of prelingually deaf children, this method was found to be particularly useful for those who had not shown progress in the previous training phase.

The use of computers is another way to evaluate and train for speechreading. About a decade ago, the first computer-assisted video tape system for speechreading instruction was reported by a group of researchers at the National Technical Institute for the Deaf (NTID) as part of a computer-assisted instruction project. The system, named DAVID (Dynamic Audio Video Instructional Device), has been used successfully in the language laboratory situation, by NTID students who had a severe-to-profound hearing impairment. (Sims, Von Felst, Dowaliby, Hutchinson and Myers, 1979; Sims, Scott and Myers, 1982).

While DAVID was designed for young adults with severe-to-profound hearing impairment, CAST (Pichora-Fuller and Cicchelli, 1986; Benguerel and Pichora-Fuller, 1988; Pichora-Fuller and Benguerel, 1989) is a system aiming at the pre-retirement adult with a mild to moderate hearing loss. The CAST lessons were constructed in accordance with the
principles of the tracking method: face to face interaction, with feedback between sender (in this case the video tape) and receiver (trainee). In the investigation regarding the effectiveness of the CAST program, a greater improvement of speechreading ability was found in the experimental group, i.e. eight normal-hearing adults who had been trained with CAST, than in the control group, i.e. eight normal-hearing adults who had not had any speechreading training (Gagné, Dinnon and Parsons, 1989).

2.3. Speechreading and auditory rehabilitation in Japan

In Japan, auditory rehabilitation/habilitation has put its major emphasis on children with congenital or acquired hearing loss, for whom the acquisition of language is the primary concern. Generally, maximum use of residual hearing through amplification is the first goal of any approach involving children. Speechreading, like manual communication methods such as fingerspelling, sign language, or cued speech, is considered as a supplemental medium for auditory speech perception. In contrast to aural (re)habilitation for children, that for adults with an acquired hearing loss tends to put more emphasis on speechreading. Unfortunately, however, the topic of speechreading is given extremely limited space in the literature (Tanaka, 1980).

As far as the development of assessment materials is concerned, very few studies are found in the literature. Training and evaluation of speechreading, in general, is totally dependent on the protocols of individual clinicians. In the 1970s, Tanaka and his associates developed a video system to assess speechreading ability with the goal of using the assessment results for an adequate hearing aid fitting (Tanaka, Shindo and Motomiya, 1973; Tanaka and Chiba, 1975). Despite the potential of the system for assessing adults, it was designed primarily for hearing-impaired children. Recently, a video system has been developed by Fukuda (1987) for assessing the speechreading ability of adults. However, it is specifically designed for adults with cochlear implants, and no application to other populations has been
attempted.

2.4. Speechreading for the individuals with cochlear implants

A change in attitude towards speechreading in Japan has occurred in the mid-1980s with the development of cochlear implantation for post-lingually deaf adults. With cochlear implants, neurons of the auditory nerve are stimulated electrically through electrodes surgically implanted in the cochlea. These cochlear implants provide an auditory input to profoundly hearing impaired individuals who cannot benefit from conventional amplification, medication, and/or surgical treatment. The data reported throughout the world thus far appear to indicate that cochlear implants are a promising tool in aural rehabilitation. Presently, there exist several design options: unipolar versus bipolar electrode systems, single versus multichannel electrode systems, solitary versus multiple simultaneous electrode stimulation, transcutaneous versus percutaneous conduction, and various sound processing strategies. Common to all designs are the surgical implantation of the stimulating electrode(s) and the externally worn microphone and sound processor.

Past research has shown that, in general, multichannel cochlear implants provide better speech recognition than single channel models, at least in post-lingually deaf adults (Eddington, 1980; Engelmann, Waterfall and Hough, 1981; Tyler, Tye-Murray, Preece, Gantz and McCabe, 1987; Youngblood and Robinson, 1988), although a successful case has also been reported with the single channel device (Tyler, 1988). Furthermore, continuing improvement in modern technology has allowed the use of implants in cases which had been considered contraindicated for implantation (Balkany, Gantz and Nadol, 1988). At the same time, however, the risk of possible medical complications should not be ignored (Cohen, Hoffman and Stroschein, 1988).

Overall, improvement of hearing in the patients with cochlear implants appears to be encouraging: some patients have been reported to acquire the ability to use the telephone
(Parkin and Stewart, 1988; Youngblood and Robinson, 1988); hearing with a cochlear implant is described as superior to hearing with a hearing aid by patients who have undergone implantation (Youngblood and Robinson, 1988). It has also been reported that some people do well in open-set speech recognition tests, without lipreading (Gantz, McCabe, Tyler and Preece, 1987; Parkin and Stewart, 1988). Nevertheless, the majority of patients benefit most from the combination of auditory and visual inputs (Funasaka, Hayashibara, Takahashi, Yukawa and Hatsushika, 1987; Funasaka and Shiroma, 1987; Fukuda, Shiroma and Funasaka, 1988; Parkin and Stewart, 1988).

In Japan, the first 22-channel cochlear implant was implanted by Funasaka, Hosoya, Hayashibara, Takahashi, Yukawa and Hatsushika (1986). In subsequent studies by Funasaka and his colleagues, the importance of speechreading in the whole rehabilitation process was emphasized (Funasaka et al., 1986, 1987; Funasaka and Shiroma, 1987; Shiroma and Funasaka, 1988).

The patients who have undergone cochlear implantation in Japan are all adults with an acquired hearing loss. It is suggested that the ability to lipread be one of the criteria for selecting the candidates for implantation. The data for six subjects studied by Fukuda et al. (1988) exhibited enhanced speech perception three months post-operatively, when auditory and visual cues were combined, as compared with auditory cues only or visual cues only. The mean correct response scores for the above three conditions were respectively 55%, 20%, 16% for words, and 69%, 23%, 8% for sentences. Correlation coefficients indicated also the dominance of auditory perception in overall speech recognition (V + A and A: 0.72, V + A and V: 0.48). Thus, post-lingually deaf adults with cochlear implants use auditory input as a primary source of speech information, and they can increase their communicative ability to a great extent by using supplementary visual cues.

Currently, cochlear implants are becoming an important option in the aural (re)habilitation of children with no or only limited residual hearing; however, no attempt to
perform implantation in children has yet been made in Japan. There are some data indicating the successful application of implantation of young children outside Japan (Eisenberg et al., 1983; Lyxford, House, Hough, Tonokawa, Berliner and Martin, 1988). Considering the importance of educational and developmental issues, however, this option needs to be assessed with special caution (Tyler, Davis and Lansing, 1987; Boothroyd, 1988; NIH, 1988; Simmons, 1988).
CHAPTER 3

TARGET POPULATION

The original version of CAST was developed specifically for pre-retirement adults with acquired mild-to-moderate hearing loss (Pichora-Fuller and Cicchelli, 1986). The premises of their approach are that speechreading is "a communication task which entails three skills: visual speech perception, use of linguistic redundancy, and effective use of feedback between message sender and receiver" (p. 11). The aim of CAST is to improve speechreading, a supplement to auditory perception of speech which can be affected by various environmental factors such as background noise, speaking rate, familiarity with the topic, etc. The potential users of CAST, therefore, are those who use the auditory channel as the primary means of communication, and the visual channel as a supplemental one. Like the English and the French versions of CAST, CAST-J will aim at the post-lingually hearing impaired adults. The trainees will be expected to have a mild-to-moderate hearing loss and post-secondary-school level linguistic competency.

Children are not included in the target population for this particular version of CAST, for the following reasons. First, the program for pre-lingually hearing impaired children must consider the course of cognitive, phonological, and other linguistic development of children. CAST-J, which is designed for those who can use linguistic redundancy effectively, is thus not ideally suited for pre-lingually deaf children. Second, Luxell and Ronnberg (1987) report that skilled guessing is critical for two speechreading conditions: (1) when messages have little contextual information, and (2) when longer messages are to be speechread. Third, the study by Berger (1972) indicates that a person with an early onset of deafness tends to be poorer at speechreading. Conrad (1979) reports a positive correlation between speechreading ability
and intelligence in pre-lingually deaf children: the more intelligent a child, the better his use of linguistic knowledge for speechreading. Conrad also points out the importance of establishing a phonological code in memory as well as acquiring other linguistic skills such as reading. Campbell (1987, p. 225) claims that speechreading is "a component set of functions that are language dependent and are localized in the left hemisphere". Fourth, CAST-J is inappropriate for post-lingual children in terms of word familiarity. If the linguistic level of a child is advanced enough, a separate CAST-J version using children's vocabulary could be useful, but since this study is only the first stage of a Japanese program, lessons specifically designed for children will not be considered.

CAST-J will consider cochlear implant patients as primary potential trainees for three reasons: (1) ability to lipread is considered to be an important criterion for potential candidates to the surgery, (2) adults who have undergone a cochlear implantation have been found to have relatively good auditory perception of speech, thus they can use speechreading as a supplementary medium for speech perception, and (3) past research has shown that these patients greatly benefit from speechreading training (Funasaka, Hayashibara, Takahashi, Yukawa and Hatsushika, 1987; Funasaka and Shiroma, 1987; Fukuda, Shiroma and Funasaka, 1988).

Agreement on the above points is not unanimous though. Lesner, Sandridge and Kricos (1987), for example, investigated the effects of an analytic approach for speechreading, i.e. the effects of emphasizing identification and discrimination of visemes at the segmental level on sentence-level speechreading in normal hearing subjects. When the recognition of consonant visemes was assessed in VCV utterances without auditory input, the subjects having had analytic training performed better than those with no training. However, when the improvement of speechreading was obtained from the pre- and the post-training scores, the results did not show any significant difference between the subjects whose training included feedback and those whose training did not include feedback, whether their responses were correct or not. When the recognition of sentences was assessed without auditory input, there
was no significant difference in performance between the three groups of subjects tested, namely those who had been trained with feedback, those who had been trained without feedback, and those without training.

Based on these results, Lesner et al. (1987) concluded that (1) the analytic approach for training speechreading is not useful, at least for those who use visual information of speech as the primary means of recognizing connected speech; (2) the ability to recognize visemes at the segmental level was not reflected in the recognition of sentences when the subjects totally relied on the visual information of speech; (3) the analytic approach may be only effective for those who can use auditory input as the primary source of speech information and visual information as the secondary one, as reported by Walden, Prosek, Montgomery, Scherr and Jones (1977) and Walden, Erdman, Montgomery, Schwartz and Prosek (1981).

Lesner et al. (1987) contend that individuals with cochlear implants perceive speech by relying primarily on speechreading and use the auditory channel for supplemental purposes; therefore, they should not be candidates for training which places the emphasis on contrast of visemes. The present author believes that individuals with cochlear implants are still potential candidates for training with CAST-J for the following reasons. Firstly, recent research reveals an improvement of the auditory perception of speech by these subjects and the possibility of their using audition as the primary channel for speech perception. Secondly, CAST-J does indeed have some analytic aspects in its approach (i.e. lessons include recognition of visemes at the segmental level), but it also includes synthetic aspects in its approach: CAST-J combines phrase- and sentence-level materials in its lessons, emphasizing the integration by the trainee of various contextual and linguistic cues available in a given discourse. Thirdly, CAST-J is designed with a consideration of the coarticulatory effects. These effects are believed to be one of the main reasons why the ability to contrast visemes at the segmental level is not reflected in the sentence level as found in the study by Lesner et al. (1987). Lastly, a subject has a large amount of freedom in combining the two aspects depending on what is the most
salient cue (e.g. visemic cue, contextual cue) in speechreading under certain circumstances.

With the rising popularity and success of cochlear implants, implementation of CAST-J should meet the needs of the patients as well as those of the clinicians.
CHAPTER 4

VISEMES

4.1. Concept of viseme in general

The term viseme, according to Fisher (1968), refers to any individual and contrastive unit of visually perceived speech. It is conceptually analogous to the term phoneme, a contrastive unit of auditorily perceived speech. In the glossary of a textbook of audiology (Sheeley, 1985, p. 1090), viseme is defined as "a group of speech sounds that all have the same appearance when spoken". The phonemes /p/, /b/ and /m/, whose labial articulations look the same, thus constitute a viseme. Sheeley also uses 'visual phoneme' as a synonym for viseme. A "word that looks the same on the lips as another word" (ibid., p. 1073) is termed a homophone of that other word. For example, "pine" is a homophone of "mine", and these two words are said to be homophenous. The first attempt to discover a 'unit of lipreading' comparable to the phoneme appears to have been made by Woodward and Barber (1960). They proposed four categories of visually perceived English consonants (i.e. viseme groups): bilabial (p, b, m), rounded labial (m, w, r), labiodental (f, v), and non-labiodental (t, d, n, l, θ, ð, s, z, č, ħ, š, ž, y, k, g, h). In these definitions, visemes are considered to be a re-categorization of phonemes, mainly with respect to place of articulation, and they are not independent of phonemes. In order to refine the definition of the viseme, possible meanings of viseme in hypothetical situations will be discussed in comparison with some other 'units' of speech and language.

Henderson (1985) discusses the use of the term grapheme, which is used with different meanings by different authors. Originally, a grapheme denoted a minimal contrastive unit in a writing system (Sense 1). For example, in hand writing, the letter "a" belongs
to the grapheme I a l, while the letter << g >> belongs to the grapheme I g l. Both << a >> and << a >> belong to the grapheme I a l, while both << g >> and << g >> belong to the grapheme I g l. In other words, both << a >> and << a >> are members (allographs) of the grapheme I a l, and both << g >> and << g >> are allographs of the grapheme I g l. In this sense, the role of graphemes in written language is considered parallel to that of phonemes in spoken language. The term grapheme, on the other hand, is sometimes used to denote a letter or a cluster of letters that refers to or corresponds to a single phoneme in speech (Sense 2). For example, << p >> in "hoping" and << pp >> in "hopping" belong to the same grapheme I p l; similarly, << gh >> in "laugh", << ph >> in "graph", and << f >> in "fish" belong to the same grapheme I f l. In other words, << p >> and << pp >> are allographs (Sense 2) of the grapheme I p l, whereas << gh >>, << ph >> and << f >> are allographs of the grapheme I f l. This definition of grapheme is based on spoken language, with a view that written language is "a mere notation for representing speech sounds" (Henderson, 1985, p. 135). Henderson claims that Sense 1 is independent of the phoneme: a grapheme is "the minimal functional distinctive unit of any writing system" (p. 137), and "not in the phoneme-representing sense" (p. 137). Although the writing system of a language might have been developed on the basis of a correspondence with spoken language, written language can exist without speaking out the script. Languages such as Latin, classical Greek, and Sanskrit are examples of this type. Extreme cases are found in written forms in music or in the formulae of physical sciences, where the script does not always need to be pronounced. (Stetson, 1937 cited in Henderson, 1985). Therefore, graphemes (in Sense 1) can exist independently, without any reference to phonemes, even though they are procedurally defined in a similar way to phonemes.

Sign language is another type of language which exists without any reference to spoken or written languages. The minimal contrastive unit of sign language is called a chereme (Stokoe, Casterline and Croneberg, 1965). The grapheme (Sense 1) and the chereme
are respectively the minimal contrastive unit of written language and sign language, while the phoneme is the minimal contrastive unit of spoken language.

In view of these definitions, how does one define the viseme? For this author, a viseme, just as a phoneme, is a unit of spoken language or speech. From the perceptual point of view, spoken language or speech can be considered from two perspectives: audible speech and visible speech. A phoneme is the minimal contrastive unit of audible speech, whereas a viseme is the minimal contrastive unit of visible speech. It is possible to conceive a type of communication that would be performed solely by means of the visual information provided by articulatory movements, i.e. without any auditory information. If a sociolinguistic group of speakers had to communicate by means of visible speech alone, without audition, a number of easy-to-see visemes would no doubt develop, apart from the regular phonology of audible speech. Phonemes (of audible speech) are determined primarily by three elements: place of articulation, manner of articulation and voicing contrast. Likewise, each viseme could be determined by parameters such as those used for speech articulations, or preferably, by similar parameters more appropriate for describing the visual aspects of these same articulations, e.g. the positioning of the articulators, the direction of articulatory movements, and/or the duration of articulation. When visemes are considered as a visual aspect of a regular spoken language (audible speech), the number of visemes is unlikely to exceed that of phonemes, because acoustic and phonetic characteristics of speech do not necessarily correspond to visually discernable features and it is not very likely that there is some visible characteristic of speech, which is not audible but is important and useful for speechreading.

Since the visemes proposed in the past have been determined specifically for the purpose of speechreading or lipreading audible speech, which consists of phonemes, it would be hard to postulate that these visemes are independent of phonemes. The definition of the viseme, as used today, appears to be fairly parallel to the definition of the phoneme used in phonology. In this study, as in most past research, a viseme is considered as a minimal unit
of speechreading which corresponds to one (or more) phoneme of audible speech (phoneme-based speech). Each viseme is a set whose members will be called allophenes, and each allophene in turn is a set of physical realizations which will be called phenes. In order to denote visemes and their corresponding (allo)phenes, backward slashes \ \ and angular brackets < > will be used, in the same way as / / and [ ] are used for phonemes and (allo)phones.

For a given language, it is not easy to define a complete and consistent set of visemes for several reasons. Great methodological differences are found in past studies on visemes. Some authors use normal hearing subjects, while others use hearing-impaired subjects; some use identification tasks, while others use discrimination tasks. Criteria for categorizing visemes vary from author to author. For example, the hierarchical clustering analysis developed by Johnson (1967) is commonly used as a criterion for grouping visemes. It is a technique which combines objects (e.g. vowels, consonants) into optimally homogenous groups on the basis of measures of similarity (or confusability) between objects. By using a confusion matrix, one can obtain percentages of how often a particular item was confused with others. The higher the percentage of confusion, the more similar the items are. In a viseme recognition task, if a consonant was frequently identified as (or confused with) another consonant, these consonants both belong to the same visemic group (or a cluster). A 65% or higher within-cluster response rate was employed by Montgomery and Jackson (1987) to group vowel visemes, whereas a 75% or higher within-cluster response rate was employed by Walden et al., (1977, 1981) and by Lesner et al. (1987) to group consonant visemes. There are other reasons for difficulties, in addition to methodological differences: (1) coarticulatory effects may lead different investigators to partition differently the 'visemic' space into visemes; (2) the visibility of articulation varies from speaker to speaker (Benguerel and Pichora-Fuller, 1982; Montgomery et al. 1987); (3) speechreading training influences the ability to contrast visemes at the monosyllable level. Walden et al. (1977) report that hearing impaired adults, who could
discriminate five visemes prior to speechreading training, were able to discriminate nine visemes after training. The visemes they obtained pre- and post-training are shown below.

<table>
<thead>
<tr>
<th>Pre-training</th>
<th>Post-training</th>
</tr>
</thead>
<tbody>
<tr>
<td>{ &lt; θ &gt;, &lt; δ &gt; }</td>
<td>{ &lt; θ &gt;, &lt; δ &gt; }</td>
</tr>
<tr>
<td>{ &lt; f &gt;, &lt; v &gt; }</td>
<td>{ &lt; f &gt;, &lt; v &gt; }</td>
</tr>
<tr>
<td>{ &lt; p &gt;, &lt; b &gt;, &lt; m &gt; }</td>
<td>{ &lt; p &gt;, &lt; b &gt;, &lt; m &gt; }</td>
</tr>
<tr>
<td>{ &lt; f &gt;, &lt; 3 &gt; }</td>
<td>{ &lt; f &gt;, &lt; 3 &gt; }</td>
</tr>
<tr>
<td>{ &lt; w &gt; }</td>
<td>{ &lt; w &gt; }</td>
</tr>
<tr>
<td></td>
<td>{ &lt; s &gt;, &lt; z &gt; }</td>
</tr>
<tr>
<td></td>
<td>{ &lt; t &gt;, &lt; d &gt;, &lt; n &gt;, &lt; k &gt;, &lt; g &gt;, &lt; j &gt; }</td>
</tr>
<tr>
<td></td>
<td>{ &lt; r &gt; }</td>
</tr>
<tr>
<td></td>
<td>{ &lt; l &gt; }</td>
</tr>
</tbody>
</table>

The number of English consonant visemes reported in the past ranges from 3 to 12 (Jeffers and Barley, 1971). For vowel visemes, Montgomery and Jackson (1983) reported 8 visemes, including diphthongs, while Nitchie (1979) reported 9 visemes. Visemes reported in these two studies are shown below.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>{ &lt; i &gt;, &lt; i &gt; }</td>
<td>narrow { &lt; i &gt; }</td>
</tr>
<tr>
<td></td>
<td>relaxed narrow { &lt; i &gt; }</td>
</tr>
<tr>
<td>{ &lt; u &gt;, &lt; Ω &gt; }</td>
<td>small, oval { &lt; u &gt; }</td>
</tr>
<tr>
<td>{ &lt; u &gt; }</td>
<td>puckered { &lt; u &gt; }</td>
</tr>
<tr>
<td>{ &lt; a &gt;, &lt; Ω &gt;, &lt; Α &gt; }</td>
<td>wide { &lt; a &gt; }</td>
</tr>
<tr>
<td>{ &lt; au &gt; }</td>
<td>elliptical { &lt; Α &gt;, &lt; Ω &gt;, &lt; Ω &gt; }</td>
</tr>
</tbody>
</table>
Despite the absence of a set of universally accepted visemes for English (or for any other language), there is a generally accepted hierarchy of visemic groups, from well-defined to less well-defined visemes (Pichora-Fuller, 1980). Benguerel and Pichora-Fuller (1982) discussed the notion of 'visual dominance' of phonemes. In their study, \{ /p/, /f/, /w/, /θ/, /u/ \} were shown to be visually dominant, i.e. they were easily recognized in all phonetic environments of their experiment, while \{ /t/, /k/, /s/, /ʃ/, /i/, /æ/ \} were less dominant or prone to variation: "visually dominant phonemes may be representative members of well-defined visemic categories" (ibid., p. 606). Labial, labiodental, and linguadental movements in natural linguistic contexts are the important features for the determination of visemes.

4.2. Visemes for Japanese

4.2.1. Phonological and phonetic characteristics of Japanese

There exist great phonological variations between dialects of Japanese. This paper basically focuses on the Tokyo dialect, which is considered to be the basis of Standard Japanese (kyōtsūgo).

Japanese phonemes and their main allophones are shown in Table I, which is based on the work of Hattori (1960, 1979, 1984). Consonants are grouped by manner of articulation. In the table, some modifications to Hattori's system were made by this author for practical purposes, based on the review of other works in the literature (Kawakami, 1977; Jouo, 1980).

For phonemes which have been modified by this author and for those which are
different from the corresponding English phoneme, further description is given for each case individually.

/ g /, / η /: In the Tokyo dialect, traditionally, / g / and / η / have been traditionally considered as phonemically distinctive, although it is not the case in some other dialects. In a recent study, however, Inoue (1983) has shown that, even in the Tokyo dialect, / g / and / η / are becoming phonemically non-distinctive, and that / g / is "taking over" / η /. This tendency was reported to be particularly noticeable in the younger generation. Inoue has further stated that the older generation is more likely to view the taking over of / η / by / g / as a deterioration. The younger generation, on the other hand, is unaware of the former difference between the two phonemes, thus it does not express any opinion about this sociolinguistic change. Since the place of articulation of these two phonemes is identical, the nasality of / η / will not be an issue for speechreading. Therefore, in this study, the phoneme / η / will no be considered: / g / and / η / will be considered to constitute a single phoneme denoted as / g /.

/ s /, / z /: In a strict sense, the phonetic transcription of Japanese / s / and / z / include [ s ], [ s ], [ ʦ ], [ z ], and [ ʣ ]. Conventionally, however, the phonetic symbols [ s ], [ s ] and [ ʣ ] are used, because [ s ], [ ʦ ] and [ ʣ ] are respectively very close to each other in terms of place of articulation, and these two groups of consonants are not phonemically distinctive in Japanese. Although the voiced fricatives [ z ] and [ ʣ ] can occur in certain phonetic contexts, they are not phonemically distinctive from [ ʣ ] and [ ʣ ], which are the major allophones of the phoneme / z /.

/ r /: The phoneme / r / is usually realized by the alveolar flap [ r ] in intervocalic position (Sawashima and Kiritani, 1985). Its place of articulation varies, depending upon the following vowel: most forward before [ i ] and gradually moving backward in the context of [ e ], [ u ], [ a ], [ o ] (Chiba and Kajiyama, 1958). In word initial position, on the other hand, a weak plosive very similar to [ d ] is observed. This weak plosive is considered to be
<table>
<thead>
<tr>
<th>PHONEMES</th>
<th>MAIN ALLOPHONES</th>
<th>PHONETIC REALIZATION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CONSONANTS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stop</td>
<td>/ p /</td>
<td>[ p ]</td>
</tr>
<tr>
<td>/ b /</td>
<td>[ b ]</td>
<td></td>
</tr>
<tr>
<td>/ t /</td>
<td>[ t ]</td>
<td></td>
</tr>
<tr>
<td>/ d /</td>
<td>[ d ]</td>
<td></td>
</tr>
<tr>
<td>/ k /</td>
<td>[ k ]</td>
<td></td>
</tr>
<tr>
<td>/ g /</td>
<td>[ g ]</td>
<td></td>
</tr>
<tr>
<td>/ s /</td>
<td>[ s ] [ c ]</td>
<td></td>
</tr>
<tr>
<td>/ z /</td>
<td>[ dz ] [ dz ]</td>
<td></td>
</tr>
<tr>
<td>/ h /</td>
<td>[ h ] [ ç ] [ ð ]</td>
<td></td>
</tr>
<tr>
<td>affricate</td>
<td>/ c /</td>
<td>[ ts ] [ tç ]</td>
</tr>
<tr>
<td>flap</td>
<td>/ r /</td>
<td>[ r ] [ r ] [ l ]</td>
</tr>
<tr>
<td>nasal</td>
<td>/ m /</td>
<td>[ m ]</td>
</tr>
<tr>
<td>/ n /</td>
<td>[ n ] [ n ]</td>
<td></td>
</tr>
<tr>
<td>/ ñ /</td>
<td>[ ñ ]</td>
<td></td>
</tr>
<tr>
<td>semivowel</td>
<td>/ j /</td>
<td>[ j ]</td>
</tr>
<tr>
<td>/ w /</td>
<td>[ w ]</td>
<td></td>
</tr>
<tr>
<td><strong>VOWELS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ i /</td>
<td>[ i ]</td>
<td></td>
</tr>
<tr>
<td>/ e /</td>
<td>[ ë ]</td>
<td></td>
</tr>
<tr>
<td>/ a /</td>
<td>[ ø ]</td>
<td></td>
</tr>
<tr>
<td>/ o /</td>
<td>[ ø ]</td>
<td></td>
</tr>
<tr>
<td>/ u /</td>
<td>[ ø ]</td>
<td></td>
</tr>
<tr>
<td><strong>SPECIAL PHONEMES</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ N /</td>
<td>[ m ] [ n ] [ ŋ ] [ ñ ] [ n ]</td>
<td></td>
</tr>
<tr>
<td>/ Q /</td>
<td>[ ? ] [ p ] [ t ] [ k ] [ s ] [ ç ]</td>
<td></td>
</tr>
<tr>
<td>/ ’ /</td>
<td>[ ? ]</td>
<td></td>
</tr>
</tbody>
</table>
a natural preparatory motion for the word initial position of the flap [ r ] (Kawakami, 1977). However, some individuals possess [ l ] and/or [ r ] as allophones of Japanese / r /, although these are not the main allophones (Hattori, 1984). Use of [ r ] can be considered to be due to dialectal variation rather than to the influence of a foreign language. Occurrence of [ l ], which is not phonemically distinctive from [ r ] and [ r ], can be considered as a variation influenced by certain phonetic environment.

/ j /, / w /: The phonemes / j / and / w / are semivowels (or semiconsonants). Besides occurring in syllable initial position as in / ja /, / ju /, / jo /, the semivowel / j / can follow consonants, as in / pj /, / bj /, / sj /, / zj /, / hj /, / cj /, / mj /, / nj /, / rj /. The other semivowel / w / is always found in syllable initial position, except in the cluster / kw /.

Vowels: Japanese has five vowels: / i /, / e /, / a /, / o /, and / u /. The vowel / u / in the Tokyo dialect has no or little lip rounding, although some other dialects do. Its major allophone is [ uu ]. Among the five vowels of Japanese, lip rounding is most obvious in / o /. The tongue position for a high back vowel is generally slightly fronted as compared to [ u ], [ uu ]. Devoicing of the vowels [ i ] and [ uu ] occurs in certain contexts (e.g. between voiceless consonants, or in word final position, when preceded by voiceless consonants, etc.), while devoicing for other vowels seldom occurs. Conventionally, Japanese vowels are transcribed phonetically as [ i ], [ e ], [ a ], [ o ], and [ uu ], thus this study will also follow this notation.

/ N /: The phoneme / N / has several phonetic realizations: [ m ], [ n ], [ n ], [ n ], [ n ] depending upon the phonetic context. If / N / is followed by a labial consonant, it will be realized as [ m ], e.g. / siNpu / : [ impu ] (BRIDE); if it is followed by an alveolar consonant, it will be realized as [ n ], e.g. / teNto / : [ tento ] (TENT); if it is followed by a vowel or is used in utterance final position, it will be realized as [ n ], / teN / : [ ten ] (POINT).

/ Q /: The phoneme / Q / always occurs after a vowel, either in front of one of the voiceless consonants [ p ], [ t ], [ k ] [ ts ], [ tf ], [ s ], [ s ], or in word final position.
Exceptions occur in a limited number of loanwords in which the phoneme /\( Q /\) precedes the voiced stop [d], [d3], [g]. The phoneme /Q/ is usually considered as an "implosive", in the sense described by Saussure (1959, pp. 51-54). In phonetic transcriptions, /Q/ is usually considered as a geminated consonant such as in the word nippon [nip\( ^{\prime} \)pon] (Japan).

In the case of plosives, the phoneme /Q/ corresponds to the first part of the geminated consonant, which is always an implosive, and to part of the closure. As a consequence, acoustically, a silent interval is created. While a geminated consonant is the most common realization of the phoneme /Q/, it can be realized as a combination of two different plosives, subject to certain constraints: [nit\( ^{\prime} \)pon] or [nik\( ^{\prime} \)pon] as long as [t] or [k] is not released; it can be realized as a combination with the glottal stop [ʔ] as in [niʔpon]. The glottal stop [ʔ] in Japanese, however, is so weak that it is not clearly observable. When the phoneme /Q/ is followed by the phoneme /c/, as in /siQci / (swamp), it is not realized as [ji\( ^{\prime} \)fi] but as [ji\( ^{\prime} \)fi]. Again, the first [t] is implosive and [\( ^{\prime} f \)] is explosive.

Unlike in the case of plosives, no obvious silent interval is observed for geminated fricatives (e.g. [issei] unanimously, [iff\( ^{\prime} \)wun] an instant), due to the continuant nature of fricatives. However, in the [ss] and [\( ^{\prime} f \)], a qualitative difference between the implosive and the explosive part can be perceived (Hattori, 1979). The phoneme /Q/ corresponds to the first fricative ([s], [\( ^{\prime} f \)]), which is closely associated to the preceding vowel [i]. Hattori states that one of the characteristics of the geminated consonant is the strong linkage, or 'fester Anschluss', of the first consonant to the preceding vowel. The second fricative ([s], [\( ^{\prime} f \)]), on the other hand, is closely associated to the following vowel. Phonetically, [iff\( ^{\prime} wun\)] is extremely close to [if\( ^{\prime} \)wun]; phonemically, however, they are different.

Loanwords in Japanese, such as baggu (bag) and beddo (bed), are realized phonetically as [bak\( ^{\prime} \)gu] and [bet\( ^{\prime} \)do], respectively.

/\( ^{\prime} /\): The phoneme /\( ^{\prime} /\) was first proposed by Hattori (1951, cited in Hattori, 1979). The purpose of proposing this phoneme was to make all the syllables start phonemically with
a consonant, e.g. /'i/, /'e/, /'a/, /'o/, /'u/, and to make a semivowel always follow a consonant, as in palatalized consonants, e.g. /'ja/, /'ju/, /'jo/, /'wo/. Although Hattori (1979) claims that the phoneme /ʼ/ is the voiced counterpart of the glottal consonant /h/, neither its phonetic realization nor its acoustic characteristics are described clearly. The existence this phoneme, however, is important as far as vowels are concerned. In Japanese, a single vowel (e.g. [o], [i]), a sustained vowel (e.g. [o:], [i:]), and a successive occurrence of two identical vowels (e.g. [oo], [ii]) are phonemically distinctive. For example, [oja] (PARENT(S)) contrasts with [o:ja] (LANDLORD). Similarly, the word [sato:ja] (SUGAR MAKER) contrasts with [sato'oja] (FOSTER PARENTS). Hattori suggests transcribing these latter two words phonemically as /sato'oja/ and /sato'oja/ respectively. Kawakami (1977) uses the glottal stop symbol [?] as a marker of separation of two vowels, although he warns that [?] in Japanese is very weak and that it "does not exist as one consonant of Japanese." This study will employ the phoneme /ʼ/ proposed by Hattori, not as a regular consonant, but as a special phoneme only for the case of the vowels described above.

4.2.2. Significance of haku (mora) in Japanese

The mora has been defined as "the smallest unit for measuring the quantity of sound in prosody" (Onishi, 1981, p. 352). In some languages, such as classical Latin, the mora is seen as a syllable, while in some other languages, such as classical Greek, the mora is seen as an accent. In Japanese, the mora is often considered to have a constant length in time. In some languages, the concept of mora and that of syllable are ostensibly identical, but in others, they are not. Japanese belongs to the latter case. For example, the word nippon (JAPAN) is considered to consist of two syllables: nip + pon, but of four morae: ni + p + po + n (Hattori, 1960).

Hattori (1960) was the first Japanese author to use the term mora in a linguistic sense. Kindaichi (1968), however, proposed to adopt the term haku to refer to the mora,
because the definition of the mora by Hattori does not always agree with the term used by other authors (e.g. Trubetzkoy, 1969) when applied to other languages besides Japanese. This study will use the term haku.

In Japanese, haku is the basic unit of rhythm, and structurally, it can be either V, C(j)V, or one of the two special consonantal phoneme / N / or / Q /. The concept of haku is significant in Japanese, not only because it is the basic unit of the rhythm, but also because it is related to the writing system called kana (phonemic script). One kana corresponds to each haku except when C is followed by the semivowel /j/, as in /pj/, /bj/, /mj/, /sj/, /zj/, /cj/, /nj/, /kj/, /gj/, /rj/, /hj/. For these cases, two kanas are used: one full-size kana, followed by a small-size kana in subscript position, indicating the presence of the semivowel. Both full-size and small-size kana are transcribed in the phonemic-alphabetic script romaji. There are three styles of romaji: hyōjun-shiki, nihon-shiki, and kunrei-shiki. Among the three styles, Hyōjun-shiki romaji best reflects the actual pronunciation of kana. Nihon-shiki romaji and kunrei-shiki romaji (modification of nihon-shiki) are less close to the actual pronunciation. For example, the word meaning "foot" in English is pronounced as [aji] in Japanese. It is written ashi in hyōjun-shiki romaji, whereas it is written asi in nihon-shiki romaji and in kunrei-shiki romaji. (See Hattori (1979) for further information.) This study will follow the hyōjun-shiki romaji, which is most commonly used today.

All possible hakus for the Tokyo dialect are listed in Tables II and III, in romaji and in phonetic symbols respectively. Japanese sentences are combinations of these hakus. A elongated vowel (e.g. [ oː ]) is counted as two hakus, and the symbol R - introduced by Kindaichi (1968) and elaborated on by other linguists - is commonly used as the second haku corresponding to the second half of an elongated vowel. Phonetically, the haku R is denoted by the symbol [:], and in romaji, R is denoted by a macron (a bar placed on the top of the preceding vowel). For example, a word kēki, which is the romaji transliteration of the English word "cake", is a three-haku word ( ke + R + ki ), and it is pronounced as [ kəːki ].
Table II: List of Hakus Written in Romaji

The list is arranged according to the traditional ordering of Japanese. All hakus except N, Q, and R, are written in romaji.

<table>
<thead>
<tr>
<th>Letter</th>
<th>Romaji</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>i</td>
</tr>
<tr>
<td>ka</td>
<td>ki</td>
</tr>
<tr>
<td>ga</td>
<td>gi</td>
</tr>
<tr>
<td>sa</td>
<td>shi</td>
</tr>
<tr>
<td>za</td>
<td>ji</td>
</tr>
<tr>
<td>ta</td>
<td>chi</td>
</tr>
<tr>
<td>da</td>
<td>de</td>
</tr>
<tr>
<td>na</td>
<td>ni</td>
</tr>
<tr>
<td>ha</td>
<td>hi</td>
</tr>
<tr>
<td>ba</td>
<td>bi</td>
</tr>
<tr>
<td>pa</td>
<td>pi</td>
</tr>
<tr>
<td>ma</td>
<td>mi</td>
</tr>
<tr>
<td>ra</td>
<td>ri</td>
</tr>
<tr>
<td>wa</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
</tr>
</tbody>
</table>
Table III: List of Hakus Written in Phonetic Symbols

The list is arranged according to the traditional ordering of Japanese. All hakus except N, Q, and R, are written in romaji. Hakus for loanwords are indicated in parentheses.

| a | i | u | e | o | ja | ju | jo | (je) |
| ka | ki | ku | ke | ko | kja | kju | kjo |
| ga | gi | gu | ge | go | gja | gju | gjo |
| sa | (si) | su | se | so |
| ja | fi | fu | (fe) | fo |
| dza | dze | dzo |
| dʒa | dʒi | dʒu | (dʒe) | dʒo |
| ta | (ti) | tu | te | to | (tjo) |
| tʃa | tʃi | tʃu | (tʃe) | tʃo |
| (tsa) | (tsi) | tsu | (tse) | (tso) |
| da | (di) | (du) | de | do | (dju) |
| na | ne | no |
| ɲa | ɲi | ɲu | (ɲe) | ɲo |
| ha | he | ho |
| ça | çi | çu | (çe) | ço |
| (φa) | (φi) | φu | (φe) | (φo) |
| ba | bi | bu | be | bo | bja | bju | bjo |
| pa | pi | pu | pe | po | pja | pjʊ | pjo |
| ma | mi | mu | me | mo | mjʊ | mjʊ | mjo |
| ra | ri | ru | re | ro | rja | rju | rjo |
| wa | (wi) | (we) | (wo) |
| N | Q | R |
The symbol N and Q are respectively used for the special phonemes / N / and / Q /. Unlike the other consonant phonemes, which are always tied to a vowel, both / N / and / Q / are independent hakus.

Kindaichi (1968) uses 112 hakus for the Tokyo dialect, while Hattori (1960) uses 110. Two hakus - R and wo - were not included in Hattori’s list. When sounds occurring only in loanwords are included, the number of hakus is increased to approximately 130 (Kindaichi, 1981). While both Hattori and Kindaichi counted g(j)V and η(j)V as separate hakus, this author will consider only 103 hakus for the Tokyo dialect, excluding η(j)V, which is not phonemically distinctive from g(j)V for some people, and wo, which is commonly seen in loanwords and is replaced by o in conventional Japanese. Possible hakus including loanwords are listed in Table III in phonetic forms. Phonetic forms are not given for the hakus N, Q and R, because their phonetic realization varies depending upon the phonetic context.

There is no definite number of hakus for loanwords; the number may vary, depending upon an individual’s linguistic experience and needs. For example, for people who frequently use musical terminology or specific German nouns, the phoneme strings / tsi /, / tse /, / tsa / and / tso / may have become a part of their total phonemic inventory. The hakus widely used by most Japanese in their daily life are listed in Table IV. Some hakus are so common that people may be unaware that they are not conventional combinations of Japanese phonemes. People who have a fair amount of exposure to foreign languages or to modern technological terms may not have any difficulty in producing the hakus listed in parentheses in Table III, despite the phonological constraint of traditional Japanese. This tendency is presumably more true for the younger generation. People unfamiliar with loanwords, on the other hand, may have a rather reduced inventory of hakus.

The labiodental fricative consonants [ f ] and [ v ] are not part of the phonetic inventory of conventional Japanese. The voiceless labiodental fricative [ f ], in the original pronunciation of loanwords, is replaced by the bilabial voiceless fricative [ φ ], and the voiced
Table IV: Examples of Loanwords Widely Used in Japanese

<table>
<thead>
<tr>
<th>HAKU</th>
<th>ORIGINAL WORDS</th>
<th>ARTICULATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ φe ]</td>
<td>FENCE</td>
<td>[ φensu: ] or [ fensu: ]</td>
</tr>
<tr>
<td>[ φa ]</td>
<td>FAN</td>
<td>[ φan ] or [ fan ]</td>
</tr>
<tr>
<td>[ φo ]</td>
<td>FOUR</td>
<td>[ φo: ] or [ fo: ]</td>
</tr>
<tr>
<td>[ wi ]</td>
<td>WEEK</td>
<td>[ wi:kuy ]</td>
</tr>
<tr>
<td>[ we ]</td>
<td>WAY</td>
<td>[ wei ]</td>
</tr>
<tr>
<td>[ wo ]</td>
<td>WATER</td>
<td>[ wota: ]</td>
</tr>
<tr>
<td>[ si ]</td>
<td>CT SCAN</td>
<td>[ si:tisukjan ]</td>
</tr>
<tr>
<td>[ ti ]</td>
<td>TEA</td>
<td>[ ti: ]</td>
</tr>
<tr>
<td>[ di ]</td>
<td>DINNER</td>
<td>[ dina: ]</td>
</tr>
<tr>
<td>[ tuu ]</td>
<td>TWO</td>
<td>[ tuu: ]</td>
</tr>
<tr>
<td>[ tʃe ]</td>
<td>CHECK</td>
<td>[ tʃek-kuy ]</td>
</tr>
<tr>
<td>[ dʒe ]</td>
<td>JET</td>
<td>[ dʒet-tø ]</td>
</tr>
<tr>
<td>[ fe ]</td>
<td>SHEPHERD</td>
<td>[ sepa:do ]</td>
</tr>
</tbody>
</table>
labiodental fricative [v] is replaced by the bilabial stop [b]. Hiki and Fukuda (1982, p. 192), however, point out that some people who are familiar with the pronunciation of English produce a sound by "pulling the lower lip backward and by touching it to the upper teeth, instead of by protruding and rounding the lips". They further report that this sound is close to English [f], but that it has shorter duration and less friction than English [f]. Inoue (1983), who investigated the production of [фи] versus [фи] in the loanword firumu (FILM), reports that the production of [f] appears to be related to an individual’s educational background. One group of subjects, who were students in a private high school, exhibited a higher usage of [f] than of [фи], whereas the other groups of subjects exhibited opposite results: a higher usage of [фи] than of [f]. Inoue (1983) also shows that the older generation tends to use a combination of two conventional hakus [фуи] + [и] for [фи] in the loanword firumu (FILM), while the younger generation uses [фи] or [фи] for it. Since the difference in place of articulation between [фи] and [f] is visually significant, further sociolinguistic research, including [v] versus [b] distinction will be valuable, particularly for the area of speechreading.

When these loanwords are transcribed in kana, usually one full-size kana for CV is followed by a subscript small-size kana for V. For example, {[фи], [f]} + {[и], [е], [а], [о]} is expressed with the kana for [фуи] plus a small size kana for {[и], [е], [а], [о]}; [w] + {[и], [е], [о]} is expressed with the kana for [wa] plus a small size kana for {[и], [е], [о]}; [tфи] is expressed with the kana for [тфи] plus a small size kana for [е].

4.2.3. Determination of visemes of Japanese

A few attempts to categorize visemes for Japanese can be found in the recent research literature, but the term viseme, or any translation of the word, is rarely used by Japanese researchers. Viseme is mentioned in the paper by Fukuda (1987) only as a mere introduction
to the concept. The present study aims at determining a set of visemes for Japanese which are suitable for the CAST-J lessons.

Whether several vowel visemes should be considered or not is still controversial. In the CAST lessons developed in the past (Benguerel and Pichora-Fuller, 1988; Pichora-Fuller and Benguerel, unpublished manuscript), English vowels were categorized as a single viseme V, and French vowels were grouped into two visemes: one for unrounded vowels, the other for rounded vowels. There are two reasons for having only one or two visemic groups for vowels: (1) the definition of the viseme, in particular for vowels, is not clear enough, and (2) auditory perception of vowels is much better than that of consonants. Although it was reported that the visual discrimination of English vowels was poor, due to coarticulatory effects (Montgomery, 1987), the results may not be directly applicable to Japanese, which differs from English on many points. The determination of visemic groupings for vowels appears to depend greatly on the linguistic features of the language under consideration. The significance of vowels in Japanese and the necessity of establishing separate visemes for the vowels will be discussed next.

Firstly, there are only five vowels in Japanese: / i /, / e /, / a /, / o /, / u /. Compared to several Indo-European languages (e.g. English, German, French), this number is considerably smaller. These five vowels play an important role in various linguistic aspects. In Japanese, each vowel or combination of vowels can be a meaningful word. Some of the most common and popular examples are given below:

/ i /  stomach, well, mind/feeling, difference
/ e /  painting/drawing, branch
/ a /  painted in white, mutism
/ o /  tail
/ u /  cormorant
Secondly, there is no or little variation across dialects for Japanese vowels. The five vowels are remarkably stable throughout Japan.

Thirdly, as discussed in Section 4.2.2, vowels are given more weight in terms of the structure of haku: V, CV, or C. The special phonemes (/ N / and / Q /) are the only consonant (C) which can be realized as a single haku on its own, and the occurrence of these two phonemes has been reported to be relatively low (Bloch, 1950). Bloch analyzed the frequency of occurrence of each phoneme, for all positions, from a text which consisted of two thousand phonemes. A calculation based on his results reveals that the ratio of occurrence was 0.1% for the final / N /, and 1.3% for the final / Q /. Therefore, Japanese can be considered to consist almost exclusively of hakus having a (C)V structure. The frequency of occurrence of vowels was also shown to be higher than for consonants (Bloch, 1950). When phonemes for word boundary and accent are ignored, 53.6% of the text consisted of vowels and 46.4%
of consonants. Although the sample used in the Bloch's study is fairly limited in size, the results illustrate one feature of Japanese: the importance of vowels in terms of frequency of occurrence. This is a remarkable difference from English, in which consonants are more frequently used than vowels. A study based on the predicted pronunciation of read text shows that in English, the relative frequency of vowels and diphthongs is 38%, while that of consonants is 62% (Flanagan, 1972).

Fourthly, Japanese vowels may sometimes become voiceless: high vowels [i] and [u] are usually devoiced between voiceless consonants. Generally, the auditory perception of vowels is better than that of consonants because of their higher intensity. In Japanese, however, vowels are not necessarily audible at all times.

Lastly, evidence has been reported that, visually, Japanese vowels are easily distinguished. Fukuda et al. (1976) examined the visual discrimination of vowels in CV monosyllables for thirteen hearing impaired children. Total correct scores for each vowel were relatively high: 91% for /i/, 83% for /e/, 81% for /a/, 88% for /o/, and 80% for /u/. The most common confusions observed in their study were that of /a/ with /e/ (10%) and that of /u/ with /o/ (9%). Fukuda and Hiki (1977) also reported high correct discrimination scores of vowels for mono- and disyllables (nonsense words) and for meaningful words. Nine subjects out of thirteen performed well for all vowels, with correct scores of 85 to 96%, although great individual variability was reported among the other four subjects. High correct scores for vowel discrimination were also found for cochlear implant patients (Fukuda et al., 1988; Funasaka et al., 1987). Stroboscopic observation of articulations in which coarticulatory effects were taken into consideration has shown that the Japanese vowels are relatively distinctive, both in the frontal and the lateral view (Fukuda and Hiki, 1982).

Montgomery et al. (1987) examined the visual discrimination of the five English vowels /i/, /u/, /a/, /u/, /o/ in various consonantal contexts. The tense vowels were found to
be more easily recognized than the lax vowels. The mean correct score for the tense vowels was 57.9%, whereas for the lax vowels, it was 35.7%. These scores are considerably lower than those obtained by Fukuda and her associates (Fukuda et al. 1976; Fukuda and Hiki, 1977). There are possible explanations for this discrepancy. One of them is that Montgomery et al. (1987) had limited their study to the discrimination between five vowels, four of which were, pairwise, adjacent in the vowel triangle, and thus had similar articulatory features. The five Japanese vowels, on the other hand, are more evenly distributed in the vowel triangle; articulatory features are thus relatively less similar. If the data presented in the study of Montgomery et al. (1987) is recomputed for three groups of vowels \{ /i/, /ɪ/ \}, \{ /a/ \}, \{ /u/, /u/ \} instead of the original five, percentage of correct responses increases to approximately 85%. Overall, the discrimination of the five vowels of Japanese can be considered as relatively easy.

All the factors discussed above support the consideration of each Japanese vowel as a different viseme. Therefore, the visemes for each vowel, with their main allophones, are described as follows:

\[
\begin{align*}
\text{i} & : <i> \\
\text{e} & : <e> \\
\text{a} & : <a> \\
\text{o} & : <o> \\
\text{u} & : <u>.
\end{align*}
\]

There is also good evidence that speechreading training for vowels is imperative. In most cases, hearing impaired people can perceive vowels better than consonants through their auditory channel, because, acoustically, vowels are low in frequency and high in intensity. Particularly with amplification, speechreading of vowels may not be indispensable for them.
Under certain circumstances, however, amplification is no longer helpful. For example, a hearing impaired person is most likely to turn off his/her hearing aid in noisy surroundings, e.g. on the train or the subway, a common mode of transportation in Japan. In this case, speechreading becomes a primary means of speech perception rather than a supplemental one.

Three studies of the categorization of consonants will now be compared. By means of 'labiograms', Sudo (1973) conducted a quantitative investigation of the labial and facial movements in C/(j)a / monosyllables. Twelve parameters (e.g. width, velocity, duration of certain articulatory movements) were investigated in his study: four parameters for the center of the lips, two for the corner of the lips, and six for the cheeks. Sudo (1973) explains that the term 'labiograms' used in his study refers to an experimental apparatus, which consists of three CdSe cells, three DC amplifiers and one electromagnetic oscillograph. The CdSe cells transform the light waves reflected from the speaker's face into electric signals. Based on his results, Sudo (1979) proposes four groups of phonemes, in terms of the similarity of the labial and the facial movements. Although not all the consonants are included in his study, Sudo's grouping can be applied to the categorization of visemes.

GROUP 1 \{ \langle p \rangle, \langle b \rangle, \langle m \rangle, \langle pj \rangle, \langle bj \rangle, \langle mj \rangle \}
GROUP 2 \{ \langle t \rangle, \langle n \rangle, \langle k \rangle, \langle g \rangle, \langle s \rangle, \langle ts \rangle \}
GROUP 3 \{ \langle kj \rangle, \langle gj \rangle, \langle rj \rangle, \langle tf \rangle, \langle z \rangle, \langle dz \rangle, \langle r \rangle \}
GROUP 4 \{ \langle w \rangle, \langle \phi \rangle \}

These four groups are further divided into 10 subgroups. Thus, Sudo's study indicates that the visemes of Japanese consonants, at least in CV monosyllables, can range from 4 to 10.

Fukuda and Hiki (1982) conducted a more extensive study of the visual characteristics of Japanese consonants in various phonetic environments: 100 single haku words which were either V or CV monosyllables, 15 two-haku words, 12 three-haku words, and 10 four-haku
words. Both frontal and lateral views of the movements of the mouth were examined from pictures taken by a special camera in which a long roll of film was driven continuously. The mouth of the speaker was illuminated by a light strobed at 5, 10, or 20 Hz. In addition to the measurement of the change in distance of various parts of the mouth, the authors utilized a type of display called 'labiogram', produced by filming small aluminum pellets placed on the important points on the lips. Using stroboscopic illumination, the traces of these points were photographed and displayed in a pseudo three-dimensional coordinate system (time, height, and width). The waveforms of the sound was also included on the display. The authors came up with eight mouth shapes. These eight groups are listed below with the 'mouth-shape symbols' assigned by Fukuda and Hiki (1982). The letters Vf stand for 'mouth shape of the following vowel'.

\[
\begin{align*}
p & : \{ <p>, <b>, <m> \}; \\
w & : \{ <w>, <\phi> \}; \\
t & : \{ <t>, <d>, <n> \}; \\
r & : \{ <r> \}; \\
s & : \{ <ts>, <dz>, <s>, <z> \}; \\
sy & : \{ <tf>, <d3>, <\eta>, <\phi>, <f>, <\varsigma> \}; \\
y & : \{ <j> \}; \\
Vf & : \{ <k>, <g>, <\eta>, <h> \}.
\end{align*}
\]

Fukuda and Hiki comment that these mouth shapes are discernible only in clearly articulated utterances, and can vary as a function of coarticulatory effects.

There are some important observations in the study of Fukuda and Hiki (1982) dealing with the determination of visemes. Fukuda and Hiki contend that the difference between \p\ and \p\+\j\ is discernable: the speed of the opening movement of the lips is slower
They report that the opening of the lower lip at 0.1 second after plosion for the utterance /pja/ is almost half that for /pa/. They also confirm an observation made by Fujimura (1961): the speed of lip opening in bilabial plosives is higher than in bilabial nasals. It is an indication that the bilabial plosives (/p/ and /b/) and the bilabial nasal (/m/) may belong to different visemes. Fukuda and Hiki (1982) further suggest that /t/ and /s/ are discernable: /s/ has longer duration than /t/ possibly due to the [+continuant] characteristic of fricatives. Although these findings have to be taken into consideration in categorizing the visemes, there is no evidence how well a speechreader perceives these features. Further research, including a comparative study of connected speech spoken by different speakers or at different speaking rates is needed.

Sato (1961) arrived at four groups for consonants when they are followed by the vowel /a/, /e/, or /o/:

- \{/w/\};
- \{/p/, /b/, /m/\};
- \{/h/, /k/, /g/, /ŋ/\};
- \{/t/, /d/, /n/, /s/, /z/, /ṣ/, /ts/, /dz/, /f/, /tʃ/, /dʒ/, /j/, /r/\}.

When they are followed by /i/ or /u/, he categorizes them into three groups, combining the latter two into one.

One of the remarkable characteristics of Sato’s approach is that he considers the haku as a unit of speechreading. Like other authors, Sato categorized vowels and consonants into several groups respectively. However, this categorization is not yet that of visemes in his sense. He further combined the two (the vowel groups and the consonant groups) to obtain visemes in his sense, i.e. a haku-level categorization. Sato (1961) categorized hakus into 14 groups and assigned a lipreading symbol, which corresponds to the haku-size viseme, to each
group. (These 14 symbols have been replaced by numbers in this study.) They are:

\[
\begin{align*}
1 & : < a > \text{ or } \{ < h >, < k >, < g > \} + < a > \\
2 & : < e > \text{ or } \{ < h >, < k >, < g > \} + < e > \\
3 & : < o > \text{ or } \{ < h >, < k >, < g > \} + < o > \\
4 & : < w > \text{ or } \{ < h >, < k >, < g >, < t >, < d >, < n >, < s >, < ts >, < dz >, < f >, < t f >, < d s >, < j >, < r > \} + < w > \\
5 & : < i > \text{ or } \{ < h >, < k >, < g >, < t >, < d >, < n >, < s >, < ts >, < dz >, < f >, < t f >, < d s >, < j >, < r > \} + < i > \\
6 & : \{ < t >, < d >, < n >, < s >, < ts >, < dz >, < f >, < t f >, < d s >, < j >, < r > \} + < a > \\
7 & : \{ < t >, < d >, < n >, < s >, < ts >, < dz >, < f >, < t f >, < d s >, < j >, < r > \} + < e > \\
8 & : \{ < t >, < d >, < n >, < s >, < ts >, < dz >, < f >, < t f >, < d s >, < j >, < r > \} + < o > \\
9 & : \{ < p >, < b >, < m > \} + < a > \\
10 & : \{ < p >, < b >, < m > \} + < e > \\
11 & : \{ < p >, < b >, < m > \} + < o > \\
12 & : \{ < p >, < b >, < m > \} + < w > \\
13 & : \{ < p >, < b >, < m > \} + < i > \\
14 & : < w > + < a >
\end{align*}
\]

Sato (1961) emphasizes the similarity between lipreading and the decoding of unfamiliar characters. The interpretation of a spoken message without auditory signal requires the skill to identify particular labial and labiodental movements and to match them to one of the many possible hakus, using various contextual cues. Sato's approach first provides a course in which a subject is required solely to identify the 14 symbols without making any connection with their meanings. Following training in matching the movements to non-linguistic symbols in
connected speech, the subjects proceed to the next level: the selection of the appropriate lexical item for the symbol identified.

Integrating these three studies, it is clear that \{<p>, <b>, <m>\} and \{<w>, <\emptyset>\} can be considered as well-defined visemes. For the less well-defined visemes, this author has decided to combine the grouping of the studies of Fukuda and Hiki (1982) and of Sato (1961), because these two studies came up with similar visemic groups, despite the fact that the authors have used totally different method for arriving at them: the study of Fukuda and Hiki (1982) is based on the quantitative measurements of articulatory movements, with a time precision of 0.1 second, whereas the study of Sato (1961) is based on his clinical research and experience with hearing impaired children learning speechreading.

The seven visemes for Japanese consonants arrived at are:

\[
\begin{align*}
p & : \{<p>, <b>, <m>\} \\
w & : \{<w>, <\emptyset>\} \\
r & : \{<r>\} \\
t & : \{<t>, <d>, <n>, <s>, <dz>, <ts>\} \\
k & : \{<k>, <g>, <h>\} \\
j & : \{<j>, <n>, <f>, <tj>, <d3>, <z>\} \\
f & : \{<f>\}
\end{align*}
\]

As discussed in Section 4.2.2., the phonemes /N/ and /Q/, and vowel elongation (symbolized by R) are recognized as hakus in Japanese. Fukuda and Hiki (1982) report that, when \p, \t and \k are preceded by N or Q, they are perceived either as consonants with a longer closure, in the case of \p and \t, or as consonants preceded by an elongated vowel, in the case of \k. Considering these observations, this author also proposes to use N, Q, and R as haku visemes. Their realization is exemplified as follows:
\[N\] \rightarrow \ p \ @ \ C \quad \text{where } C \text{ is } p, \ b, \text{ or } m; \\
\ t \ @ \ C \quad \text{where } C \text{ is } t, \ d, \ n, \ p, \ s, \ dz, \ f, \ tf, \ d3, \ r, \ ç, \ j; \\
\ k \ @ \ C \quad \text{where } C \text{ is } k, \ g, \ h, \ i, \ e, \ a, \ o, \ w \\
\[Q\] \rightarrow \ p \ @ \ C \quad \text{where } C \text{ is } p; \\
\ t \ @ \ C \quad \text{where } C \text{ is } t, \ s, \ f, \ tf, \ (d), \ (d3); \\
\ k \ @ \ C \quad \text{where } C \text{ is } k, \ (g) \\
\[R\] \rightarrow \ i \ @ \ i; \\
\ e \ @ \ e; \\
\ a \ @ \ a; \\
\ o \ @ \ o; \\
\ u \ @ \ u;
CHAPTER 5

EQUIPMENT AND INPUT/FEEDBACK MODE

5.1. Equipment

The hardware components required by CAST are: (1) a microcomputer including a hard
disk, a colour graphics card and a keyboard; (2) a high-quality colour video monitor (14''); (3)
a video disc player; (4) an amplifier and a speaker; and (5) a printer. (See Figure 1, p. 3.)
For the English version of CAST, an IBM PC-AT microcomputer with 512-kb RAM, BCD
Videolink RS 232 video controller, Sony PVM 1271Q TV monitor (14''), Sony SLO 325 video
cassette recorder, and Epson FX 286 printer were used. For CAST-J, a hardware setup
functionally equivalent to that for the English version of CAST will be used. The main
difference is that CAST-J replaces the video cassette player by a video disc player, which
provides a faster access every time the user wants to jump from one place to another.

The software components of CAST-J system are: (1) an operating system; (2) a set of
CAST-J programs written in a high-level programming language (C); (3) a set of CAST-J
lessons consisting of the spoken materials recorded on video discs, and of the lesson texts
stored on the computer with the programs.

In CAST-J, the user is provided with easy-to-follow instructions on screen. The format
of the screens and key functions are consistent throughout the lessons. The lesson texts are
read by the speaker on screen, either with or without audio signal. The trainee inputs his/her
response to the computer via the keyboard. The input is echoed on screen, in written form.
The feedback provided by the computer, as well, appears on screen in written form. The detail
of the lessons is described in the following chapter. For CAST-J, some modifications to the
format of the screens have been made to the English version of CAST, on the basis of the
study by Aldham, Rochford and Warling (1987), in order to make instructions even more
efficient. Further information regarding the operation of the system and the detail of the
instructions given on screen are available in the owner's manual of the English version of
CAST (Pichora-Fuller and Cicchelli, 1986) and in the study by Aldham et al. (1987).

Currently, various highly-sophisticated pieces of equipment are on the market, and their
usefulness for further research in speechreading appears promising. Sun Video, for example,
allows the user to have the video image on a workstation, or to combine video images with
computer-generated graphics. This type of apparatus may not be necessary for the regular
training of subjects in the clinic, but it will be invaluable for research purposes.

5.2. Input and feedback modes

There are several options regarding the keyboard mode and the feedback mode. The
following sections will discuss which mode is most appropriate for input via the keyboard and
for feedback on screen.

5.2.1. Kanji, kana, and romaji

As mentioned in the previous chapter, Japanese has three different writing systems:
kanji (ideographic script), kana (phonemic-syllabic script), and romaji (phonemic-alphabetic
script). Kana is further divided into hiragana and katakana. These different scripts are
combined in accordance with the function of the parts of the message. Kanjis (i.e. Chinese
characters adapted for use in Japanese) are used to distinguish homophonous words, which
can be differentiated neither by kana nor by romaji. In general, hiragana is used to write
anything that cannot be expressed with kanjis (e.g. verb inflections, function words), or to write
words whose kanji form is difficult to read or write correctly. Katakana is used to write
loanwords and proper nouns from foreign languages. Most printed matters issued in Japanese
include a combination of these three systems. Romaji, which is an alphabetization of kana, is
rarely used by the native speakers of Japanese for a purpose of general communication. This alphabetic script, however, may be used for street signs and/or for textbooks of Japanese specifically written for foreigners.

In addition to these conventional writing systems of Japanese, foreign words spelled in their original Western form are widely used in the Japanese society, usually for commercial and/or aesthetic reasons. There exist a number of magazines which are titled in English or in French (e.g. " more ", " croissant "). Although Japanese people are getting more and more familiar with the words written in the alphabet of the original language(s), they are not necessarily getting more familiar with the words written in the alphabet used for romaji. For instance, the English word " STOP " written as in English will be recognized by Japanese people as easily as the word written in kana or in kanji. However, when the word " STOP " is transliterated into romaji, i.e. *sutoppu* or *SUTOPPU*, or when the Japanese word meaning " to stop " is written in romaji, i.e. *tomare*, Japanese people will have difficulty determining what the word means.

Generally, in Japan, both the alphabetic mode and the kana mode are provided as input mode to computers. Although kanjis are also available, they require the conversion from either kana or romaji as well as the search of the target character from a number of homophones.

Kanjis are not suitable for the CAST-J lessons. In terms of input, it is important to keep to a minimum the time needed by a trainee to input his/her responses. In terms of feedback, one kanji usually corresponds to more than one haku, thus it is difficult to provide feedback to the trainee with viseme-size units.

This leaves kana and romaji as the better options for input and feedback. Advantages and disadvantages of each mode will be discussed below. As stated in the 'notation' section at the beginning of this study, *italics* denote romaji, whereas *bold type italics* denote kana, with an italic exclamation mark (!) indicating separations between kana characters. (cf. Notation, p. viii). For example, the Japanese word meaning " rain " in English is pronounced [ ame ], thus
it will be transliterated *ame* in romaji. In kana, the same word will be written with two kanas: one kana corresponding to the haku " a " and another kana corresponding to the haku " me "; thus, in this study, it will be written as *a!me*.

5.2.2. Input

As far as the selection of a keyboard is concerned, neither kana nor romaji can be said to be superior to the other in all situations. Advantages and disadvantages vary, depending on the trainee’s familiarity with either type of keyboard. For example, individuals who are used to typing letters in English may prefer romaji input to kana input. On the other hand, individuals who have used the kana mode from the very beginning will prefer kana input to romaji for the reverse reason.

In terms of the programming of the computer, the following five aspects have to be considered: (1) There are two particular kana characters, each of which corresponds to two pronunciations ( [ he ] and [ e ]; [ ha ] and [ wa ] ), depending on the syntactic context. In the case of [ he ] and [ e ], the former corresponds to a two-viseme sequence (consonant and vowel), while the latter corresponds to only one viseme. In the case of [ ha ] and [ wa ], the consonant visemes of these two hakus belong to different visemes: < h > belongs to \k\, while < w > belongs to \w\ (2) With kana input, the program would have to be sensitive to the size of each kana (full-size kana versus small-size kana), because the difference in size has phonemic significance in kana. (3) There is no adequate single Roman letter for the haku N. The upper case " N " cannot be used for the haku N, because in romaji, as in English, upper case characters are used in sentence initial position and for proper nouns. Lower case " n " is usually used for the haku N in romaji, but it is not possible to distinguish " na " ( *na* in kana and phonetically [ na ] ) from " na " ( *n/a* in kana and phonetically [ na ] ). This ambiguity can be resolved by placing a hyphen between " n " and the vowel " a " for [ na ], (i.e. "n-a"), and by leaving " na " ( [ na ] ) as it is. The hyphenation, however, may not be
a common usage for all trainees. (4) In romaji, the haku R (vowel elongation) has to be indicated by a macron (a horizontal bar on the top of the vowel being elongated) as in to5i (FAR). This particular symbol is not available on normal keyboards. The use of a colon after the vowel (e.g. to:i) is one possible solution to this problem. In kana, the haku R is replaced by an additional vowel, or by a dash, which indicates elongation of the preceding vowel. (5) As for the phonetic realization of vowel elongation, there is not necessarily a one-to-one correspondence between kana and romaji. For example, the Japanese word meaning "senior high school" in English will be pronounced [ koːkoː ], although in kana, it must be always written ko!u!ko!u. Spelling rules for romaji, however, are not as strict as those for kana. Some people may write the word as it is written in kana, thus koukou. Others may write the word as it is pronounced, thus ko5ko5 (or ko:ko: in the present study). Furthermore, some phonetically naive people may claim that the word is pronounced [ koukou ] instead of [ ko:ko: ], based on the strong association of their intuition with kana. In such cases, it is not possible to decide whether the trainee perceived the target viseme correctly but simply wrote it out in romaji following the kana convention, or whether the trainee uses o: and ou interchangeably, regardless of the visual perception of the target. Therefore, a new set of orthographic rules of romaji has to be adhered to, in order to establish a consistent input-feedback relationship.

Points (1) and (2) above can be looked at as disadvantages of the kana input, or advantages of the romaji input, while points (3) to (5) can be looked at as disadvantages of the romaji input, or advantages of the kana input.

5.2.3 Feedback

An advantage of using kana for feedback messages is that it is easy for the trainee to comprehend their meaning. Since a kana can represent any of the strings (j)V, CV, Q or N,
the trainee can decode a kana message more quickly than a romaji message. A disadvantage of kana feedback, however, is that it is difficult to provide viseme-level feedback: most of the time, one kana contains information about more than one viseme. For example, when a trainee responds incorrectly to the consonant viseme but correctly to the vowel viseme of a given CV target, it is impossible for the kana feedback to inform the trainee that his misperception was not the vowel but only the consonant. If the feedback provides the answer (target kana) even though the response is not completely right, the feedback is too 'generous'. On the other hand, if the feedback does not provide the answer until the trainee perceives all the visemes of a given target kana correctly, the feedback is too 'demanding'. This problem, however, may be solved by utilizing various techniques of screen display such as different colours and/or different appearance of lines. For example, a red underscore could indicate that the response for a given kana was incorrect for both the consonant and the vowel; a white underscore could indicate that only the consonant was incorrect for the given kana; finally, a yellow underscore could indicate that only the vowel was incorrect for the given kana. Alternatively, a doubled solid underscore could indicate that both the consonant and the vowel were identified incorrectly; a single solid line could indicate that the consonant was identified incorrectly; and a dotted line could indicate that the vowel was identified incorrectly.

The main advantage of romaji feedback is that it can provide viseme-level feedback. Its main disadvantage is that people are not used to reading romaji, thus it will take longer for a trainee to decode a feedback message, and possibly interfere in other ways with his/her learning.

As far as speechreading is concerned, there is no evidence for deciding which mode is more efficient or preferable for training Japanese subjects. For this first set of CAST-J lessons, romaji input and romaji feedback will be employed for the following reasons: (1) It is probable that a trainee can learn the layout of a keyboard without much difficulty along with the lessons. (2) Even for those who have experience only with the kana keyboard, it is advantageous to
become familiar with alphabetic keyboards, which are available for any computer. (3) New rules for romaji writing that have to be learned by a trainee are not complicated and limited in number. Therefore, proper instruction or the provision of a proper "list of rules" will help the trainee to learn and apply the rules correctly, without much difficulty. (4) Romaji feedback can provide viseme-level feedback most conveniently. (5) There is hope that a trainee may feel gradually less uncomfortable reading romaji messages as the training sessions progress. It is thus important for CAST-J, which is still at the pilot study stage, to continue to investigate how successfully trainees perform with different input and feedback modes.

Ideally, in future implementations of CAST-J, a trainee should be allowed to choose which input mode (kana versus romaji) he/she wants to use. Likewise, it should be possible to provide the 'demanding' kana feedback as default option, and to have the trainee choose later if he/she wants to see viseme-level feedback in romaji. In the case of romaji feedback, the trainee can see which viseme he/she has misperceived. Some trainees may already know which viseme is most likely to have been misperceived, while others may not have a clue. In the first case, the demanding kana feedback (more than one viseme together) may be sufficient, whereas additional romaji feedback may be necessary in the second case.

5.3. Orthographic rules for kana - romaji conversion

In order to have romaji input and feedback, it must be confirmed that several orthographic rules are known not only by the trainee, who inputs his/her responses to the computer in romaji, but also by the designer of the lessons (the author, in this study), who inputs target words/phrases to the computer. The stages where these orthographic rules are necessary are shown in Figure 2, indicated with arrows. Since not many Japanese people are used to writing Japanese in romaji, it is very important to establish orthographic rules for romaji, at least among the users of the CAST-J lessons, to avoid those input-feedback discrepancies which would be due to orthographic-level confusion rather than to the
The 'intended texts' of the lessons, which are given to the speaker of the lessons for video-recording, are first written in a combination of kana and kanji, so that the speaker has no uncertainty as to the contents. Although the speaker is expected to read the intended texts as accurately as possible, in the way it is written (e.g. make a pause at phrase boundaries, these being indicated by spaces), it is probable that, after the video-recording, the intended texts will need to be revised in accordance with the performance of the speaker. The revised texts will be transliterated into romaji and will become the actual texts of the lessons, following the orthographic rules for the designer (Appendix A). These romaji-form-texts will, then, be transferred to the computer via the keyboard. Similarly, the trainee will input his/her response to the computer via the keyboard, following a subset of orthographic rules identical to those for the designer (Appendix A).

The CAST-J program contains various other rules. These rules include the rules for romaji-to-visemic code conversion, both for target and response, the rules for target-response matching of visemic code, and the rules for providing feedback in romaji. A further description of these rules is provided in Chapter 6.
Figure 2. Writing Systems Used at Different Stages and Conversion Rules.

DESIGNER'S SIDE

INTENDED TEXT
(Kanji & kana)

(RECORDING)

VIDEO DISC
(Video & Audio)

(TRANSCRIPTION & REVISION)

REVISED TEXT
(Kana)


ACTUAL TEXT
(Romaji)


KEYBOARD
(Romaji)

COMPUTER PROGRAM
(CAST-J)

Target

Response

TRAINEE'S SIDE

PSYCHOLOGICAL REPRESENTATION OF RESPONSE STRING

RESPONSE STRING REPRESENTATION
(Kana or Romaji)

RESPONSE STRING
(Romaji)

KEYBOARD
(Romaji)

ORTHOGRAHPIC RULES FOR ROMAJI

Rules for romaji-visemic code conversion
Rules for target-response matching
Rules for feedback
CHAPTER 6

LESSONS

6.1. Course structure

There are 16 training lessons in CAST-J: 5 lessons for vowel visemes, 8 lessons for consonant visemes, and 3 lessons for haku visemes. The lessons progress from the easiest and least controversial viseme to the more difficult and controversial ones, for both vowels and consonants. Haku visemes, which are considered to be advanced and supplemental, come last. The lessons cover the 15 visemes as follows:

1. \p\ < p >, < b >, < m >
2. \a\ < a >
3. \i\ < i >
4. \u\ < ï >
5. \w\ < w >
6. \o\ < o >
7. \e\ < e >
8. \r\ < ŭ >
9. \t\ < t >, < d >, < n >, < s >, < dz >, < ts >
10. \k\ < k >, < g >, < h >,
11. \w\ < w >, < F > in loanwords
12. \f\ < f > in loanwords
13. \j\ < j >, < J >, < S >, < ts >, < dz >, < C >
14. \N\ haku N
15. \Q\ haku Q
16. \R\ haku R
Two other lessons (Lesson 17 and Lesson 18) are available. These 'general lessons' are designed to be used for pre- and post-evaluation of speechreading ability, or for further practice for the trainees who have completed the training lessons.

Lesson 1 through 10 constitute the minimum normal sequence: all trainees are expected to go through this set of lessons. These ten lessons include the five vowel visemes and most of the consonant visemes found in conventional Japanese. Lesson 11 through 16 are optional, i.e. the trainee should be allowed to skip some or all of these lessons when the lessons are too difficult for him/her, or when the instructor judges that these lessons are inadequate for the level of skill of the trainee.

6.2. Lesson structure

Basically, the four sections of each lesson are designed (1) to review previously learned visemes (old visemes), (2) to practice the recognition of a new viseme, in syllables, (3) to practice old and new visemes by the discourse tracking method, and (4) to recap the lesson materials, with a listing of the trainee's performance statistics.

6.2.1. Review of previously learned visemes

The reviewing of old visemes is accomplished by presenting the visemes in visual-only mode, in single-haku format (CV) for the consonant visemes, in two-haku format (VV) for the vowel visemes, and in three-haku format (VCCV) for the haku visemes. For this section, only one allophene is selected for each viseme, e.g. \( \langle p \rangle \) is used as a typical representative for the viseme \( p \). There is no review section for the first lesson and for the two general lessons. Viseme tokens are demonstrated and the trainee is simply asked to identify them. He/she may re-play the demonstration of a particular viseme any number of times.

6.2.2. Practice of a new viseme
Practising the recognition of a new viseme is accomplished by presenting pairs of visemes in visual-only mode, namely C1a - C2a single-haku contexts for consonant visemes, and V1a - V2a two-haku contexts for vowel visemes (V1i - V2i for the viseme \a\), where C1 or V1 always stands for the target viseme, while C2 or V2 stands for the same or another viseme. The order of the pairs is random. For the training of the haku visemes, two- and three-haku sequences are contrasted, e.g. anpa - apa (or anpa) for \N\; appa -apa (or appa) for \Q\; and a:i - ai (or a:i) for \R\. The trainee is required to discriminate whether the two members of a particular pair of visemes are the same or different. After all pairs have been judged, pairs responded to incorrectly are repeated until they are judged correctly. The trainee may re-play any pair any number of times until he/she feels confident that he/she can discriminate the two visemes.

After the discrimination task, all members of the target viseme are presented audio-visually. The trainee may re-play the demonstration any number of times.

6.2.3. Practice of old and new visemes in discourse

Practice of old and new visemes is accomplished by presenting one or two paragraphs. Each paragraph (text) is characterized by a high frequency of occurrence of the newly learned viseme. Words selected for the texts are words which have a statistically high frequency of occurrence in Japanese (Nakano, 1975; Sato, 1982; The National Language Research Institute, 1984). In this section of the lesson, the trainee is expected to integrate visual information of speech with linguistic redundancy and effective sender-receiver feedback in perceiving speech.

Each text is centered around a given topic, so that all sentences in the text are semantically related. All topics are considered to be within the everyday experience and linguistic level of the target population. The topic of each lesson is provided at the beginning of the text. Craig (1988) reports that word recognition performance was better for high-predictability sentences than low-predictability sentences and sentences with neutral-predictability
carrier phrase. He also reports that older adults can be misled by semantic cues just as well as by acoustic cues. In the study of Sparks, Ardell, Bourgeois, Wiedmer and Kuhl (1979), a significant change in speechreading performance was found between linguistically simple and linguistically complex sentences. The tracking performance of their subjects (the number of words correctly repeated per minute) dropped drastically, when the text was changed from linguistically simple material (e.g. novels written for children and young adults) to linguistically complex material (e.g. stories from magazines written for adults).

CAST-J includes both high- and low-predictability sentences, as well as a variety of syntactic structures and phrase lengths. The text for each lesson is divided into phrases of various length (major strings), which are predetermined by the designer. Each major string is numbered, and used as a set of target words/phrases. A major string is further divided into words or phrases (minor strings), which are also predetermined by the designer. In the text, minor strings are separated by spacing. Although an indication of word/phrase boundaries is not mandatory in Japanese writing, the use of 'minor string' is imperative for the computer program to provide better feedback: the computer can compare the target and the response more accurately for short strings of characters than for long ones. The text for each lesson consists of approximately 30 to 50 major strings. In Lesson 1 through 10, each lesson has two paragraphs, the first consisting of syntactically simple sentences and the second consisting of syntactically complex sentences. The syntactically simple sentences contain few embedded sentences, while the syntactically complex sentences contain a number of embedded sentences. In Lesson 11 through 16, each lesson has one single paragraph with sentences of various syntactic structures. Lesson 17 and Lesson 18 are also one-paragraph lessons. The paragraphs used in these two lessons are considered to be visemically well balanced, and to be equivalent in terms of occurrence of each viseme, syntactic complexity, and word familiarity. They are suitable for pre- and post-training comparison. (Further information is provided in Table VI, p. 62, and in Appendix D.)
Jackson (1988) suggests that data on coarticulatory effects be used to select contexts: one should proceed from easy-to-see contexts to hard-to-see contexts. In general, consonants are harder to perceive when followed by /u/ than when followed by /a/ or /ə/; vowels are harder to perceive when surrounded by consonants such as /f/, /v/, /s/, /tʃ/, /z/, /dz/ than in the vicinity of /h/ and /g/. Although various phonetic environments are included in CAST-J, more easy-to-see contexts are used for the more controversial visemes than for the less controversial visemes. For example, in Lesson 13 (Viseme \(j\)), \(C + /a/\) context in sentence initial position and /a/ + \(C + /a/\) context are frequently used, while in Lesson 1 (viseme \(p\)), different vowel contexts are used. In the case of viseme \(w\) \(< w >\) and \(< F >\), practice of \(< w >\) in the conventional Japanese haku \(wa\), is emphasized in Lesson 2, whereas practice of \(< w >\) followed by other vowels and practice of \(< F >\) is emphasized in loanwords in Lesson 12.

In CAST-J, in the third section of the lesson (practice of old and new visemes in discourse), the texts are presented in paragraph and/or phrase format, as in the English and the French versions of CAST. After practising recognition of a new viseme in the (previous) second section, the paragraph format is used for the introduction and the conclusion of the this section. The phrase format is used for the main part of the practice i.e. interactive training for speech perception.

The speaking rate of the texts (both at paragraph level and phrase level) can be selected. The reading of the texts is recorded at two speaking rates: slow normal speech and fast normal speech. The trainee or the instructor selects the appropriate rate for each lesson.

The modality of presentation can also be selected. Visual-only and audio-visual presentation are available. The trainee or the instructor can select the adequate modality for each lesson. The presentation of lessons can be made with or without amplification, and/or with or without background noise.

Both responses and feedback are always displayed on screen (worksheet), throughout the
lessons. Compared with the live face-to-face procedure, this display is helpful in two ways: (1) it eliminates the possible ambiguity of the response; (2) it eliminates the misunderstanding of feedback by the trainee, who, in the face-to-face situation, may have the same difficulty understanding the spoken feedback as understanding the target phrases.

For this particular section of the lesson, there is a ceiling to the number of times that the trainee is allowed to repeat the same target when his/her response is not correct. However, the trainee is given the freedom to skip a phrase and to come back to it later.

6.3. **Romaji - viseme conversion and target response matching for feedback**

In the tracking procedure used in CAST-J, feedback is provided to the trainee regarding the accuracy of his/her responses. Target string and response string are converted into visemic code, following the rules of romaji - viseme conversion given in Appendix B. Vowels are coded by alphabetic characters: A, I, U, E, and O; consonants are coded by numbers: 1 through 7; and the haku visemes \N\, \Q\, and \R\ are coded by 8, 9, and a colon (:), respectively. Feedback is provided on the basis of the visemic match between target and response.

The matching algorithm compares response and target, character by character. Visemes which have not yet been presented, in previous lessons or in the current one, are coded as C for consonants, /N/ and /Q/, and as V for vowels and /R/. When the trainee fails to identify the existence of a certain viseme, i.e. when he does not input any response for the target viseme, the feedback for this viseme is an underscore, regardless of the level of the viseme. A typical interaction in Lesson 8 may be as follows:

<table>
<thead>
<tr>
<th>Transaction</th>
<th>Presentation Mode</th>
<th>Item</th>
<th>Visemic code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target:</td>
<td>video image</td>
<td>yoso:sareru</td>
<td>7040:4A3E3U#</td>
</tr>
</tbody>
</table>
Response: keyboard/screen osareru 0 4A3E3U#
Feedback: screen _o__sareru _0__4A3E3U#
Response: keyboard/screen yoso:sareru 7040:4A3E3U#
Feedback: screen daihoyo:tekina 4AI70:4EI4A# (matched!)

However, when the trainee identifies the presence of a viseme but misidentifies it, i.e. puts the wrong viseme for the target, the feedback provides the answer for visemes which are not yet learned (C and V), but it only provides an underscore for previously learned and currently learned visemes. The following exchange illustrates this point for a trainee studying Lesson 5:

<table>
<thead>
<tr>
<th>Transaction</th>
<th>Presentation Mode</th>
<th>Item</th>
<th>Visemic code (not shown on screen)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target:</td>
<td>video image</td>
<td>shimainitottewa</td>
<td>7I1AI7I4094E2A# (C11A1C1C1VCCV2A#)</td>
</tr>
<tr>
<td>Response:</td>
<td>keyboard/screen</td>
<td>hikaritonde</td>
<td>5I5A 3I4084E #</td>
</tr>
<tr>
<td>Feedback:</td>
<td>screen</td>
<td>shi_a_nitotte__</td>
<td>5I_A_7I4094E__#</td>
</tr>
<tr>
<td>Response:</td>
<td>keyboard/screen</td>
<td>shidainitottewa</td>
<td>7I4AI7I4094E2A</td>
</tr>
<tr>
<td>Feedback:</td>
<td>screen</td>
<td>shi_ainitottewa</td>
<td>5I_AI7I4094E2A#</td>
</tr>
</tbody>
</table>

Thus, feedback is more demanding in later lessons, where the trainee is expected to know a certain number of visemes, than in earlier lessons, where the trainee is expected to identify only a few visemes.
In addition to a simple position-by-position match of the visemic code, a rightward search takes place. If the visemic code for the response does not match the visemic code for the target in the corresponding position, the matching program continues to check the next characters, until the end of the minor string (i.e. until the next word/phrase boundary). If a match is found, (a number or C for a consonant, a letter or V for a vowel), the search re-start from the visemic code which is immediately right of the one matched last. The following exchange illustrates this point for a trainee studying Lesson 9:

<table>
<thead>
<tr>
<th>Transaction</th>
<th>Item</th>
<th>Visemic code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target:</td>
<td>do:yara watashio</td>
<td>40:7A3A#2A4A7IO#</td>
</tr>
<tr>
<td>Response:</td>
<td>yaya toshio</td>
<td>7A7A# 4O7I0#</td>
</tr>
<tr>
<td>Feedback:</td>
<td>___ya_a ___t_shio</td>
<td>___7A7A#__4O7I0#</td>
</tr>
</tbody>
</table>

Although the matching provided by the system is not always ideal, it can provide feedback which is not too far from that could be given by an actual instructor.

6.4. Texts of the lessons

The material of each lesson is listed in Appendix C. The text is listed in the phrase format (major string) and its corresponding visemic transcription. The number of occurrences for the previously learned visemes, the target viseme, and the visemes which have not yet been learned, is listed in Table V, for each lesson individually. The haku visemes \N\ and \Q\ are included with the consonants, and the haku viseme \R\ is included with the vowels.
The viseme distribution for Lesson 17 and Lesson 18 is shown in Table VI. An English translation for the entire texts of Lesson 17 and Lesson 18 is given in Appendix D. The texts were translated literally, so as to reflect better the linguistic complexity of the original texts.
### Table V: Lesson-by-Lesson Inventory of Viseme Tokens

Distribution of visemes are given in number of token (columns 3 to 6), and in percent (column 8). For each pair of lines, the first one is for consonants, and second one for vowels.

<table>
<thead>
<tr>
<th>Lesson</th>
<th>Visemes studied in earlier lessons</th>
<th>Visemes studied in current lessons</th>
<th>Visemes studied in later lessons</th>
<th>Total number of visemes in lesson</th>
<th>Distribution in percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 \ p \</td>
<td>---</td>
<td>93</td>
<td>291</td>
<td>384</td>
<td>0 / 24 / 76</td>
</tr>
<tr>
<td></td>
<td>---</td>
<td>---</td>
<td>370</td>
<td>370</td>
<td>0 / 0 / 100</td>
</tr>
<tr>
<td>2 \ a \</td>
<td>47</td>
<td>---</td>
<td>326</td>
<td>373</td>
<td>13 / 0 / 87</td>
</tr>
<tr>
<td></td>
<td>---</td>
<td>91</td>
<td>232</td>
<td>323</td>
<td>0 / 28 / 72</td>
</tr>
<tr>
<td>3 \ i \</td>
<td>35</td>
<td>---</td>
<td>316</td>
<td>351</td>
<td>10 / 0 / 90</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>92</td>
<td>186</td>
<td>353</td>
<td>21 / 26 / 53</td>
</tr>
<tr>
<td>4 \ u \</td>
<td>54</td>
<td>---</td>
<td>342</td>
<td>396</td>
<td>14 / 0 / 86</td>
</tr>
<tr>
<td></td>
<td>153</td>
<td>95</td>
<td>154</td>
<td>402</td>
<td>38 / 24 / 38</td>
</tr>
<tr>
<td>5 \ w \</td>
<td>42</td>
<td>76</td>
<td>239</td>
<td>357</td>
<td>12 / 22 / 66</td>
</tr>
<tr>
<td></td>
<td>186</td>
<td>---</td>
<td>109</td>
<td>295</td>
<td>63 / 0 / 47</td>
</tr>
<tr>
<td>6 \ o \</td>
<td>47</td>
<td>---</td>
<td>248</td>
<td>295</td>
<td>16 / 0 / 84</td>
</tr>
<tr>
<td></td>
<td>160</td>
<td>71</td>
<td>70</td>
<td>301</td>
<td>53 / 24 / 23</td>
</tr>
<tr>
<td>7 \ e \</td>
<td>70</td>
<td>---</td>
<td>252</td>
<td>322</td>
<td>22 / 0 / 78</td>
</tr>
<tr>
<td></td>
<td>181</td>
<td>77</td>
<td>21</td>
<td>279</td>
<td>65 / 28 / 7</td>
</tr>
<tr>
<td>8 \ r \</td>
<td>58</td>
<td>81</td>
<td>214</td>
<td>353</td>
<td>17 / 23 / 60</td>
</tr>
<tr>
<td></td>
<td>301</td>
<td>---</td>
<td>21</td>
<td>322</td>
<td>93 / 0 / 7</td>
</tr>
<tr>
<td>9 \ t \</td>
<td>82</td>
<td>80</td>
<td>127</td>
<td>289</td>
<td>28 / 28 / 44</td>
</tr>
<tr>
<td></td>
<td>278</td>
<td>---</td>
<td>18</td>
<td>296</td>
<td>94 / 0 / 6</td>
</tr>
<tr>
<td>10 \ k \</td>
<td>173</td>
<td>86</td>
<td>84</td>
<td>343</td>
<td>51 / 25 / 24</td>
</tr>
<tr>
<td></td>
<td>320</td>
<td>---</td>
<td>19</td>
<td>339</td>
<td>94 / 0 / 6</td>
</tr>
<tr>
<td>11 \ w \</td>
<td>89</td>
<td>19</td>
<td>24</td>
<td>132</td>
<td>67 / 15 / 18</td>
</tr>
<tr>
<td></td>
<td>107</td>
<td>---</td>
<td>14</td>
<td>121</td>
<td>88 / 0 / 12</td>
</tr>
<tr>
<td>12 \ f \</td>
<td>84</td>
<td>24</td>
<td>25</td>
<td>133</td>
<td>64 / 18 / 18</td>
</tr>
<tr>
<td></td>
<td>104</td>
<td>---</td>
<td>18</td>
<td>122</td>
<td>85 / 0 / 15</td>
</tr>
<tr>
<td>13 \ j \</td>
<td>98</td>
<td>33</td>
<td>16</td>
<td>147</td>
<td>67 / 22 / 11</td>
</tr>
<tr>
<td></td>
<td>124</td>
<td>---</td>
<td>9</td>
<td>133</td>
<td>93 / 0 / 7</td>
</tr>
<tr>
<td>14 \ N \</td>
<td>157</td>
<td>40</td>
<td>4</td>
<td>201</td>
<td>78 / 20 / 2</td>
</tr>
<tr>
<td></td>
<td>149</td>
<td>---</td>
<td>16</td>
<td>165</td>
<td>90 / 0 / 10</td>
</tr>
<tr>
<td>15 \ Q \</td>
<td>155</td>
<td>37</td>
<td>---</td>
<td>192</td>
<td>81 / 19 / 0</td>
</tr>
<tr>
<td></td>
<td>147</td>
<td>---</td>
<td>9</td>
<td>156</td>
<td>94 / 0 / 6</td>
</tr>
<tr>
<td>16 \ R \</td>
<td>185</td>
<td>---</td>
<td>---</td>
<td>185</td>
<td>100 / 0 / 0</td>
</tr>
<tr>
<td></td>
<td>170</td>
<td>40</td>
<td>---</td>
<td>210</td>
<td>81 / 19 / 0</td>
</tr>
</tbody>
</table>
Table VI: Viseme Inventory for Lesson 17 and Lesson 18

Distribution of each viseme by count and percentage, in each lesson. Percentages are given in parentheses.

<table>
<thead>
<tr>
<th>Viseme</th>
<th>Lesson 17</th>
<th>Lesson 18</th>
</tr>
</thead>
<tbody>
<tr>
<td>\p \</td>
<td>22 (7.7)</td>
<td>19 (6.7)</td>
</tr>
<tr>
<td>\w \</td>
<td>7 (2.4)</td>
<td>7 (2.4)</td>
</tr>
<tr>
<td>\r \</td>
<td>13 (4.5)</td>
<td>12 (4.2)</td>
</tr>
<tr>
<td>\t \</td>
<td>33 (11.5)</td>
<td>33 (11.5)</td>
</tr>
<tr>
<td>\k \</td>
<td>25 (8.8)</td>
<td>29 (10.1)</td>
</tr>
<tr>
<td>\f \</td>
<td>1 (0.3)</td>
<td>2 (0.7)</td>
</tr>
<tr>
<td>\j \</td>
<td>25 (8.8)</td>
<td>22 (7.6)</td>
</tr>
<tr>
<td>\a \</td>
<td>27 (9.5)</td>
<td>32 (11.1)</td>
</tr>
<tr>
<td>\i \</td>
<td>29 (10.1)</td>
<td>29 (10.0)</td>
</tr>
<tr>
<td>\u \</td>
<td>23 (8.1)</td>
<td>25 (8.7)</td>
</tr>
<tr>
<td>\o \</td>
<td>30 (10.5)</td>
<td>29 (10.0)</td>
</tr>
<tr>
<td>\e \</td>
<td>25 (8.8)</td>
<td>23 (8.0)</td>
</tr>
<tr>
<td>\N \</td>
<td>12 (4.2)</td>
<td>11 (3.8)</td>
</tr>
<tr>
<td>\Q \</td>
<td>7 (2.4)</td>
<td>5 (1.7)</td>
</tr>
<tr>
<td>\R \</td>
<td>7 (2.4)</td>
<td>10 (3.5)</td>
</tr>
</tbody>
</table>

Total 286 (100.0) 288 (100.0)

Consonants 145 (50.6) 140 (48.6)
Vowels 141 (49.4) 148 (51.4)
CHAPTER 7

CONCLUSIONS

This study is the first attempt to design a computer-aided speechreading training system for Japanese (CAST-J). In order to implement it and to further refine it, several factors must continue to be investigated.

The first factor is the definition of the viseme, in relation to coarticulatory effects. It is known that some visemes are difficult to recognize in certain phonetic contexts, and that the visibility of visemes varies, depending on the speaker (Benguerel and Pichora-Fuller, 1982; Montgomery et al., 1987). Considering the significance of haku, or the strict (C)(j)V structure, in Japanese, it may be desirable to group visemes in accordance with their visemic context. For example, as assumed in this study, one could have a single bilabial viseme \( p \) for all phonetic contexts, i.e. for any combination with the five vowels. Instead of the alveo-dental viseme \( t \) assumed in this study, however, one could consider two separate visemes: one viseme \( t_1 \) could be used in those contexts where it is followed by /a/, /e/, or /o/, and another viseme \( t_2 \) could be used in those contexts where it is followed by /i/ or /u/.

Although this approach appears similar to that of Sato (1961), the difference between the two is that the approach described above still considers the viseme as a phoneme-size unit, while Sato considers it as a haku-size unit. Unfortunately, for Japanese, research on visemes and/or on coarticulation is extremely limited. In order to apply this approach to the definition of viseme, a close observation of coarticulatory effects are indispensable, and such a study must await the availability of data describing the coarticulatory effects of the preceding sound(s) as well as of the following one(s). At the same time, speaker variability and the effects of speaking rate should also be investigated closely.
The second factor is related to sociolinguistic changes in Japanese. Some sociolinguistic variations can be crucial for the categorization of visemes. For instance, the fricative consonant \[ f \], which is not included in the phonetic inventory of conventional Japanese, has been shown to occur in certain sociolinguistic groups. Since \[ f \] is considered to be one of the well-defined visemes in English (Benguerel and Pichora-Fuller, 1982), occurrence of this articulation in Japanese has a great influence on speechreading in Japanese. Rapid expansion of the use of loanwords in Japanese society and increasing opportunities to expose oneself to foreign languages may have a long-term impact on the phonetic inventory of Japanese.

The third factor is the adequacy of input and feedback mode, i.e. kana versus romaji. Although it is often suspected that Japanese subjects are more familiar with kana than with romaji, there is no evidence on how much the kana mode is superior to the romaji mode, or how the romaji mode can interfere with one’s expression and comprehension of the message.

The fourth factor is the relation between syntactic complexity of the text and speechreading performance. Past research has shown that linguistic complexity affects the speechreading performance (De Filippo and Scott, 1978; Sparks et al., 1979). However, a study of the detail of syntactic structure (such as the number of embedded sentences) or of the structure of relative clauses is not found in the literature. Further research in this area will be helpful for the development of adequate lesson materials and the selection of an appropriate group of subjects.

The fifth factor is the application of CAST-J to different populations. In the near future, serious consideration should be given to developing a similar system for children and young adults who are post-lingually deaf. Application of CAST-J to this group may simply require the adequate selection of vocabulary (for the text) that be suitable for the target population. Other potential target population for other versions of CAST-J include adults with a severe-to-profound hearing loss who cannot benefit from amplification, and adults and children with a congenital hearing loss. With a better understanding of speech perception and
language development in the above populations, CAST-J will make a contribution to the enhancement of their communicative ability.
BIBLIOGRAPHY

NOTE: Japanese references for which no English title has been provided by the original author(s) are listed with their Japanese title first, followed by an English translation by this author in parentheses.


APPENDIX A

Orthographic Rules for Kana - romaji Conversion

NOTE: These rules are applied by the designer of the lessons when the actual texts are written in romaji for computer input, after the speaker has read the intended texts written in kanji and kana. They are also applied by the trainee when he/she inputs his/her responses to the computer in romaji.

(1) Use ひつじんしき romaji.

(2) Place a word/phrase boundary (space), where the speaker made a pause.

(3) Always use a colon (:) for vowel elongation in loanwords.
   e.g.  しょ: (show)

(4) When transliterating a vowel sequence such as あ'あ, い'い, う'う, え'え, お'お,
   (a) re-write it as a vowel plus a colon (:],
   e.g.  お'し'し ---› お'し: (DELICIOUS)
   unless separation of the vowels is semantically important;
   in such a case,
   (b) re-write it as two vowels.
   e.g.  い'い ---› い (GOOD)

(5) When transliterating the kana string お'う,
   (a) re-write it as the first vowel plus a colon (]:) when in the context Cy_,
   e.g.  きょ'う ---› きょ: (TODAY)
   or when the phonetic realization of お'う is customarily [o:];
   e.g.  くお'くお' ---› く'くo: (SENIOR HIGH SCHOOL)
   (b) re-write it as おu when the separation of these two vowels is semantically
       important.
   e.g.  お'も'お' ---› お'もu (TO THINK)

(6) When transliterating the kana string え'い,
(a) re-write it as the first vowel plus a colon (e:) only in Lesson 16;
   e.g.  se!i!do ---+ se:do  (system)
(b) re-write with ei in all other cases.
   e.g.  e!i!ga ---+ eiga  (movie)

(7) When transliterating the kana string i!u,
   (a) re-write it as yu: in the context of to!_(_/ko/to);
   e.g.  to!i!u!ko!to!ni ---+ toyu:kotoni  (IT IS THAT...)
   (b) re-write with iu in other contexts.
   e.g.  mi!u!shi!na!u ---+ miushinau  (LOSE SIGHT OF)

(8) When transliterating the kana denoting the haku N,
   (a) re-write it as n followed by a hyphen when the haku N is followed by a
       (semi)vowel;
       e.g.  an-i  [ ani ]  (CARELESS) instead of ani  [ ani ]  (BROTHER)
            kon-ya  [ konja ]  (TONIGHT) instead of konya  [ kona ]  (non-sense word)
   (b) re-write it as n when followed by a consonant, or when in utterance final
       position.
       e.g.  sonna  (SUCH)
            ichiban  (BEST)

(9) When transliterating the kana for the haku Q, duplicate the following consonant.
   e.g.  nippon  (JAPAN)

   ikkai  (FIRST FLOOR)

The rules (8) and (9) are illustrated without giving the kana string, due to problems with their
transliteration.
APPENDIX B

Romaji - Visemic Code Conversion Rules

```c
viseme_rule visja[] = {
  /* replace code, matched string, no. chars to skip */
  /* a */ 'A', "a", 1,
  /* i */ 'I', "i", 1,
  /* u */ 'U', "u", 1,
  /* e */ 'E', "e", 1,
  /* o */ 'O', "o", 1,
  /* b */ 'B', "b", 1,
  /* c */ '9', "cch", 2,
  /* d */ '9', "dd", 1,
  /* f */ 'f', 1,
  /* g */ '9', "gg", 1,
  /* h */ '7', "hy", 2,
  /* j */ '7', "j", 1,
  /* k */ '9', "kk", 1,
  /* m */ '1', "m", 1,
  /* n */ '4', "ny", 1,
  '8', "nb", 1,
  '8', "nc", 1,
  '8', "nd", 1,
  '8', "nf", 1,
  '8', "ng", 1,
  '8', "nh", 1,
};```
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Audio</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>'8'</td>
<td>&quot;nj&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'8'</td>
<td>&quot;nk&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'8'</td>
<td>&quot;nm&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'8'</td>
<td>&quot;nn&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'8'</td>
<td>&quot;np&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'8'</td>
<td>&quot;nr&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'8'</td>
<td>&quot;ns&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'8'</td>
<td>&quot;nt&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'8'</td>
<td>&quot;nw&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'8'</td>
<td>&quot;nz&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'8'</td>
<td>&quot;n-&quot;</td>
<td>2,</td>
</tr>
<tr>
<td>'7'</td>
<td>&quot;ni&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'4'</td>
<td>&quot;na&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'4'</td>
<td>&quot;nu&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'4'</td>
<td>&quot;nc&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'4'</td>
<td>&quot;no&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'8'</td>
<td>&quot;n&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>/* p */</td>
<td>'9',</td>
<td>&quot;pp&quot;</td>
</tr>
<tr>
<td>'1'</td>
<td>&quot;p&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>/* r */</td>
<td>'3',</td>
<td>&quot;r&quot;</td>
</tr>
<tr>
<td>/* s */</td>
<td>'9',</td>
<td>&quot;ssh&quot;</td>
</tr>
<tr>
<td>'9'</td>
<td>&quot;ss&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'7'</td>
<td>&quot;sh&quot;</td>
<td>2,</td>
</tr>
<tr>
<td>'4'</td>
<td>&quot;s&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>/* t */</td>
<td>'4',</td>
<td>&quot;ts&quot;</td>
</tr>
<tr>
<td>'9'</td>
<td>&quot;tt&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'4'</td>
<td>&quot;t&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>/* w */</td>
<td>'2',</td>
<td>&quot;wa&quot;</td>
</tr>
<tr>
<td>'6'</td>
<td>&quot;wi&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'6'</td>
<td>&quot;we&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>'6'</td>
<td>&quot;wo&quot;</td>
<td>1,</td>
</tr>
<tr>
<td>/* y */</td>
<td>'7',</td>
<td>&quot;y&quot;</td>
</tr>
<tr>
<td>/* z */</td>
<td>'4',</td>
<td>&quot;z&quot;</td>
</tr>
<tr>
<td>/* : */</td>
<td>':',</td>
<td>&quot;::&quot;</td>
</tr>
</tbody>
</table>
APPENDIX C
Lesson Script and Corresponding Visemic Code

LESSON 1

Review

There is no review for Lesson 1.

Training

Visual Only Presentation:

\[
\begin{align*}
\text{pa} & \quad \text{pa} \\
\text{pa} & \quad \text{na} \\
\text{pa} & \quad \text{ya} \\
\text{pa} & \quad \text{sha} \\
\text{pa} & \quad \text{ra} \\
\text{pa} & \quad \text{wa} \\
\text{pa} & \quad \text{pya} \\
\text{pa} & \quad \text{ta} \\
\text{pa} & \quad \text{kya} \\
\text{pa} & \quad \text{ma} \\
\text{pa} & \quad \text{nya} \\
\text{pa} & \quad \text{ha} \\
\text{pa} & \quad \text{hya} \\
\text{pa} & \quad \text{ga} \\
\text{pa} & \quad \text{ja} \\
\text{pa} & \quad \text{bya} \\
\text{pa} & \quad \text{fa} \\
\text{pa} & \quad \text{mya} \\
\text{pa} & \quad \text{ka} \\
\text{pa} & \quad \text{gya} \\
\text{pa} & \quad \text{za} \\
\text{pa} & \quad \text{ba} \\
\text{pa} & \quad \text{cha} \\
\text{pa} & \quad \text{tsa} \\
\text{pa} & \quad \text{rya} \\
\text{pa} & \quad \text{da} \\
\text{pa} & \quad \text{sa}
\end{align*}
\]

Audio-visual Presentation:

\[
\begin{align*}
\text{pa} \\
\text{ba} \\
\text{ma}
\end{align*}
\]
SUMBA MIYAKO (PLACE TO LIVE)

1. Mishiranutochide
2. seikatsu hajimerubai,
3. mazu shinpaininarunoga
4. tabemonoto
5. kotobadato omoimasu.
6. Donnabashodemo
7. sumebamiyakoto
8. iwarerumonono,
9. bimyo:na tokorode
10. atarashi: mituni
11. najimezu,
12. komarukotoga arimasu.
13. Mottomo nipponwa
14. motomoto shimagunide
15. tan-itsuminzokukoka
16. demoarunode,
17. ippanshiminno
18. kurashiburiniwa
19. amari sawa
20. mitomeraremasen.
21. Mononomikata,
22. uketomekatamo
23. konpontekiniwa
24. kyo:tsu:no bubunga
25. o:ku miraremasu.
26. Mushiro mondainanowa
27. uminomuko:ni
28. utsurisumu baaidesho:
29. Mochiron ima
30. sekaino musubitsukiwa
31. koremadeninaku
32. missetsunai:arimashita.
33. Shinbun, terebinado
34. masukomino okagede
35. samazamana jo:ho:ga
36. mainichi
37. meto mimikara
38. to bikondekimasu.
39. Demo,
40. mattakubetsuno bunkani
41. to kekomunowa
42. totemo muzukashi monodesu.
NEZUMINO YOMEIRI (WEDDING OF A MOUSE)

1. Mukashi mukashi, 1U5A7I#1U5A7I#
2. okubukai yamano 05U1U5A7I#7A1A40#
3. arumurani A3U1U3A7I#
4. nezuminoyakoga 4E4U1I4007A50A#
5. kurashiteimashita. 5U3A7I4E1I1A7I4A#
6. Muraichibanno 1U3A17I1A840#
7. bijintoyu: 1I7I8407U:#
8. hyo:banno 70:1A840#
9. musumenotameni 1U4U1E40A1E7I#
10. konoyode ichiban 5040704E#1I7I1A8#
11. takumashiku 4A5U1A7I5U#
12. rippanahitoo 3I91A4A51400#
13. omukosanni 01U50A487I#
14. mukaetaito 1U5AE4AI40#
15. oyanezumiwa O7A4E4U1I2A#
16. kibo:shiteimashita. 5I10:7I4E1A7I4A#
17. Mazuhajimeni 1A4U5A7I1E7I#
18. erabaretanowa E3A1A3E4A402A#
19. ohisamadeshiha. O5I4A1A4E7I4A#
20. Ohisamawa, O5I4A1A2A#
21. kumoni o:waretara 5U107I#0:2A3E4A3A#
22. jibunwa 7I1U82A#
23. tachimachi 4A7I1A7I#
24. mienakunatte 1IE4A5U4A94E#
25. shimainasuyo 7I1AI1A4U70#
26. toi:mashiha. 40I:1A7I4A#
27. Demo, kumowa kumode 4B10#5U102A#5U104E#
28. bokawa kazeniwa 105U2A#5A4E7I2A#
29. isumo maketabakariiru. I4U10#1A5E4A1A5A3I3U#
30. Attoyumani A9407U:1A7I#
31. tobasarete 401A4A3E4E#
32. shimaimasukarane 7I1AI1A4U53A4A#
33. tonobemashita. 40401E1A7I4A#
34. Ippe, kazewa I910#5A4E2A#
35. donnamononimo 4084A10407I10#
36. bikutomoshinainowa 1I5U0A107I4AI402A#
37. kabedesuyo. 5A1E4E4U70#
38. Bokutowa 105U4O2A#
39. kurabemononi 5U3A1E1O407I#
40. narimasen-yo 4A3I1A4E8#70#
41. tokotaemashita. 40504AE1A7I4A#
42. Mottomo, 1094010#
43. nezumiga yowamino 4E4U1I5A#7O2A1I4O#
44. kabewa i.mashita.
45. Kimitachinezumini
46. masarumonowa
47. imasen-yo.
48. Bokuga
49. yononakade mottomo
50. subarashi:to omounowa
51. kimitachidesu.
LESSON 2

Review

Visual Only Presentation:

*pa*

Training

Visual Only Presentation:

```
ai  ui
ai  ei
ai  oi
ai  ii
ai  ai
ai  ui
ai  ei
ai  ei
ai  oi
ai  ii
ai  ai
```

Audio-visual Presentation:

```
ai
```
### Practice A

**ARU AMERIKAJIN (AN AMERICAN)**

1. Anekara kiita  
   A4E5A3A#5II4A#
2. hanashidearu.  
   5A44A7I4E#A3U#
3. Arutoshino aki,  
   A3U407I4O#A5I#
4. anewa amerikano  
   A4E2A#A1E3I5A4O#
5. shiraiino ieni  
   7I3IAI4O#IE7I#
6. asobini ittakotogaru.  
   A401I7I#I94A5O405AA3U#
7. Arunichiyobo bino asa,  
   A3U7I7I7O:1I4O#A4A#
8. meo samasuto  
   1E0#4A1A4O4O#
9. atarini  
   A4A3I7I#
10. panno kaoriga  
   1A84O#5A03I5A#
11. tadayotteita.  
   4A4A7094E#A4A#
12. Asagohanni okuretewa  
   A4A5O5A87I#05U3E4E2A#
13. taihento,  
   4AI5E84O#
14. awatete  
   A2A4E4E#
15. kaoo arai,  
   5A00#A3AI#
16. heyao deta.  
   5E7A0#4E4A#
17. Daidokorono  
   4AI405O3040#
18. doao akete,  
   4OA0#A5E4E#
19. aratto  
   A3A94O#
20. koeo ageteshimatta.  
   5OE0#A5E4E7I1A94A#
21. Okusamadenaku  
   05U4A1A4E4A5U#
22. dannasamaga arekore  
   4A84A4A1A5A#A3E5O3E#
23. asagohanno shitakuo  
   A4A5O5A84O#7I4A5U0#
24. shiteitanoda.  
   7I4EI4A4O4A#
25. Aneomiruto,  
   A4E01I3U4O#
26. sassoku akarui egaode  
   4A94O5U#A5A3UIE5A04E#
27. atsui ko:hi:o iretukureta.  
   A4UI#5O:5I:E0I#3E4E5U3E4A#
28. Mada atatakai panni  
   1A4A#A4A4A5AI#1A87I#
29. bata:to  
   1A4A:4O#
30. amai jamuo  
   A1AI#7AIU0#
31. tappuri nuttekureta.  
   4A91U3I#4U94E5U3E4A#
32. Okusamawa  
   05U4A1A2A#
33. mada arawareni.  
   1A4A#A3A2A3E4AI#
34. Hajimewa,  
   5A7I1E2A#
35. tsunamichi atamagawa  
   4U1A7I#A4A1A5A#
36. agaranainokato  
   A5A3A4A4I5O5A4O#
37. omottaga,  
   01O94A5A#
38. mainichiyo bino  
   1AI7I7I7O:1I4O#
39. asagohanzukariwa  
   A4A5O5A84U5U3I2A#
40. kareni tottewa  
   5A3E7I#4O94E2A#
41. makkusu atarimaeno  
   1A94A5U#A4A3I1AE4O#
42. kotorashi:  
   50403A7I: #
43. Anewa,  
   A4E2A#
44. konotokibakariwa,
45. amerikajintono
46. kekkonni akogaretarashi:
47. Ma:, imadewa
48. akirametayo:dearuga.

50404051A5A3I2A#
A1E3I5A7I84O4O#
5E95087I#A5O5A3E4A3A7I:#
1A:#I1A4E2A#
A5I3A1E4A7O:4EA3U5A#
1. O: mukashi,  O:1U5A7I#
2. sagakuninawa moji ga  4A5A5U7I7I2A#107I5A#
3. arimasendeshita.  A3I1A4E84E7I4A#
4. Hajimete tsukawareta moji wa,  5A7I1E4E4#4U5A2A3E4A#107I2A#
5. imakara  11A5A3A#
6. nisennen chikakumo ma ni,  7I4E84E8#7I5A5U10#1AE7I#
7. chu: gokukara tsuwawat t a  7U:505U5A3A#4U4A2A94A#
8. kanjideshita.  5A87I4E7I4A#
9. Furukukara  6U3U5U5A3A#
10. hanasaretekita  5A4A4A3E4E5I4A#
11. kotobao  50401AO#
12. kakiarawasutameni  5A5I3A3A2A4U4A1E7I#
13. atarashiku kanjiga  A4A3A7I5U#5A87I5A#
14. a tehameraremashita.  A4E5A1E3A3E1A7I4A#
15. Jidaiga kawari,  7I4AI5A#5A2A3I#
16. yagate kanaga  7A5A4E#5A4A5A#
17. hatsumeisaremashita.  5A4U1EI4A3E1A7I4A#
18. Kananiwa  5A4A7I2A#
19. hiraganato katanaga  5I3A5A4A40#5A4A5A4A5A#
20. arimasuga,  A3I1A4U5A#
21. dochiraga hayaku  407I3A5A#5A7A5U#
22. tsukuraretanokawa  4U5U3A3E4A405A2A#
23. amari yokuwa  A1A3I#705U2A#
24. wakatteimasen.  2A5A94E1I4A4E8#
25. Shikashinagara,  7I5A7I4A5A3A#
26. imano katch ngi  11A4O#5A4A7I5A#
27. dekiagaru kateiniwa,  4E5I5A5A3U#5A4E7I7I2A#
28. kanarina gai jikanga  5A4A3I4A5AI#7I5A85A#
29. kakattakotowa  5A5A94A50402A#
30. tashikadesu.  4A7I5A4E4U#
31. Hiraganawa,  5I3A5A4A2A#
32. dochirakatoieba,  407I3A5A40IE1A#
33. onnanohitoga, tegami  084A4O5I4O5A#4E5A1I#
34. aruwa wakao  A3UI2A#2A5A0#
35. kakunoni tsukare shamashita.  5A5U407I#4U5A2A3E1A7I4A#
36. Ip po:, katakanawa,  I91O:#5A4A5A4A2A#
37. otokonohitoga  04O5O4O5I4O5A#
38. kanbun-o yomutameni  5A81U8#0#701U4A1E7I#
39. kangaedasarreta moji desu.  5A85AE4A3E4A#107I4E4U#
40. Kanamojino okagede,  5A4A107I40#5O5A5E4E#
41. kanarinagai jikanga  2A3E3A5A#404E84A7I4O#
42. yomikakino chikarawa  701I5A5I40#7I5A3A2A#
43. hi j o:ni takamarimashita.  5I7O:7I#4A5A1A3I1A7I4A#
LESSON 3

Review

Visual Only Presentation:

$aa$

$pa$

Training

Visual Only Presentation:

$ia$ $oa$

$ia$ $ea$

$ia$ $ia$

$ia$ $aa$

$ia$ $ua$

$ia$ $ia$

$ia$ $ua$

$ia$ $oa$

$ia$ $ea$

$ia$ $aa$

Audio-visual Presentation:

$ia$
1. Kyo:towa nishinihonno
570:402A#71717150840#
2. daihyo:tekina kanko:chino
4AI70:4E514A#5A850:7I40#
3. hitotsudesu.
5I404U4E4U#
4. Nagai rekisio mochi,
4AI5AI#3E517I0#107I#
5. machino naigai
1A7I40#4AI5AI#
6. itarutokoroni
I4A3U05O307I#
7. kicho:na bunkaisanga
5I70:4A#1U85AI4A85A#
8. nokosareteimasu.
40504A3E4E1A4U#
9. Shinkansenno hikarigo:nara
7I85A8E840#5I5A3I50:4A3A#
10. Tokyo:ekikara
40570:E5I5A3A#
11. sanjikaninaini tsukimasu.
4AI87I5AI7I#4U5I1A4U#
12. Furuitodo:jini
6U3UI4040:7I7I#
13. kindaitoshidemo
5I84AI407I4E10#
14. arimasukara
A3I1A4U5A3A#
15. wakai hitotachinimo
2A5AI5I4O4A7I7I10#
16. hipo:ni
5I70:7I#
17. ninkiga arimasu.
7I85I5A#A3I1A4U#
18. Ichinen-o tsujite
I7I4E8#0#4U:7I4E#
19. iroirona gyro:jiga
I3O13O4A#570:7I5A#
20. toriokonawaremasu.
403I0504A2A3E1A4U#
21. Aomatsuri,
AOI1A4U3I#
22. Jidaimatsuri,
7I4A11A4U3I#
23. Daimonjyiakiwa
4AI1087I7A5I2A#
24. ippanno hitonimo
I91A840#5I407I10#
25. hiroku shirareteimasu.
5I3O5U#7I3A3E4E1A4U#
26. Ichidodei:kara
I7I4O4EI:5A3A#
27. jissaini mitemiruto
7I94AI7I#1I4E1I3U40#
28. i:desho:
I:4E70:#
29. Kyo:toniwa
570:407I2A#
30. shikioriorno
7I5I03O3I4O#
31. utsukushisaga ari,
U4U5U7I4A5A#A3I#
32. donokisetsaga itemo
40405I4E4U7I#194E10#
33. kokoro hikarezuniwa
505O3O#5I5A3E4U7I2A#
34. iraremasesu.
I3A3EI4A4E8#
35. Toriwake
403I2A5E#
36. momijino jikawa
10117I40#7I5I2A#
37. insho:tekidesu.
1870:4E5I4E4U#
38. Akiniwa
A5I7I2A#
39. sumikitta ku:kinonaka
4U1I5I94A#5U:5I404A5A#
40. achikochi
A7I507I#
41. saikuringusurunomo
4AI5U3I85U4U3U4O10#
42. tanoshiku
4A407I5U#
43. kimochinoi.monodesu.
5I107I40I:10404E4U#
| 44. | Watashino shiraiwa          | 2A4A7I4O#7I3IAI2A# |
| 45. | itsumo ujikintokini         | I4U1O#U7I5I4O5I7I# |
| 46. | shitatsuuzumio uchi,        | 7I4A4U4U1O#U7I#   |
| 47. | kiyomizuyaki,              | 5I70I4U7A5I#      |
| 48. | nishijin-ori,              | 7I7I7I8#O3I#     |
| 49. | soshite yatsuhashio        | 407I4E#7A4U5A7I0# |
| 50. | omiyageni kaimasu.         | 01I7A5E7I#5AI1A4U# |
Practice B

HIKO:KI (AN AIRPLANE)

1. Hiko:kiwa ochirukara
2. kiraidatono
3. ikenga o:i.
4. Ririkushitakeara
5. chakurikusurumadewa
6. ikitakokochiga
7. shinaitoyu:
8. koemo
9. hinpanni kikareru.
10. Tashikani
11. ittan nanikaga
12. okiruto
13. inochio ushinau
14. kakuritsuwa takai.
15. Ichikabachikano
16. seishitsuga
17. tsuyoinowa
18. jijitsuda.
20. mainichi ikutsumo
21. okiteiru
22. jido:shajikono
23. hindoni
24. hikakushitara
25. ko:ki:ki tsuirakuno
26. hanashiwai
27. mettani kikanai.
28. Jitsuni anshinshite
29. riyo:dekiru
30. ko:tsu:kikanto ierunoda.
31. Hiko:kiwa tsugitsugini
32. seino:no takai
33. shingataga
34. kaihatsusareteiru.
35. Atarashi: enjin-o
36. riyo:shi,
37. iroirona
38. kaizen-o hete
39. tsukuriagerarea
40. yori hayai binga,
41. ima sekaino
42. kakutoshimezashite
43. tondeiru.
LESSON 4

Review

Visual Only Presentation:

aa
ia
pa

Training

Visual Only Presentation:

ua   oa
ua   ea
ua   ia
ua   aa
ua   ua
ua   ia
ua   ua
ua   ea
ua   oa
ua   aa

Audio-visual Presentation:

ua
1.  Atsui natsuga sugi, A4UI#4A4U5A#4U5I#
2.  kugatsumo 5U5A4U10#
3.  chu:junni naruto 7U:7U87I#4A3U40#
4.  akino kehaiga A5I40#5E5AI5A#
5.  kokunaru. 50504A3U#
6.  Tentakaku umakoyuru kisetsuno 4E84A5A5U#U1A507U3U#5I4E4U40#
7.  otozuredearu. 0404U3E4EA3U#
8.  Masshiroi kumoga 1A97I3OI#5U105A#
9.  suna aozorani ukabu. 4U84A#A0O3A7I#U5A1U#
10. Sujigumo, urokogumo, 4U7I5U10#U3O505U10#
11. iwashigumo nadodearu. 12A715U10#4A404EA3U#
12. Sugasugashi: tenkiga 4U5A4U5A7I:#4E85I5A#
13. tsuzukutame, 4U4U5U4A1E#
14. supo:tsuni 4U10:4U7I#
15. aseonagasuhitoga A4E04A5A4U5I405A#
16. tokuni medatsu. 40507I#1E4A4U#
17. Zenkokukakuchide, 4E8505U5A5U7I4E#
18. undo:kaiga okonawaru. 0840:5A15A#O504A2A3E3U#
19. Mata, ko:rokuchiwa 1A4A#50:3A5U7I2A#
20. kazokuzureya 5A405U4U3E7A#
21. guru:pudef 5U3U:1U4E#
22. nigiwau. 7I5I2AU#
23. Ku:kimo oishiku 5U:5I1O#O17I5U#
24. shokuyokumo waku. 7050705U10#2A5U#
25. Kudamonoohajime 5U4A104005A7I1E#
26. akino mikakuga A5I40#155A5U5A#
27. kakushu ho:huni desorou. 5A5U7U#50:5U7I#4E8303U#
28. Shufutachiwa, 7U6U4A7I2A#
29. matsutakeya 1A4U4A5E7A#
30. kurio tsukatte 5U3I0#4U5A94E#
31. ryo:ri no 370:3I4O#
32. udeofuruu. U4E0U3UU#
33. Geijutsusakuhinni 5E17U4U45U5I87I#
34. fureru chansumo 6U3E3U#7A84U10#
35. o:kunaru. 0:5U43A3U#
36. Ongakukai, 085A505AI#
37. bijutsutenbergodo 17U4U4E84A4010#
38. issuninaku konzatsusuru. I4U7I4A5U#5084A4U4U3U#
39. Kuwaete, 5U2AE4E#
40. dokushonetsumo 4O5U704E4U10#
41. ky:ni takamaru. 57U:7I#4A5A1A3U#
42. Yoruwa tsukiga 7O3U2A#4U5I5A#
43. akaruku kagayaku. A5A3U5U#5A5A7A5U#
Mangetsuno utsukushisawa kakubetsu dearu.
Shizentono fukai musubitsukio itsumademo ushinawazuni itaimonodearu.
ANRI = FA:BURU (Jean Henri Fabre)

1. Konchu:kide 5087U:5I4E#
2. yokushirareteiru 705U7I3A3E4EI3U#
3. Fa:buruwa, 6A:1U3U2A#
4. Furansuno 6U3A84U4O#
5. mazushi: no:funo 1A4U7I:#40:6U40#
6. musukotoshite umareta. 1U4U5O407I4E#U1A3E4A#
7. Itsunadetattemo I4U1A4E4A94E10#
8. karenouchino 5A3E40U7I4O#
9. kurashimukiwa 5U3A7I1U5I2A#
10. hitotsumo rakuniwa 5I4O4U10#3A5U7I2A#
11. naranakatta. 4A3A4A5A94A#
12. Kurushi: seikatsuno 5U3U7I:#4E5A4U4O#
13. renzokunimokakavarazu, 3E8405U7105A5A2A3A4U#
14. karewa gakumonde 5A3E2A#5A5U1084E#
15. miotatetai toyu: 1I04A4E4AI#407U:#
16. tsuouikibio: 4U70I5I1O:0#
17. isumo motteita. I4U1O#1094E14A#
18. Fa:buruga zutto 6A:1U3U5A#4U94O#
19. idakitsuzuketeita yumewa I4A5I4U5E4I4A#7U1E2A#
20. su:gakushani 4U:5A5U7A7I#
21. narukotodeaatta. 4A3U5O404EA94A#
22. Shikashi Faburuwa, 7I5A7I#6A1U3U2A#
23. jitsuwa 7I4U2A#
24. do:butsugaku, 40:1U4U5A5U#
25. shokubutsugakunitsuitemo 705U1U5A5U7I4U1I4E10#
26. sugureta saino:no 4U5U3E4A#4AI4O:40#
27. mochunishideaatta. 107I4U7I4EA94A#
28. Aruhi, A3U5I#
29. aru yu:meina A3U#7U:1EI4A#
30. do:butsgakushakara 40:1U4U5A5U7A5A3A#
31. konchu:no kenkyu:o 5087U:4O#5E857U:0#
32. susumerareta karewa, 4U4U1E3A3E4A#5A3E2A#
33. mayoukotonaku 1A70U5O404A5U#
34. mashino kansatsun A7I4O#5A84A4U7I#
35. torikunda. 403I5U84A#
36. Katatsumuri, matsumushi, 5A4A4U1U3I#1A4U1U7I#
37. mitsubachinadoo 1I4U1A7I4A4O0#
38. kuruhimo kuruhimo 5U3U5I1O#503U5I1O#
39. mitsuzuketa. 1I4U4U5E4A#
40. Hitotshuitotsuno 5I4O4U5I4O4U4O#
41. tsumikasanega 4U1I5A4A4E5A#
42. sanju:nengo 4A87U:4E85O#
43. jukkanno 7U95A84O#
44. konchu:kitoshite
45. shuppansareta.
46. Senkyu:hyakuju:gonen,
47. kagakunoshijinto
48. utawareta Fa:buruwa
49. kyu:ju:nisaino sho:gaini
50. shu:shifuo utta.

5087U:5I4O7I4E#
7U91A84A3E4A#
4E857U:7A5U7U:5O4E8#
5A5A5U4O7I7I84O#
U4A2A3E4A#6A:1U3U2A#
57U:7U:7I4AI4O#7O:5AI7I#
7U:7I6U0#U94A#
LESSON 5

Review

Visual Only Presentation:

aa
ia
ua
pa

Training

Visual Only Presentation:

wa pa
wa sha
wa ra
wa wa
wa pya
wa ta
wa kya
wa ma
wa nya
wa ha
wa hya
wa ga
wa ja
wa bya
wa fa
wa mya
wa ka
wa gya
wa za
wa ba
wa cha
wa tsa
wa rya
wa da
wa sa
wa ya
wa na

Audio-visual Presentation:

wa
WATASHINO KAZOKU (MY FAMILY)

1. Watashino iewa 2A4A7I40#1E2A#
2. Kanagawaken 5A4A5A2A5E8#
3. Fujisawashini aru. 6U7I4A2A7I7I#A3U#
4. Ryo:shinwa futaritomo 370:7I82A#6U4A3I4O1O#
5. sho:waumaredo, 7O:2A01A3E4E#
6. kodomowa watashito 5O401O2A#2A4A7I4O#
7. futatsuchigaino imo:to. 6U4A4U7I5AI4#I1O:4O#
8. Chichiwa 7I7I2A#
9. Wakayamaken shusshin. 2A5A7A1A5E8#7U97I8#
10. Hahawa Ishikawakenno 5A5A2A#I7I5A2A5E84O#
11. Wajima shusshin. 2A7I1A#7U97I8#
12. Chichino shigotowa 7I7I4O#7I504O2A#
13. kiwamete tenkinga o:ku, 5I2A1E4E#4E85I85A#O:5U#
14. watashitachiwa 2A4A7I4A7I2A#
15. kitawa hokkaido:no 5I4A2A#5O95AI4O:4O#
16. Wakkanai, 2A95A4AI#
17. minamiwa Okinawemade 1I4A1I2A#05I4A2A1A4E#
18. mawatta. 1A2A94A#
19. Chichiwa 7I7I2A#
20. wakarazuyadaga, 2A5A3A4U7A4A5A#
21. hahawa wariai 5A5A2A#2A31AI#
22. monowakairiga i:. 1O4O2A5A3I5A#I1:
23. Wakaikorowa 2A5A15O3O2A#
24. sorewa sorewa 4O3E2A#4O3E2A#
25. kawairashikattato 5A2A13A7I5A94A4O#
26. hitono uwasanikiku. 5I4O4O#U2A4A7I5I5U#
27. Demo wareware 4E10#2A3E2A3E#
28. shima:nitottewa 6U4A2A7I7I4O94E2A#
29. kowai sonzaideshika nakatta. 5O2A1I4O84AI4E7I5A#4A5A94A#
30. Ma:, warerano chichiyo:aniwa 1A:#2A3E3A4O#7I7I07A7I2A#
31. fusawashi: hitodato omowareru. 6U4A2A7I7I:#5I4O4A4O#O102A3E3U#
32. Waga imo:towa 2A5A#I1O:4O2A#
33. mawari:jukara 1A2A3I7U5A3A#
34. kawai:gararetebarari. 5A2A15A3A4E1A3A3I#
35. Iwayuru wagemamamamumuseda. 12A7U3U#2A5A1A1A1U4U1E4A#
36. Monowasureno hidosa:wa 1O4O2A4U3E4O#5I4O4A2A#
37. niwatoridesae 7I2A4O3I4E4AE#
38. kanawanaidaro:to 5A4A2A4AI4A3O:4O#
39. omowaseruhododa. 0102A4E3U5O4O4A#
1. Chu: gokudewa mukashi 7U: 505U4E2A#1U5A7I#
2. wagakunino kotoo 2A5A5U7I40#50400#
3. "wa" toyon de imashita. "2A" 407084E11A7I4A#
4. "Wajin" toieba, "2A7I8" 40IE1A#
5. "wa" nokuninisumu "2A" 405U7I7I4U1U#
6. warewarenokotoo 2A3E2A3E4050400#
7. sashitawakedesu. 4A7I4A2A5E4E4U#
8. Tsukawareru mejikoso 405A2A3E3U#107I5O4O#
9. kawarimashitaga, 5A2A3I1A7I4A5A#
10. "wa" notsuku kotobawa "2A" 404U5U#50401A2A#
11. nihonfu: toyu: imio 7I5O86U:#407U:11IO#
12. arawasukotoga A3A2A4U5O4O5A#
13. wakarimasu. 2A5A3I1A4U#
14. Imawa II2A2#
15. "heiwa" no "wa" nojiga "5E12A" 4O#"2A" 4O7I5A#
16. tsukawareteimasu. 4U5A2A3E4E11A4U#
17. Korewa, iroirona kotobatono 503E2A#I30I304A#50401A4040#
18. kumiwasega 5U1IA2A4E5A#
19. kano: na mojidesu. 5A4O:4A#107I4E4U#
20. Tabemonoo arawasutokiwa 4A1E10400#A3A2A4U4O5I2A#
21. washoku. 2A705U#
22. Kimonono baiwa 5I1O4O40#1AAI2A#
23. wafuku. 2A6U5U#
24. Hokiniwa, 5O5A7I2A#
25. waka, wabun, washitu, 2A5A#2A1U8#2A7I4U#
26. wasai, wagashi. 2A4AI#2A5A7I#
27. Kuwaete, saikindewa, 5U2AE4E#4A15I84E2A#
28. wafu: yo: gashi toka 2A6U:70:5A7I#4O5A#
29. yo: fu: wagashi, 70:6U:2A5A7I#
30. aruiwa A3U12A#
31. wafu: yo: shoku, 2A6U:70:7O5U#
32. yo: fu: washokunado 70:6U:2A7O5U4A4O#
33. wakenowakaranai monoga 2A5E4O2A5A3A4AI#104O5A#
34. demawanetteimasa. 4E1A2A94E11A4U#
35. Aruhitoni iwasereba, A3U5I4O7I#2A4E3E1A#
36. koremo wayo: secchu: to 503E10#2A7O:4E97U:4O#
37. yu: kotorashi: 7U:5O4O3A7I#:7:
38. "Wa" toyu: kotobawa "2A" 4O7U:5O4O1A2A#
39. sunawachi 4U4A2A7I#
40. cho: watoyu: kotonoi 7O:2A4O7U:5O4O7I#
41. naruwakeka. 4A3U2A5E5A#
LESSON 6

Review

Visual Only Presentation:

\[
\begin{align*}
aa & \quad pa \\
i a & \quad wa \\
u a &
\end{align*}
\]

Training

Visual Only Presentation:

\[
\begin{align*}
o a & \quad ua \\
o a & \quad ea \\
o a & \quad oo \\
o a & \quad ia \\
o a & \quad aa \\
o a & \quad ua \\
o a & \quad aa \\
o a & \quad ia \\
o a & \quad oo \\
o a & \quad ea
\end{align*}
\]

Audio-visual Presentation:

\[
o a
\]
KYODO TENKI (WEATHER FORECAST)

1. Zenkoku no minasama 4E8505U40#1I4A4A1A#
2. ohayo gozaimasu 05A70:504A1A4U#
3. Kyono otenkio 570:40#04E851O#
4. otsuaitashimasu 04U4AEI4A7I1A4U#
5. Kyono zenkokutekini 570:10#4E8505I4E517I#
6. yoiotenkito 70I04E8514O#
7. narumikomidesu 4A3U1I501I4E4U#
8. Yoso saiko kionwa 7040:4AI5O:5I082A#
9. To kyoga nijuyo yondo 40:570:5A#7I7U:7084O#
10. o sakaga onajiku 0:4A5A5A#04A7I5U#
11. nijuyo yondo 7I7U:7084O#
12. fukuokaga 6U5U05A5A#
13. nijuyo gododesu 7I7U:50404E4U#
14. Ko suikakuritsuwa 5O:4U15A5U3I4U2A#
15. gozen gototomoni 504E8#505040107I#
16. zero kara 4E305A3A#
17. gopa sentodesu 501A:4E8404E4U#
18. Kononotokoro 50504040503O#
19. zenkokutekini 4E8505U4E517I#
20. odayakanayo kiga 04A7A5A4A7O:5I5A#
21. tsuzuiteimashita ga 4U4UI4E1A7I4A5A#
22. koremo do yara 5O3E10#40:7A3A#
23. konshuippainoyo desu 5087U:191A4O7O:4E4U#
24. Nihon retto o ootteita 7I5083E94O:0#0094E14A#
25. ko kiat suga 5O:5IA4U5A#
26. higashini nuke 5I5A7I7I#4U5E#
27. nishinoho kara 7I7I4O5O:5A3A#
28. otenki o habani 04E85I#O:5A1A7I#
29. kuzurerumikomidesu 5U4U3E3U1I501I4E4U#
30. Kionmo godohodo 5I0810#50405040#
31. hikukunaru to no 5I5U5U4A3U4O4O#
32. yoso desu 7040:4E4U#
33. Soroso 40304030#
34. oshimeriga hoshi to 071I3I5A#5O7I:4O#
35. onowareitakorodesuga 0102A3E4E4I4A503O4E4U5A#
36. kisho cho no 5I7O:7O:4O#
37. ho kokuniyorimasuto 5O:505U7I7O3I1A4U4O#
38. tokoronyottowa 4O5O3O7I7O942A#
39. o mameninaru osoremo 0:1A1E7I4A3U#04O3E10#
40. arutoyu kotodesunode A3U4O7U:5O40E40404E#
41. do zo gochu ikudasai 4O:4O#5O7U:15U4A4AI#
KOINOBORI (CARP STREAMERS)

1. Gogatsuitsukawa
2. kodomonohide oyasumidesu.
3. Konohiwa
4. "tangonosekku"toyobareru
5. otokonokono omatsuridemoarimasu.
6. Kintaro:nadono
7. oningyo:o kazari,
8. koinoborio tatete
9. oiwaishimasu.
10. Iroazayakana
11. o:kinakoiga
12. aozorao ikiioyoku
13. oyogu yo:sowa,
15. Musukoya magono
16. sukyakana seicho:o negau
17. chichiyoa, hahaoyo,
18. oji:chan, oba:channo
19. koko:ro goko wakarimasu.
20. Tokorode,
21. onnanokoga yorokobu
22. omatsuriwa
23. nikagetsuhodo
24. hayaku okonawareru
25. momonosekudesu.
26. Kochirano oiwaika,
27. go:kana ohinasamao
28. kazarimasu.
29. Hinaningyo:niwa
30. odairisamato ohimesamanohokani,
31. sanninkanjoya
32. goninbayashinadoo
33. fukumukotoga o:idesu.
LESSON 7

Review

Visual Only Presentation:

\begin{align*}
aa & \quad pa \\
ia & \quad wa \\
ua & \\
oa & 
\end{align*}

Training

Visual Only Presentation

\begin{align*}
ea & \quad ua \\
ea & \quad ea \\
ea & \quad oa \\
ea & \quad ia \\
ea & \quad ea \\
ea & \quad aa \\
ea & \quad ia \\
ea & \quad aa \\
ea & \quad oa \\
ea & \quad ua 
\end{align*}

Audio-visual Presentation:

\begin{align*}
ea & 
\end{align*}
1. Watashiwa bo:ekigaishao 2A4A7I2A#10:E5I5A17AO#
2. keieishiteiru. 5EIE17I4E13U#
3. Eikokuni EI5O5U7I#
4. dekakerukotoga o:ku, 4E5A5E3US0405A#O:5U#
5. ieni irunowa IE7I1I3U402A#
6. kazoerukuraida. 5A40E3U5U3AI4A#
7. Nippon-e 7I9108#E#
8. kaettatokiwa 5AE94A405I2A#
9. kazokueno sa:bisuo 5A405UE40#4A:1I4U0#
10. dekirudake kokorogakeru. 4E5I3U4A5E#5050305A5E3U#
11. Minnano egaowa 1I84A40#E5AQ02A#
12. menimienai sasaeda. 1E7I1IE4AI#4A4AE4A#
13. Kesawa rokujimaekara 5E4A2A#305U7I1AE5A3A#
14. suemusumeno 4UE1U4U1E40#
15. genkinakoega kikoeru. 5E85I4A5OE5A#5I50E3U#
16. So:ieba, 40:IE1A#
17. yo:chienno 7O:7IE84O#
18. ensokuda. E8405U4A#
19. Do:butsuen-e 40:1U4UE8#E#
20. tsuretette moraerurashi:. 4U3E4E194E#103AE3U3A7I:#
21. Kaeriwa 5AE3I2A#
22. watashiga ekimade 2A4A7I5A#E5I1A4E#
23. makaenideru. 1U5AE7I4E3U#
24. Sukeju:ruo 4U5E7U:3U0#
25. machigaenaiyo: 1A7I5AE4A17O:#
26. kiotsukakenareba. 5I04U5E4AE3E1A#
27. Asawa ueno musumeto A4U2A#UE40#1U4U1E40#
28. kodomomukeno 5040101U5E40#
29. eigaomini E15A01I7I#
30. machie dekakeru. 1A7IE#4E5A5E3U#
31. Benkyo:zukueo 1E857O:4U5UE0#
32. kaikaeru yokusokumo 5A15AE3U#7A5U4O5U10#
33. shikkari oboeteoko:. 7I95A3I#01E4E050:#
1. Kemuriwa chi:sana tsubuno
   5E1U3I2A#7I:4A4A#4U1U40#
2. atsumaridearu.
   A4U1A3I4EA3U#
3. Kemurino tsubuno
   5E1U3I4O#4U1U40#
4. hitotsuihitotsua
   5I4O4U5I4O4U2A#
5. ningenno medewa
   7I85E840#1E4E2A#
6. miwakerukotowa dekinai.
   1I2A5E3U50402A#4E5I4AI#
7. Keredomo,
   5E3E4010#
8. kenbikyo:o tsukaeba,
   5E8I157O:O#4U5AE1A#
9. sono keitaiga
   404O#5E1A4I5A#
10. hakkiri mieru.
    5A95I3I#1IE3U#
11. Entotsukara deru
    E84O4U3A#4E3U#
12. kuroi kemurio
    5U30I#5E1U3I0#
13. keiseishiteiru tsubuwa
    5IE4I7I4EI3U#4U1U2A#
14. kotaidearu.
    504AI4EA3U#
15. Ippo:
    I910:#
16. tabakono kemurio
    4A1A5O4O#5E1U3I0#
17. keiseishiteirunowa
    5EI4I7I4EI3U4O2A#
18. dochirakatoieba
    40713A5A40IE1A#
19. ekitaibo tsubuto
    5I4AI4O#4U1U40#
20. kagareareteiru.
    5A85EA3A#4E4I3U#
21. Yugemo mata
    7U5E10#1A4A#
22. kemurino nakamato ieru.
    5E1U3I4O#4A5A1A4O#IE3U#
23. Kemurito kenko:no
    5E1U3I4O#5E85O:40#
24. kankeini tsuiteno
    5A85E1I7I#4U14E40#
25. jikkenya kenkyu:mo
    7I95E47A#5E857U:10#
26. fueteiru.
    6UE4E13U#
27. Kenko:ni
    5E85O:7I#
28. warui eikyo:o
    2A3UI#E157O:O#
29. ataeru kemurimonakade,
    A4AE3U#5E1U3I4O4A5A4E#
30. mazu agerarerunoga
    1A4U#5E3A3E3U4O5A#
31. tabakodearu.
    4A1A5O4EA3U#
32. Kin-enga
    5I8#E85A#
33. sakebareru gendaiwa
    4A5E1A3E3U#5E84AI2A#
34. ken-enken toyu:
    5E8#E85E8#4O7U:#
35. kotobasei umareta.
    50401A4AE#U1A3E4A#
36. Seiketsuna ku:kio eru
    4E15E4U4A#5U:5IO#E3U#
37. kenrinitsuiewa
    5E83I7I4I4E2A#
38. samazamana ikenga
    4A1A4A1A4A#I5E85A#
39. aruyo:daga,
    A3U7O:4A5A#
40. kono kenriwa
    5O4O#5E83I2A#
41. kesshite
    5E97I4E#
42. mushishitewa ikenai.
    1U71I4E2A#I5E4AI#
43. Nipponmo
    7I910810#
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>44.</td>
<td>dekirudake hayaku</td>
<td>48.</td>
</tr>
<tr>
<td>45.</td>
<td>kitsuenmo</td>
<td></td>
</tr>
<tr>
<td>46.</td>
<td>kin-enmo</td>
<td></td>
</tr>
<tr>
<td>47.</td>
<td>eraberu shakaini</td>
<td></td>
</tr>
<tr>
<td>48.</td>
<td>kae te ikanakerebananai.</td>
<td></td>
</tr>
</tbody>
</table>
LESSON 8

Review

Visual Only Presentation:

\[
\begin{array}{ll}
aa & pa \\
ia & wa \\
ua & \\
oa & \\
ea & \\
\end{array}
\]

Training

Visual Only Presentation:

\[
\begin{array}{ll}
aa & pa \\
aa & sha \\
aa & ra \\
aa & wa \\
aa & pya \\
aa & ia \\
aa & kya \\
aa & ma \\
aa & nya \\
aa & ha \\
aa & hya \\
aa & gya \\
aa & ka \\
aa & ja \\
aa & bya \\
aa & fa \\
aa & mya \\
aa & tsa \\
aa & rya \\
aa & za \\
aa & ba \\
aa & cha \\
aa & da \\
aa & ga \\
aa & sa \\
aa & ya \\
aa & na \\
\end{array}
\]

Audio-visual Presentation:

\[
ra
\]
KAZENO YOBO: (HOW NOT TO CATCH A COLD)

1. **Daredemo kenko:**niwa
   4A3E4E10#5E85O7I2A#
2. **sukunakarazu**
   4U5U4A5A3A4U#
3. **kanshin-o motteiru.**
   5A87I8#O1094EI3U#
4. **Nenreiga**
   4E83E15A#
5. **takakunarunitsure,**
   4A5A5U43U7I4U3E#
6. **henkaga arawareru.**
   5E85A5A3A2A3E3U#
7. **Korekarawa masumasu**
   503E5A3A2A1A4U1A4U#
8. **ko:re:kaga susumuto**
   50:3E5A5A4U4U1U4O#
9. **yoso:sar eru.**
   704O:4A3E3U#
10. **Soreni tom onai**
   403E7I40104AI#
11. **do:shitar a yamaini**
   40:7I4A3A7A1AI7I#
12. **kakarazunisumukari**
   5A5A3A4U7I4U1U5A7I#
13. **ju:tenga okarerudaro:**
   7U:4E85A05A3E3U4AO:
14. **Tokorode,**
   4050304E#
15. **arayuru byo:okinakade,**
   A3A7U3U170:5I404A5A4E#
16. **kazewa dochirakato yu:to**
   5A4E2A407I3A5A4O7U40#
17. **karuku miraregachida.**
   5A3U5U1I3A5E5A7I4A#
18. **Ker edomo,**
   5E3E4O10#
19. **korewa mukashikara**
   503E5U3AI4O#
20. **manby:onomoto toiware,**
   1A817:0401O404O12A3E#
21. **mareni taibyo:ni**
   1A3E7I4AI170:7I#
22. **tsunagaranaitomo i:kirenai.**
   4U4A5A3A4AI4010#5I3E4AI#
23. **Kazeni kakattara,**
   5A4E7I5A5A94A3A#
24. **korekuraito**
   503E5U3AI4O#
25. **hottarakashini shinaikotoda.**
   5094A3A5A7I7I5I4AI5040A#
26. **Yobo:nimo fudankara**
   7010:7I1O6U4A85A3A#
27. **chu:shinakerebanaranai.**
   7U:7I4A5E3E1A4A3A4AI#
28. **Sotokara kaet tara**
   40405A3A5AE94A3A#
29. **ugaito reiko:shi,**
   U5A1O#3EI5O:7I#
30. **kanarazu teo**
   5A4A3A4U44E0#
31. **kireini arau.**
   5I3E17I#A3AU#
32. **Soreigainimo,**
   403E15AI711O#
33. **eiyo:no baransuno**
   EI70:40#1A3A84U4O#
34. **yokutor ethos shoku jio**
   705U403E4A705U71O#
35. **toranakerebanaranai.**
   403A4A5E3E1A4A3A4AI#
36. **Izureniseyo,**
   I4U3E14I7E0#
37. **byo:genkinni furetemo**
   170:5E8518I#6U3E4E1O#
38. **soreni makenai chikarao**
   403E7I1A5E4AI7I5A3AO#
39. **higorokara tsuket e okukotoda.**
   5I5O3O5A3A4U5E4E05U5O4O4A#
LESSON 9

Review

Visual Only Presentation:

\[
\begin{align*}
aa & \quad pa \\
\text{ia} & \quad \text{wa} \\
\text{ua} & \quad \text{ra} \\
oa & \\
e a &
\end{align*}
\]

Training

Visual Only Presentation:

\[
\begin{align*}
\text{ta} & \quad \text{ja} \\
\text{ta} & \quad \text{fa} \\
\text{ta} & \quad \text{wa} \\
\text{ta} & \quad \text{rya} \\
\text{ta} & \quad \text{ra} \\
\text{ta} & \quad \text{mya} \\
\text{ta} & \quad \text{hya} \\
\text{ta} & \quad \text{nya} \\
\text{ta} & \quad \text{ha} \\
\text{ta} & \quad \text{na} \\
\text{ta} & \quad \text{gya} \\
\text{ta} & \quad \text{hya} \\
\text{ta} & \quad \text{bya} \\
\text{ta} & \quad \text{tsa} \\
\text{ta} & \quad \text{ya} \\
\text{ta} & \quad \text{ka} \\
\text{ta} & \quad \text{pya} \\
\text{ta} & \quad \text{sha} \\
\text{ta} & \quad \text{ba} \\
\text{ta} & \quad \text{cha} \\
\text{ta} & \quad \text{ta} \\
\text{ta} & \quad \text{da} \\
\text{ta} & \quad \text{za} \\
\text{ta} & \quad \text{ga} \\
\text{ta} & \quad \text{pa} \\
\text{ta} & \quad \text{sa}
\end{align*}
\]

Audio-visual Presentation:

\[
\begin{align*}
a s a & \quad z a & \quad n a \\
t a & \quad t s a & \quad d a
\end{align*}
\]
KYU:YU: (AN OLD FRIEND)

1. Ototoi, 040401#
2. daigakujidaino 4A15A5U7I4AI40#
3. do:kyu:seiga 40:57U:4E15A#
4. Sendaikara 4E84A15A3A#
5. tazunetekitekureta. 4A4U4E5I4E5U3E4A#
6. Tanakatyu: namaeda. 4A4A5A4O7U:#4A1AE4A#
7. Nakanakano shu:saide 4A5A4A5A40#7U:4A1E4#
8. medatsu sonzaidatta. 1E4A4U#4084AI4A94A#
9. Tamatama shigotode 4A1A4A1A#7I5O404E#
10. To:kyo:ni 40:570:7I#
11. detekitatsuinden 4E4E5I4A4UI4E7I#
12. tachiyottanodatoyu:. 4A7I7O94A40A4O7U:#
13. Totsuzendattanode, 4O4U4E84A9A4O4E#
14. isasaka awateteshimatta. I4A4A5A#A2A4E4E7I1A94A#
15. Semete hitokoto 4E1E4E#5I4O5O4O#
16. tegamika denwade 4E5A1I5A#4E82A4E#
17. jizenni 7I4E87I#
18. shirasetekureteitara 7I3A4E4E5U3E4E1A3A#
19. motenashimodekitanoni 1O4E4A7I1O4E5I4A4O7I#
20. zannende naranakatta. 4A84E84E#4A3A4A5A94A#
21. Do:yara watashio 4O:7A3A#2A4A7IO#
22. odorokasetakatarashi:. 0403O5A4E4A5A94A3A7I:#
23. Tanakano 4A4A5A40#
24. takuramukotowa 4A5U3A1U5O402A#
25. daitai so:zo:gatsuku. 4AI4A1#4O:40:5A4U5U#
26. Sorenishitemo 4O3E7I7I4E1O#
27. hisabisano saikaida. 5I4A1I4A40#4A15AI4A#
28. Nanto 4A84O#
29. ni:j:sannenburinaru. 7I7U:4A84E81U3I7I4A3U#
30. Otagai toshio 04A5AI#4O7I0#
31.ottomanoda. 4O94A1O4O4A#
32. Ochitsuite mierunomo 07I4UI4E#1I3U4O1O0#
33. to:zendaro:. 40:4E84A30:#
34. Tadashi, 4A4A7I#
35. futaritomo madamada 6U4A3I4O1O1#A4A1A4A#
36. hatarakizakari. 5A4A3A5I4A5A3I#
37. Toshiyoriatsukaidakewa 4O7I7O3IA4U5AI4A5E2A#
38. zettai saretakunai. 4E94AI#4A3E4A5U4AI#
1. Hichigatsunanokani
2. umaretanowa
3. tashika anatano
4. oto:tosandeshitane.
5. Watashiga
6. tanjo:bio tadashikute
7. oboe teiruhitowa
8. hotondo inainodesuga,
9. gu:zen,
10. watashinoaneto onaji
11. seinengappidattanode
12. kiokushite irunodesu.
13. Hitogotonagara,
14. tanabatano
15. umaretowa
16. nakanaka
17. fuzeigaatte
18. i:to omoimasenka?
19. Tadashi soremo,
20. maitoshi tsuyuno
21. massaichu:dewa
22. nantonaku
23. monotarinasa
24. kanjite shimaimasune.
25. Tanjo:bio
26. motsuimiwa
27. hitosorezore
28. kotonarudesho:shi,
29. iwaikatamo
30. samazamadesho:
31. Demo,
32. darenitottemo
33. tokubetsunahide
34. arukotowa tashikadesu.
35. Otonaninaruto
36. hitosu hitotsu
37. tosho torutabini,
38. jinseiitowa nanikato
39. kangaesareraemasune.
40. Tokoro de,
41. anatano seinengappio
42. oshieteitadakemasenka?
LESSON 10

Review

Visual Only Presentation:

\[
\begin{align*}
aa & \quad pa \\
iu & \quad ia \\
uu & \quad ua \\
oa & \quad oa \\
eu & \quad ea
\end{align*}
\]

Training

Visual Only Presentation:

\[
\begin{align*}
ka & \quad pya \\
ka & \quad ya \\
ka & \quad sha \\
ka & \quad ra \\
ka & \quad wa \\
ka & \quad na \\
ka & \quad ta \\
ka & \quad kya \\
ka & \quad ha \\
ka & \quad ma \\
ka & \quad nya \\
ka & \quad hya \\
ka & \quad ga \\
ka & \quad ja \\
ka & \quad bya \\
ka & \quad fa \\
ka & \quad mya \\
ka & \quad ka \\
ka & \quad gya \\
ka & \quad za \\
ka & \quad ba \\
ka & \quad cha \\
ka & \quad pa \\
ka & \quad rya \\
ka & \quad da \\
ka & \quad tsa \\
ka & \quad sa
\end{align*}
\]

Audio-visual Presentation:

\[
ka \\
ga \\
ha
\]
1. ako hachinenkanni 5A50#5A7I4E85A87I#
2. kyu:kao gaikokude 57U:5AO#5AI505U4E4#
3. sugosuhiitoga 4U504U5I405A#
4. zo:kashimashita. 40:5A7I1A7I4A#
5. Seikakuna kazuwa 4E15A5U4A#5A4U2A#
6. wakarimasenga, 2A5A3I1A4E85A#
7. kazokudedekakeru ke:suga 5A40S50E4E5A5E3U#:5E4U5A#
8. kekko: aruyo:desu. 5E950:#A3U70:4E4U#
9. Kaigairyoko:wa, 5A5AI37050:2A#
10. mukashiwa 1U5A7I2A#
11. kanari okanega 5A4A3I#05A4E5A#
12. kakarumono 5A5A3U04040#
13. kangaerareteimashitaga, 5A85AE3A3E4I1A7I4A5A#
14. konogorowa 504050302A#
15. endakano okagede E8A5A4O#05A5E4E#
16. kanarazushimo 5A4A3A4U7I10#
17. takakuitsukanaku narimashita. 4A5A5U4U5A45U#4A3I1A7I4A#
18. Seikatsuga 4E15A4U5A#
19. yutakaninari, 7U4A5A7I4A3I#
20. yokanotsuakitagata 705A4O#4U5AI5A5A#
21. henkashitakotomo 5E85A7I4A5O4010#
22. kankeishitemasu. 5A85E17I4E1IU4U#
23. Gakuseiwa 5A504E12A#
24. nagaiyasumiga 4A5AI7A4U1I5A#
25. arimasukara A3I1A4U5A3A#
26. arubaitode A3U1AI404E#
27. okaneotame 05A4EO4A1E#
28. nakamato kigaruni 4A5A1A4O#5I5A3U7I#
29. kakuchio 5A5U710#
30. kanko:surukotoga 5A850:4U3U0405A#
31. kano:desu. 5A40:4E4U#
32. Tokoroga, 405O305A#
33. shigotogaaruto 7I5O405AA3U4O#
34. nagakayasumu 4A5A5U7A4U1U#
35. wakenwa ikimasen. 2A5E7I2A#I5I1A4E8#
36. Kekkyoku, 5E95705U#
37. gokatsuno renkyu:nadoni 505A4U4O#3E857U:4A4O7I#
38. kagiraretshimaimasu. 5A5I3A3E4E7I1AI1A4U#
39. Amerikaga A1E3I5A5A#
40. ninkikatakaku, 7I85I5A445A5U#
41. nakademo nishikaiganno 4A5A4E1O#7I7I5AI5A840#
42. kariforuniaga 5A3I6O3U7I5A#
43. kanko:no mekkadesu. 5A850:40#1E95A4E4U#
44. Sonohokani,
45. chikakuno kankokuemo
46. kanshinga
47. mukerareteimasu.
1. Higoro
2. nanigenaku suetei
3. gominitsuit
4. kanshin-o mottakotoga
5. koremadeni arimasu?
6. Dokode
7. ikanaru ketroohete
8. shoriga
9. okonawarete irunodeshoka.
10. Gominawa
11. daidokorokaraderu
12. namagominohokani
13. reizo:koo hajimetosuru
14. o:gata, kogatano
15. denkaseihin,
16. kagunadoga fukumaresu.
17. Kakuchide
18. kaishu:sareri goto
19. hobo gojuppa: sentoga
20. kanengomide,
21. korerawa
22. kindaikasareta seiso: ko:jo: ni
23. hakobikomaresu.
24. Gominonakano tetsubunwa
25. jishakuo tsukatte
26. kaishu: sare,
27. shigeno sairiyo: kaga
28. hakarareteimasu.
29. Chikagorowa
30. shinkaihatsusareta
31. sozaiga ichihayaku
32. kateiseikatsuni hairikomi,
33. atarashi: ko: gaito
34. narunodewanaikati
35. kenensareteimasu.
36. Konoyo: ni,
37. gominikanshi: temo,
38. kaiketsushinakerebanaranai
39. kadaiga
40. kanari arunodesu.
LESSON 11

Review

Visual Only Presentation:

\[
\begin{array}{ll}
aa & pa \\
 ia & wa \\
 ua & ra \\
oa & ta \\
e a & ka \\
\end{array}
\]

Training

Visual Only Presentation:

There is no visual only presentation for this lesson.

Audio-visual Presentation:

\[
\begin{array}{llllll}
[\text{wa}] & [\text{wi}] & [\text{we}] & [\text{wo}] \\
\end{array}
\]
FURANKU-SAN (FRANK)

1. Furankusanwa
2. Furoridaumare.
3. Sofutoweano kaishani
4. tsuometeiru.
5. Itsumo po:ka:feisude
6. go:ingumaiwei.
7. Ha:dowa:ka:de,
8. faitono katamari.
10. Wanmandakeredomo
11. feminisutodemoaru.
12. Kaiwawa wittonitomi,
13. fo:marunabadem
14. info:marunabademo ninkimono.
15. Maiasa tore:ninguweakite
16. wo:kuman-o kikinagara
17. jogenuosuru.
18. Supo:tsude
19. we:tokontoro:ru.
20. Asagohanwa kafeoreto
21. kurowassan.
22. Wi:kuendowa
23. ba:dowocchinguo tanoshimu.
24. Sukinaosakewa
25. wiski:to wokka.
26. Omizuwa kanarazu
27. mineraruwo:ta:
LESSON 12

Review

Visual Only Presentation:

<table>
<thead>
<tr>
<th>aa</th>
<th>pa</th>
</tr>
</thead>
<tbody>
<tr>
<td>ia</td>
<td>wa</td>
</tr>
<tr>
<td>ua</td>
<td>ra</td>
</tr>
<tr>
<td>oa</td>
<td>ta</td>
</tr>
<tr>
<td>ea</td>
<td>ka</td>
</tr>
</tbody>
</table>

Training

Visual Only Presentation:

NOTE: f this lesson is [ f ] not [ φ ]

<table>
<thead>
<tr>
<th>fa</th>
<th>cha</th>
</tr>
</thead>
<tbody>
<tr>
<td>fa</td>
<td>wa</td>
</tr>
<tr>
<td>fa</td>
<td>ya</td>
</tr>
<tr>
<td>fa</td>
<td>ra</td>
</tr>
<tr>
<td>fa</td>
<td>na</td>
</tr>
<tr>
<td>fa</td>
<td>pya</td>
</tr>
<tr>
<td>fa</td>
<td>ta</td>
</tr>
<tr>
<td>fa</td>
<td>kya</td>
</tr>
<tr>
<td>fa</td>
<td>ma</td>
</tr>
<tr>
<td>fa</td>
<td>nya</td>
</tr>
<tr>
<td>fa</td>
<td>ba</td>
</tr>
<tr>
<td>fa</td>
<td>hya</td>
</tr>
<tr>
<td>fa</td>
<td>ga</td>
</tr>
<tr>
<td>fa</td>
<td>ja</td>
</tr>
<tr>
<td>fa</td>
<td>bya</td>
</tr>
<tr>
<td>fa</td>
<td>fa</td>
</tr>
<tr>
<td>fa</td>
<td>mya</td>
</tr>
<tr>
<td>fa</td>
<td>ka</td>
</tr>
<tr>
<td>fa</td>
<td>gya</td>
</tr>
<tr>
<td>fa</td>
<td>za</td>
</tr>
<tr>
<td>fa</td>
<td>sha</td>
</tr>
<tr>
<td>fa</td>
<td>sa</td>
</tr>
<tr>
<td>fa</td>
<td>tsa</td>
</tr>
<tr>
<td>fa</td>
<td>rya</td>
</tr>
<tr>
<td>fa</td>
<td>da</td>
</tr>
<tr>
<td>fa</td>
<td>pa</td>
</tr>
<tr>
<td>fa</td>
<td>cha</td>
</tr>
</tbody>
</table>

Audio-visual Presentation:

fa
KANATANAKOTOBANO HANRAN (A FLOOD OF LOANWORDS)

1. Gendaino fasshonaburuna
2. raijusutanritowa
3. katakanakotobao
4. fi:ringude tsukakotorashi:
5. Ofisuuden oshigotowa
6. wa:puroto fakkusu.
7. Afuta:faibuu
8. fittonesukurabu.
9. Dina:wa furansuryo:ri
10. furuko:su.
11. Furu:too kikinagara
12. naifuto fo:kude
13. foagurao shishoku.
14. Shu:matsuwa gorufuya
15. sa:finde
16. rifuresshuu.
17. Amerikanfutoboruwata terebikansen.
18. Yakyu:no fannara
19. fainpure:ni
21. Machidewa furawa:sho:,
22. fotokontesutosu,
23. furu:tsufe,
24. firumufesutibarugamoyo:sare,
25. kokusaif:ramumo hirakareru.
26. Hatashite korewa
27. kokusaikananodaro:ka.
LESSON 13

Review

Visual Only Presentation:

\[
\begin{align*}
aa & \quad pa \\
ia & \quad wa \\
uu & \quad ra \\
oa & \quad ta \\
ea & \quad ka \\
ea & \quad fa \\
\end{align*}
\]

Training

Visual Only Presentation:

\[
\begin{align*}
ya & \quad na \\
nya & \quad ya \\
ya & \quad sha \\
ya & \quad ra \\
ya & \quad wa \\
ya & \quad pya \\
ya & \quad ta \\
ya & \quad kya \\
ya & \quad ma \\
ya & \quad nya \\
ya & \quad fa \\
ya & \quad hya \\
ya & \quad ga \\
ya & \quad ja \\
ya & \quad bya \\
ya & \quad ha \\
ya & \quad mya \\
ya & \quad ka \\
ya & \quad gya \\
ya & \quad za \\
ya & \quad ba \\
ya & \quad cha \\
ya & \quad sa \\
ya & \quad rya \\
ya & \quad da \\
ya & \quad pa \\
ya & \quad tsa \\
\end{align*}
\]

Audio-visual Presentation:

\[
\begin{align*}
ya & \quad pya \\
sha & \quad cha \\
hya & \quad rya \\
bya & \quad ja \\
mya & \quad nya \\
gya & \quad gya \\
\end{align*}
\]
TONARINO OYAKO (MY NEIGHBOR)

1. Kyabetsuno
2. yasaisaradao tsukuri,
3. gyo:zao
4. yakinagara kangaeta.
5. Otonarino oyakowa
6. yasuminiwa
7. go:kakyakusenno
8. tabio surutoyu:
10. Kyanpuni ikunoga
11. sekinoyamano
12. wagayatowa
13. o:kina gyappuda.
14. Kyanserumachio shite
15. yatto jitsugenshitasoda.
16. Yahari nenpaino
17. okyakubakari nanodaro:ka.
18. Yuttarishita kyabinde
19. yasumunowa
20. sazo kaitekidaro:
21. Oyaomoino
22. yasashi: musukosanda.
23. Gyakuni
24. yasashisugiru tomoieru.
25. Wagayaniwa
26. kyaibao omiyageni
27. kattekitekeruru
28. yakuokuo
29. chanto shitekureta.
30. Shashinmo
31. hyakumaikurai
32. tottekuryamyo shirenu.
33. Ichinichimo hayaku
34. shuppatsushitaine chigainai.
LESSON 14

Review

Visual Only Presentation:

\[
\begin{array}{ll}
aa & pa \\
i a & wa \\
u a & ra \\
o a & ta \\
e a & ya \\
f a & ka \\
\end{array}
\]

Training

Visual Only Presentation:

\[
\begin{array}{ll}
anna & anna \\
anna & a:na \\
anna & ana \\
anpa & apa \\
anpa & anpa \\
anpa & a:pa \\
anta & ata \\
anta & a:ta \\
anta & anta \\
anka & aka \\
anka & anka \\
anka & a:ka \\
an-a & a: \\
an-a & ana \\
an-a & an-a \\
\end{array}
\]

Audio-visual Presentation:

\[
\begin{array}{ll}
anpa & anwa \\
anra & anta \\
anka & anfa \\
anya & an-a \\
\end{array}
\]
MOTTO ARUKO: (LET'S WALK)

1. Bunmeino hattatsuni tomonai
2. Jinruino undo:ryorwa
4. Undo:busokuwa
5. Himanya seijinbyo:ni
6. Tsunagaru shinpaimo aru.
7. Ippanni
8. Gendaijinwa
9. Kenko:o ishikishite,
10. Jibunkara susunde
11. Undo:ni hagendeiru.
12. Sararinanno nakaniwa
13. Suimingukurabuya
14. Tore:ningukurabuno
15. Kaininnaru
16. Hitomo ireba
17. Man-Indenshao sake
18. Jitenshade
20. Mochiron, kojinno
21. Seikatsukanky:ya nenrei
23. Sentakuga kanjinda.
24. Saikinwa
25. Wo:kingu toyu:mononi
27. Arukukotowa
28. Ningenneno kenko:no
29. Kihonda
30. Donna hitodemo
31. Donna tokorodemo
32. Kantanni dekirushi,
33. Junbimo
34. Hotondo iranai.
35. Ashini
36. Futanno kakaranai
37. Kutsugakichinto
38. Erandeareba
40. Juppundemo
41. Ju:gojundemo
42. Jikan-o mitsukete
43. Shu:kanni shiyo:.
**LESSON 15**

**Review**

Visual Only Presentation:

- *anpa*
- *anwa*
- *anra*
- *anta*
- *anfa*
- *anka*
- *an-a*

**Training**

Visual Only Presentation:

<table>
<thead>
<tr>
<th>appa</th>
<th>apa</th>
</tr>
</thead>
<tbody>
<tr>
<td>appa</td>
<td>appa</td>
</tr>
<tr>
<td>appa</td>
<td>a:pa</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>atta</th>
<th>ata</th>
</tr>
</thead>
<tbody>
<tr>
<td>atta</td>
<td>a:ta</td>
</tr>
<tr>
<td>atta</td>
<td>atta</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>akka</th>
<th>aka</th>
</tr>
</thead>
<tbody>
<tr>
<td>akka</td>
<td>akka</td>
</tr>
</tbody>
</table>

Audio-visual Presentation:

<table>
<thead>
<tr>
<th>appa</th>
<th>assa</th>
<th>asska</th>
</tr>
</thead>
<tbody>
<tr>
<td>atta</td>
<td>adda</td>
<td>accha</td>
</tr>
<tr>
<td>akka</td>
<td>agga</td>
<td></td>
</tr>
</tbody>
</table>
### MACHINO HENKA (CHANGE OF MY HOMETOWN)

<table>
<thead>
<tr>
<th>No.</th>
<th>Sentence</th>
<th>Kana</th>
<th>Romaji</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Hisashiburini kaetta kokyo:wa</td>
<td>ひさしぶりに かいeta こきyo:wa</td>
<td>ひさしぶりに かいeta こきyo:wa</td>
</tr>
<tr>
<td>2.</td>
<td>sukkari kawattemishimatta.</td>
<td>すくかり かわたてしめiatatta.</td>
<td>すくかり かわたてしめiatatta.</td>
</tr>
<tr>
<td>3.</td>
<td>Ekini oritatte</td>
<td>えきini おったて</td>
<td>えきini おったて</td>
</tr>
<tr>
<td>4.</td>
<td>bikkurishitemimatta.</td>
<td>びっくりししitemimatta.</td>
<td>びっくりししitemimatta.</td>
</tr>
<tr>
<td>5.</td>
<td>Imamadetowa</td>
<td>いままだてowa</td>
<td>いままだてowa</td>
</tr>
<tr>
<td>6.</td>
<td>mottaku chigatta</td>
<td>もたくちがっatta</td>
<td>もたくちがっatta</td>
</tr>
<tr>
<td>8.</td>
<td>Isshun,</td>
<td>いしゅん,</td>
<td>いしゅん,</td>
</tr>
<tr>
<td>9.</td>
<td>machigatte resshao</td>
<td>まちがって れっしょう</td>
<td>まちがって れっしょう</td>
</tr>
<tr>
<td>10.</td>
<td>oritesshimattanodewanaikato</td>
<td>おりてしめまったなodeないくた</td>
<td>おりてしめまったなodeないくた</td>
</tr>
<tr>
<td>11.</td>
<td>sakkakushitahododa.</td>
<td>さっかくしとたほどだ.</td>
<td>さっかくしとたほどだ.</td>
</tr>
<tr>
<td>12.</td>
<td>Chotto Urashimataroni</td>
<td>ちょっと うらしまたに</td>
<td>ちょっと うらしまたに</td>
</tr>
<tr>
<td>14.</td>
<td>Machiwa shoppinguo suruhitode</td>
<td>まちわ しょうひんぎょうゆう するひとで</td>
<td>まちわ しょうひんぎょうゆう するひとで</td>
</tr>
<tr>
<td>15.</td>
<td>gottagaheiteiru.</td>
<td>ごたとがへっている.</td>
<td>ごたとがへっている.</td>
</tr>
<tr>
<td>16.</td>
<td>Itai dokokara</td>
<td>いたい どこか-</td>
<td>いたい どこか-</td>
</tr>
<tr>
<td>17.</td>
<td>konnippai hitoga</td>
<td>こんねっぴがい ひとが</td>
<td>こんねっぴがい ひとが</td>
</tr>
<tr>
<td>18.</td>
<td>atsumatte kurunorodoka.</td>
<td>あつまって くろんのどか.</td>
<td>あつまって くろんのどか.</td>
</tr>
<tr>
<td>19.</td>
<td>Maeniwa tatta mitsushika</td>
<td>まえにわ たった みつしんか</td>
<td>まえにわ たった みつしんか</td>
</tr>
<tr>
<td>20.</td>
<td>nakattabiru</td>
<td>なかって びる</td>
<td>なかって びる</td>
</tr>
<tr>
<td>21.</td>
<td>imawa acchikocchini</td>
<td>いわ あちっこちに</td>
<td>いわ あちっこちに</td>
</tr>
<tr>
<td>22.</td>
<td>bisshirito</td>
<td>びしりと</td>
<td>びしりと</td>
</tr>
<tr>
<td>23.</td>
<td>tatteiru.</td>
<td>たてている.</td>
<td>たてている.</td>
</tr>
<tr>
<td>24.</td>
<td>Omottemomakakatta henkani</td>
<td>おもてтемおまかた へんかん</td>
<td>おもてтемおまかた へんかん</td>
</tr>
<tr>
<td>25.</td>
<td>hakkiritte</td>
<td>はかきりって</td>
<td>はかきりって</td>
</tr>
<tr>
<td>26.</td>
<td>gakkari.</td>
<td>がっかり.</td>
<td>がっかり.</td>
</tr>
<tr>
<td>27.</td>
<td>Mukashinomamao</td>
<td>むかしのままで</td>
<td>むかしのままで</td>
</tr>
<tr>
<td>28.</td>
<td>tamotteite</td>
<td>たまっていて</td>
<td>たまっていて</td>
</tr>
<tr>
<td>29.</td>
<td>hoshikattanoni.</td>
<td>ほしけたたなに</td>
<td>ほしけたたなに</td>
</tr>
<tr>
<td>30.</td>
<td>Mottomo,</td>
<td>もっとも,</td>
<td>もっとも,</td>
</tr>
<tr>
<td>31.</td>
<td>subetega zutto</td>
<td>すべてが ずっと</td>
<td>すべてが ずっと</td>
</tr>
<tr>
<td>32.</td>
<td>onaji:to:taide</td>
<td>おなじ:と:たide</td>
<td>おなじ:と:たide</td>
</tr>
<tr>
<td>33.</td>
<td>nokotteiruyou:dewa</td>
<td>のくっているゆう:でわ</td>
<td>のくっているゆう:でわ</td>
</tr>
<tr>
<td>34.</td>
<td>issai hattenga</td>
<td>いかい はてんが</td>
<td>いかい はてんが</td>
</tr>
<tr>
<td>35.</td>
<td>naikotoninate,</td>
<td>ないるときおにた,</td>
<td>ないるときおにた,</td>
</tr>
<tr>
<td>36.</td>
<td>kaette</td>
<td>かえて</td>
<td>かえて</td>
</tr>
<tr>
<td>37.</td>
<td>komatesshimaunodaro:.</td>
<td>こまてししまうどろ:</td>
<td>こまてししまうどろ:</td>
</tr>
<tr>
<td>38.</td>
<td>Ningendatte i:shoda.</td>
<td>ねいがだて いし:どや</td>
<td>ねいがだて いし:どや</td>
</tr>
<tr>
<td>39.</td>
<td>Henkaga attekoso,</td>
<td>へんかが あてっこそ,</td>
<td>へんかが あてっこそ,</td>
</tr>
<tr>
<td>40.</td>
<td>rippana seicho:nanoda.</td>
<td>ริปปันา ｾｲﾁｮ:ﾅﾝﾀﾞ</td>
<td>ริปปันา ｾｲﾁｮ:ﾅﾝﾀ</td>
</tr>
</tbody>
</table>
LESSON 16

Review

Visual Only Presentation:

anpa   appa
anwa   atta
anra   akka
anta
anfa
anka
an-a

Training

Visual Only Presentation:

a:i     a:i
a:i     ai
a:i     ai
a:i     a:i
a:i     a:i
a:i     ai
a:i     a:i
a:i     ai

Audio-visual Presentation:

a:
i:
u:
e:
o:
NIHONNO KYO:IKUSE:DO (EDUCATIONAL SYSTEM)

1. Nihondewa    7I5084E2A#
2. sho:gakko: rokunen, 70:5A950:#305U4E8#
3. chu:gakko: sannen, 7U:5A950:#4A84E8#
4. go:ke: kyu:nenga 50:5E:#57U:4E85A#
5. imukyo:ikudesu. 5I1U570:I5U4E4U#
6. Kono ho:ritsuwa 5040#50:3I4U2A#
7. sengon se:te:saremashita. 4E8507I#4E:4E:4A3E1A7I4A#
8. jugyo:ryo:wa 7U570:370:2A#
9. to:zen muryo:desu. 40:4E8#1U370:4E4U#
10. Korewa ky:o:tu:no 50:3I4U40#5A950:7I#
11. ny:shino shinpaiosuru 47U:7I40#7I81AI04U3U#
12. Watashiwa e:go:so 5I4U70:2A#A3I1A4E8#
13. jo:zunidekizu, 70:4U7I4E5I4U#
14. na:ma:se:sekideshita. 1A:1A:40#4E:4E5I4E7I4A#
15. ko:shita ke:kenga arimasu. 5US30:7I4A#5E:5E85A#A3I1A4U#
16. Demo su:ga:kuto 4E10#4U:5A5U40#
17. shissu se:seyomasha 5U30:4U3U#4E:4010#
18. sukoshidemo yume:na 4U507I4E10#7U:1E:4A#
19. daigakushi hairo:to 4A15A5U7I#5AI30:40#
20. jukensenso:no kaikakuto yu:keredo 7I947U:5A105U4E4U#
21. koko:shingakuwa 7U5E84E4A5U:
22. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
23. koko:shingakuwa 7U5E84E4A5U:
24. jukensenso:no kaikakuto yu:keredo 7I947U:5A105U4E4U#
25. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
26. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
27. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
28. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
29. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
30. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
31. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
32. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
33. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
34. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
35. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
36. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
37. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
38. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
39. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
40. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
41. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
42. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
43. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
44. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
45. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
46. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
47. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
48. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
49. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
50. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
51. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
52. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
53. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
54. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
55. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
56. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
57. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
58. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
59. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
60. kai:se:kenme:desu. 7I507I4E10#7U:5I4E84E4A5U:
LESSON 17

General

WATASHINO SONKEISURU HITO (A PERSON I RESPECT)

1. Anatawa  A4A4A2A#
2. rekishijo:no  3E5I7I7O:40#
3. jinbutsuno nakade,  7I81U4U40#4A5A4E#
4. dareo ichiban,  4A3E0#I7I1A8#
5. sonkeishimasukato kikaretara,  4085E17IIA4U5A4O#5I5A3E4A3A#
6. nanto kotaemasuka.  4A840#504AE1A4U5A#
7. Tokuteino  405U4EI4O#
8. hitono namaega,  5I4O40#4A1AE5A#
9. satto detekurudesho:ka.  4A940#4E4E5U3U4E70:5A#
10. Watashiwa,  2A4A7I2A#
11. konoshuno shitsumonwa  504O7U4O#7I4U1082A#
12. hijo:ni kiraide,  5I7O:7I5I3AI4E#
13. itsuno, sonoba, sonobano  I4U10#4O401A#4O401A4O#
14. omoitsukide  010I4U5I4E#
15. kaito:shiteimasu.  5AI40:7I4E1I4A4U#
16. Kekkyoku,  5E95705U#
17. mottomo sonkeisuru ningenga  1094O10#4O85E14U3U#7I8E85A#
18. nanninmo to:jo:surukotoni  4A87I81O#40:70:4U3U5O4O7I#
19. natteshimaundesune.  4A94E7I1A4U84E4U4E#
20. Tatoeba, aruhi,  4A4O4E1A#A3U5I#
21. tokugawaiaeyasu toittakatoomoeba,  405U5A2AIE7A4U#4O194A5A40010E1A#
22. yokujitsuwa  7O5U7I4U2A#
23. fukuzawayukichi toittafu:ni  6U5U4A2A7U5I7I#4O194A6U:7I#
24. mattaku betsuno jinmeio  1A94A5U#1E4U4O#7I81E1O#
25. ageteshimaawakedesu.  A5E4E7I1A4U2A5E4E4U#
26. Moshimo,  107I10#
27. onaji aitekara,  04A7I#AI4E5A3A#
28. do:itsuno shitsumon-o  4O:14U4O#7I4U108#:O#
29. nikai saredemoshitara,  7I5AI#4A3E4E1O7I4A3A#
30. taihenakotoni narimasuyo.  4AI5E84A5O407I#4A3I1A4U7O#
31. Naizenara,  4A4E4A3A#
32. jibunno  7I1U84O#
33. saishono kotaeo  4AI7O4O#504AE0#
34. oboete  010E4E#
35. irareinodesukara.  I3A3E4AI4O4E4U5A3A#
1. Anatawa
2. sekaiju:no
3. tabemonono nakade,
4. naniga ichiban,
5. sukidesukato kikaretara,
6. satto kaito:dekimasuka.
7. Watashiwa,
8. konoyo:na shitsumonwa
9. hijo:ni nigatede,
10. itsumo shibaraku
11. kangaenaito
12. kotaega detekimasen.
13. Tokubetsuna
14. monono namaega,
15. suguni omoitsukanainodesuyo.
16. Kekkyoku,
17. korezotioeru tabemonoga
18. naniihitasu sonzaishinaikotoni
19. natesshimandesune.
20. Tadashi,
21. onaji imino shitsumonmo,
22. sentakuhan-io kagittara
23. tasho:wa rakuni
24. kotaerarerukamo shiremasen.
25. Tatoeba, kudamononara
26. ichigo toyakotogade kuirushi,
27. kunibetsudewa,
28. Mekishikoryo:ri toittaguaini
29. sukoshiwa hayaino:mo
30. fukane:dewa nainodesu.
31. Konoteidodeareba
32. jibunmo
33. hitotsuya futatsu
34. nanika agerukotoga
35. dekirudesho:
Lesson Script and Literal Translation of Lesson 17 and Lesson 18

Lesson 17

General

WATASHINO SONKEISURU HITO


A PERSON I RESPECT

What would you say, when you are asked who is the person in history you respect most? Do you come up with a particular name immediately? Since I hate this sort of question, I always name one person randomly every time I am asked. Ultimately, it may mean that there exists no single person I really respect. For example, one time, my answer will be "Tokugawa Ieyasu", and the next day, it can be "Fukuzawa Yukichi", who are totally different characters. If the same interviewer asks me the same question twice, I will be in a big trouble, because I can never remember my first answer.
Lesson 18

General

WATASHINO SUKINA TABEMONO


MY FAVOURITE FOOD

WHEN YOU ARE ASKED WHAT YOUR MOST FAVOURITE FOOD IS, CAN YOU ANSWER QUICKLY? SINCE I AM NOT GOOD AT RESPONDING TO THIS KIND OF QUESTION, I ALWAYS HAVE TO THINK A WHILE. NO PARTICULAR NAME OF FOOD COMES UP IMMEDIATELY. ULTIMATELY, IT MAY MEAN THAT THERE EXISTS NOTHING THAT I REALLY LOVE TO EAT. HOWEVER, THE SITUATION CAN BE A LITTLE BIT DIFFERENT, IF THE INQUIRY IS MORE SPECIFIC IN TERMS OF CATEGORIES. FOR EXAMPLE, IT IS NOT IMPOSSIBLE FOR ME TO SAY THAT I LIKE STRAWBERRIES BEST AMONG FRUIT, OR THAT I PREFER ITALIAN CUISINE TO ANY OTHER ETHNIC CUISINE. WITH THESE QUESTIONS, I CAN PROBABLY NAME AT LEAST ONE OR TWO ITEMS.